DRAFT: PLEASE DO NOT POST ONLINE

Dynamic Inconsistency in Food Choice: Experimental Evidence from a Food Desert

Sally Sadoff*
University of California, San Diego

Anya Samek†
University of Wisconsin-Madison

Charles Sprenger‡
University of California, San Diego

September, 2014
This Version: January 31, 2015

Abstract

Despite the great deal of research on dynamic inconsistency in time preferences, few studies have ventured into investigating the question in a natural context. To address this gap, we conduct a natural field experiment with over 200 customers at a grocery store to investigate dynamic inconsistency and the demand for commitment in food choice. Over a 3 week time period, subjects are invited to allocate and re-allocate food items received as part of a grocery delivery program. We observe substantial dynamic inconsistency in our experiment, as well as a demand for commitment among a non-negligible number of subjects. Interestingly, individuals who demand commitment are more likely to be dynamically *consistent* in their prior behavior. For academics, our work provides direct evidence of dynamic inconsistency in consumption choices in the field and points towards potential extensions to models of temptation. For policy-makers, our findings provide insights on innovations to alter food choices.

JEL classifications: C91, D12, D81

Keywords: temptation, dynamic inconsistency, commitment demand, field experiment

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

†University of Wisconsin-Madison, School of Human Ecology, 1300 Linden Drive, Madison WI 53706; asamek@wisc.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 (Laibson, 1997; O'Donoghue and Rabin, 1999a; 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, despite stated desires to make these changes.

While a large literature has developed investigating hallmarks of dynamic inconsistency in laboratory experiments, (see Frederick, Loewenstein and O'Donoghue (2002) for an overview), limited empirical evidence exists to inform these models in a natural context.[1] To address this gap, we conduct a natural field experiment to investigate dynamic inconsistency and the demand for commitment in food choice.[2]

Our experimental setting is a food delivery service for low-income subjects in urban Chicago, Illinois. This setting is important for two reasons. First, our experimental neighborhood previously qualified as a 'food desert' according to the United States Department of Agriculture, implying both a high rate of poverty and potentially limited access to fresh fruits and vegetables.[3] The limited grocery access of our study sample is critical, as we offer subjects choices between healthy and unhealthy items. Limited access to healthy foods constrains the opportunities for subjects to arbitrage the experiment by making use of external trading opportunities. Second, our experimental

[1]Identification of dynamic inconsistency relies on subjects choosing between rewards at time t and at time $t+k$. Disagreement in these choices is evidence of dynamic inconsistency. Relatively few studies provide such longitudinal tests, as noted by Halevy (2012) and Sayman and Onculer (2009). Sayman and Onculer (2009) identifies only three prior longitudinal studies of time inconsistency: Ainslie and Haendel (1983) for monetary choices within subjects, Read, Loewenstein and Kalyanaraman (1999) for "highbrow" and "lowbrow" movies between subjects, and Read and Van Leeuwen (1998) for snack choices within subjects. Another early work is Solnick, Kannenberg, Eckerman and Waller (1980), for irritating noises between subjects. Recent experimental advances with longitudinal within subjects designs include Sayman and Onculer (2009) for cafe rewards and money, Halevy (2012) for money, Gine, Goldberg, Silverman and Yang (2010) for money, and Augenblick, Niederle and Sprenger (2013) for effort and money. A recent study with a longitudinal between subjects design is Duflo, Kremer and Robinson (2011) for fertilizer purchase. We are aware of only three field studies in which subjects potentially did not know they were participating in an experiment: Sayman and Onculer (2009), Read and Van Leeuwen (1998) and Duflo et al. (2011).

[2]Related work on food choice includes McClure, Laibson, Loewenstein and Cohen (2004), who find neural evidence for diminishing impatience surrounding beverage consumption and Brown, Chua and Camerer (2009) who show that food decisions (also consuming beverages) exhibit hallmarks of dynamic inconsistency in consumption-savings problems.

[3]See the USDA's website, http://apps.ams.usda.gov/fooddeserts/fooddeserts.aspx. A food desert is officially defined as having a) a poverty rate of 20% or greater and b) at least 500 persons and/or at least 33% of the census tract lives more than one mile from a supermarket or large grocery store. Estimates by the USDA suggest that approximately 11.5 million low-income Americans live more than 1 mile from a supermarket.

neighborhood sits in the cross-hairs of debates on food policy. Obesity is at an all-time high in the United States, contributing to chronic illnesses such as coronary disease and diabetes, and disproportionately affecting low income communities (Jia and Lubetkin, 2005; Mokdad, Ford, Bowman, Dietz, Vinicor, Bales and Marks, 2003; Kumanyika and Grier, 2006). Research in such communities may be uniquely positioned to influence the policy debate on the sources of low-income obesity and the interventions that may prove successful in combatting it.

Two hundred eighteen subjects completed a three-week food delivery program conducted in collaboration with a neighborhood grocery store, Louis' Groceries.[4] Subjects were given a budget of $10 and asked to construct a bundle of goods from a list of 20 potential foods for home delivery in one week. All items cost $1 each. Ten of the foods were fresh fruits and vegetables (which we call healthy) and ten were packaged snack foods (which we call unhealthy). Prior to the study, subjects also provided a rating for how much they liked each food.

Delivery was made at a time when subjects were sure to be home. On the day of delivery, the delivery-person brought the chosen bundle of goods, and also surprised subjects with four additional foods available for exchange. These four items were tailored for each subject to be their highest rated healthy item, their highest rated healthy item not included in their original chosen bundle, their highest rated unhealthy item, and their highest rated unhealthy item not included in their original chosen bundle. Subjects were given the opportunity to make up to four exchanges. Importantly, every bundle that could be constructed with immediate exchanges is one that was available at the time of advance choice. As such, dynamic inconsistencies are identified as violations of revealed preference between advance and immediate choices.

In the second week of the study, subjects again made advance choices. However, during the delivery reminder call, they were asked if they would like the driver to *not* bring the four additional items. That is, subjects were asked if they would like to commit to their advance choice. Thus, commitment is identified as a binding restriction to the singleton choice set consisting only of the advance bundle.

Our analysis focuses on dynamic inconsistencies that alter the proportion of healthy or unhealthy items in a subject's bundle. Forty-six of 218 subjects (21%) exhibited dynamic inconsistencies, making immediate choices that were either less healthy or

[4]Two hundred twenty five subjects were initially recruited into the study. Of these 225, 5 subjects did not provide a working phone number of address, so we consider them attrited from the sample before receiving any food deliveries. One subject requested their data not be used in analysis and one subject did not make a commitment decision. Our relatively low level of attrition is partly due to a completion bonus of $20. See section 2 for further details.

more healthy than their advance choices. Of these 46, 44 (96%) subjects exchanged for less healthy bundles, giving a clear direction for the nature of inconsistency. Consistent with models of temptation and self-control, subjects' immediate decisions are tilted towards less healthy alternatives than their advance decisions.

Of the 218 subjects, 73 (33%) chose to commit, restricting themselves to their advance choice.[5] Such demand for commitment is a key prediction of models of dynamic inconsistency. Recognizing the possibility of immediate temptation, a decisionmaker may be willing to restrict herself to her advance choice.

Exploration of the relationship between dynamic inconsistency and commitment reveals an interesting pattern. Committing subjects are disproportionately likely to be dynamically *consistent*. While only 9 of 46 (20%) dynamically inconsistent subjects demand commitment, 64 of 172 (37%) dynamically consistent subjects do so. Food ratings data indicate that committing subjects do indeed try to restrict themselves when making their advance decisions, choosing their highest rated items less frequently than their non-committing counterparts. Hence, it seems committing subjects are systematically those that overcome temptation, rather than those who succumb to it.

Our findings relate to two prominent strands in the experimental literature on dynamic inconsistency. The first strand attempts to utilize consumption rather than monetary choices to investigate intertemporal preferences. Most notably, Read and Van Leeuwen (1998) identify dynamic inconsistency in the surprise re-allocations of snack choices for Dutch workers. Across conditions, 50.5% of subjects chose unhealthy snack items in advance choice and 83% of subjects chose unhealthy snack items in immediate choice, 37.5% of subjects were dynamically inconsistent and 93% of violations were towards less healthy snack items. The patterns of dynamic inconsistency and violations towards less healthy items are consistent between our study and that of Read and Van Leeuwen (1998). However, several design differences might generate our clearly less dramatic findings. First, in Read and Van Leeuwen (1998) subjects are told that the researcher had no record of their prior decisions.[6] This might give the decisionmaker additional license to give in to temptation, or may generate exper-

[5]For the remaining 145 non-committing subjects, we can investigate stability of dynamic inconsistency, revealing substantial correlation in both choice and the inconsistency in choice over time. Section 4 provides details.

[6]The procedure notes "At this point, we made no reference to the fact that they had already chosen a snack and told them that they could have any snack they wanted. Many people felt that they should take the snack they had originally chosen, but we reminded them that they could have any snack they wanted and emphasized that we had plenty of each kind of snack. Some asked if we knew what their advance choices were, but we denied any such knowledge and 'proved' it by showing them a list containing only their names with no record of their choices. Once the subjects had made their immediate choice we gave them the snack they chose" (Read and Van Leeuwen, 1998).

imenter demand effects. A second difference is an issue with selection in Read and Van Leeuwen (1998), as 20% of the sample attrits during the study. Choices are made at two points in time, either 'after lunch time' or 'in the late afternoon around 4:30 or 5:00'.[7] Subjects present at work at those times may be selected on potential for temptatio, either by not having eaten lunch or not having left work for the day.

A second strand of literature links hallmarks of dynamic inconsistency to commitment demand. Such empirical exercises are instructive as they provide an opportunity to both validate models of dynamic inconsistency and discriminate between competing theories. In particular, under perfectly sophisticated quasi-hyperbolic discounting (Laibson, 1997; O'Donoghue and Rabin, 1999a), more present-biased subjects should be more likely to demand commitment, while under the models of Gul and Pesendorfer (2001) and Fudenberg and Levine (2006) all decisionmakers, including those who don't succumb to temptation, value commitment. With partial sophistication, the prediction depends on the relationship between sophistication and the extent of dynamic inconsistency. In all prior exercises, it appears that more dynamically inconsistent individuals are more aware of their inconsistency, yielding the observed positive correlation between the extent of inconsistency and the extent of commitment demand (Ashraf, Karlan and Yin, 2006; Augenblick et al., 2013; Kaur, Kremer and Mullainathan, 2010).[8] Our evidence indicates that those with the largest self-control problems may be *less* likely to be aware, as it is those who overcome temptation who disproportionately demand commitment. Importantly, in all of the conducted exercises, including our own, the proportion of explained variance in commitment choice is small, indicating limited correlation between dynamic inconsistency and sophistication. Clearly, more research is required in this domain before any conclusions can be drawn.

Our findings also provide policy implications. Critically, our results show clear potential health effects from making choices in advance. In our study, advance bundles contain significantly less calories, fat, saturated fat, carbohydrates, protein and added sugar; and significantly more natural sugar and fiber compared to immediate bundles. Policy makers wishing to alter patterns of food consumption in these directions could do so by encouraging advance choice. Our findings also suggest that the observable

[7] In section 4.4, we analyze differential behavior by time of day of food delivery. However, the timing is not exogenous as it is determined partly by the decisionmaker and may be an avenue for commitment.

[8] Ashraf et al. (2006) show a limited, but significant, correlation between diminishing impatience in monetary choices and take-up of a savings commitment device for women. Augenblick et al. (2013) show a small, but significant, correlation between dynamic inconsistency and commitment demand for real effort tasks. Kaur et al. (2010) document a relationship between take-up of a dominated wage contract and increased effort close to paydays.

benefits of providing commitment may be limited. Commitment in our setting allows subjects who would overcome temptation anyway an alternative method of doing so. Hence, provision of commitment devices may not dramatically affect health outcomes. This conclusion resonates with one recent finding on commitment demand in gym attendance by Royer, Stehr and Sydnor (2012), who document commitment being disproportionately demanded by subjects who already exercise regularly.

In what follows, Section 2 provides an overview of the experimental design, Section 3 couches our design in theories of temptation preferences, Section 4 describes our results and Section 5 concludes.

2 Design of Experiment

2.1 Experimental Setup

To examine dynamic inconsistency in food choice, we conducted a natural field experiment with 218 subjects at a local grocery store in Chicago, Illinois.[9] The experiment was implemented in the summer of 2014 at Louis' Groceries, a small-format neighborhood grocery store in the low-income community of Greater Grand Crossing.[10] The grocery store is a pilot program of the non-profit organization Louis' Groceries NFP, whose mission is to increase community access to healthy food options. Before Louis' opened its doors in 2012, the area was classified as a 'food desert,' meaning that 'a substantial share of residents have low levels of access to a grocery store or healthy, affordable retail outlet.'[11] Unfortunately, Louis' closed its doors in the winter of 2014, which means that the area has again become a 'food desert.'

Louis' Groceries carried out a promotion inviting customers to sign up for a free, 2-week home food delivery program. Two research assistants worked at the grocery store to conduct the experiment and deliver the food items. Subjects for the study were recruited through flyers posted in the store during a 6-week period. We assured that foods were fresh and produce was not bruised at time of delivery by working

[9]Two hundred twenty-five subjects were initially recruited into the study. Of these 225, 5 subjects attrited from the sample before receiving any food deliveries could be made. One subject requested their data not be used in analysis and one subject did not make a commitment decision.

[10]The Greater Grand Crossing community area has a total population of 35,217 (Census, 2010), the majority of whom are African Americans (97.8%). The neighborhood has high rates of poverty, with 28.5% of its residents living on an income below the poverty level. The community also has high incidences of obesity and basic prophylactic health neglect, which are higher than the city of Chicago overall.

[11]USDA website, http://apps.ams.usda.gov/fooddeserts/fooddeserts.aspx. The nearest grocery store is over 1 mile away.

closely with the grocery store and preparing deliveries as close to the delivery time as possible. In keeping with natural field experiment methodology, subjects did not know that they were in an experiment. Thus, we are able to observe subjects in their natural environment as they made a series of intertemporal food allocation decisions.

In return for participating in the promotion, including selecting items, receiving 2 weekly deliveries and completing a pre- and post- survey, subjects received a $20 participation payment. At the conclusion of the study, subjects were notified that they had participated in an experiment and had the option to withdraw their data. One subject chose to withdraw, and this subject's data is not in the dataset.

In our experiment, we observe allocation choices but not actual consumption of food items. One may worry that subjects' choices do not represent their true preferences, but rather reflect their external opportunities to trade healthy and unhealthy items. For example, a subject who can trade healthy for unhealthy items more advantageously outside of the experiment may choose a bundle consisting only of healthy items, conduct appropriate trades, and generate for herself an opportunity set that dominates that provided by experimental choice. A similar arbitrage argument is used to question the use of monetary payments in studies of discounting (Cubitt and Read, 2007; Chabris, Laibson and Schuldt, 2008; Andreoni and Sprenger, 2012; Augenblick et al., 2013).

Several aspects of the experimental environment largely minimize the possibility of arbitrage. Louis' groceries is one of few neighborhood places where fresh fruits and vegetables can be obtained and the prices in the store are identical to those faced in the experiment. Hence, external exchanges are unlikely to be advantageous. Additionally, conducting exchanges with others in the neighborhood is practically difficult, including 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 study. Hence, if choice is driven 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 either all healthy or all unhealthy items. Such bundle concentration is, in fact, quite infrequent.[12] One additional concern posed by not observing food consumption is that if foods are not consumed immediately temptation may be limited. This is a clear limitation of the present design and may indicate that observed dynamic inconsistencies and commitment demand represent lower bounds on the potential prevalence of such phenomena.

[12]Only 10 of 218 bundles consist of either all healthy or all unhealthy items.

2.2 Foods and Ratings

A total of 20 different food items were used in the study. Figure 1 reproduces the promotion sheet of relevant foods - each cell in the figure corresponded to $1 worth of foods. Ten of the food items are considered healthy items: {bananas, cucumbers, Granny Smith apples, green peppers, oranges, pears, plums, Red Delicious apples, red peppers, tomatoes}. The other ten are considered unhealthy items: {Cheetos, Cheez-Its, Doritos, fudge brownies, Honey Buns, potato chips, Nutter Butters, Oreo cookies, PayDay bars, Snickers bars}. The healthy and unhealthy foods differ dramatically in their nutritional content. Table 1 provides health information for the food quantities used in the experiment, demonstrating substantially higher calories, fat, saturated fat, and added sugar for unhealthy foods and substantially lower fiber and natural sugar.

Table 1: Nutritional Information

Food	Calories	Fat (g)	Saturated Fat (g)	Carbohydrates (g)	Fiber (g)	Natural Sugar (g)	Added Sugar (g)	Protein (g)
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 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

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

Upon signing up for the promotion in Week 1, subjects were asked to rate all 20 foods on a scale from 1 to 7. The use of Likert scales to rate different foods has been promoted in the nutrition literature as a means for assessing dietary preferences (Geiselman, Anderson, Dowdy, West, Redmann and Smith, 1998). 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 ratings data provide a critical source of information for our design, as it conveys what items subjects may find tempting. The data was also used to inform the set of

Figure 1: Study Foods

additional items that were available for exchange at the time of delivery. Completing a rating for all foods was voluntary; nevertheless, most subjects rated a large number of foods, with 191 of 218 rating 15 or more foods. However, ratings data is less reliable for some subjects, since some rated few or no foods, and others gave many foods the same rating. The composition of additional items was chosen randomly for those subjects who did not rate any foods (9), or did not rate enough foods. Seventy-seven of 218 subjects gave every food in their advance bundle their highest rating. For these subjects, tempting alternatives were also chosen at random from the set of highest rated items. We cannot be sure exactly what foods subjects with unusable ratings like most and find tempting. Intuitively, not providing sufficiently tempting alternatives may limit our ability to identify dynamic inconsistencies. In our robustness tests we investigate the possibility that our potentially tempting alternatives are not sufficiently tempting (see Section 4.4.2).

2.3 Experimental Timeline

As summarized in Table 2, the experiment was conducted over a series of three weeks.

Week 1, Advance Choice: Upon signing up for the promotion in Week 1, subjects received an order sheet and brochure listing available items and made a decision about

Table 2: Summary of Experiment

Week 1	**Week 2**		**Week 3**
* Select Delivery 1 items * Pre-Questionnaire Food Ratings	* Receive Delivery 1 items * Decide about changes to Delivery 1 items * Select Delivery 2 items	* Commitment choice for Delivery 2	* Receive Delivery 2 items * If no commitment: Decide about changes to Delivery 2 items * Post-Questionnaire

which items to receive in their first delivery. The selection was made from the 20 possible items listed in Figure 1. All items were also currently available at the store, and the fresh fruits and vegetables were visible to the subjects as they made their selections. To simplify the selection process, each item was valued at $1, with cheaper foods bundled into several for $1 (e.g., 3 bananas together cost $1). All items, both healthy and unhealthy, were priced as close as possible to their market price. Subjects were asked to create a 'basket' of items valued at $10 in total, and could choose from any of the 20 items, including selecting the same item more than once. Following their selections, subjects also completed a questionnaire about their preferred dates and times of delivery and provided a rating of their food preferences (as described in Section 2.2).

Subjects were informed that they would need to be home during their delivery, and would need to show a picture ID to the delivery person to receive their basket. Delivery was scheduled as close to 7 days post-sign up as possible, taking into account the constraints faced by the grocery store (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 in Week 1 to confirm enrollment, which also allowed us to validate their phone number.

Week 2, Immediate Choice: The following week (Week 2), 1-2 days before scheduled delivery, we initiated a reminder call to ensure that subjects would be home and then proceeded with the delivery.

For each subject, we created a customized box of 4 additional items (as in Week 1, cheaper foods were bundled into several for $1) based upon their prior food ratings. This box contained their highest rated healthy item, their highest rated healthy item not included in their original bundle, their highest rated unhealthy item, and their highest rated unhealthy item not included in their original bundle. Upon delivery, subjects were given the opportunity to make up to four exchanges between their bundle and these additional items. Subjects were not told in advance that they would have this

opportunity to exchange. The opportunity to exchange was conducted 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 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.]*

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.

Week 2.5, Commitment Choice: 1-2 days prior to the Week 3 delivery, we again initiated a reminder call to ensure that subjects would be home. This time, we also elicited the demand for commitment by asking subjects whether they would like to receive the box of additional items during the Week 3 delivery. This question was again fully scripted as:

> *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?*

If a subject answered 'yes,' the box of additional items was prepared in the same way as before and presented at the time of delivery. If a subject answered 'no.' a box of additional items was not delivered and we classified this subject as one who demands commitment. In the case that subjects did not pick up the phone for the reminder call at least 1 day prior to the delivery date, we continued calling and re-scheduled the delivery for 1 day after we reached the subject to elicit commitment.

At the time of the Week 3 delivery, subjects completed another questionnaire and received a voucher that could be redeemed for $20 in cash at Louis' Groceries upon their next visit.

3 Theoretical Predictions

Our experiment offers subjects a choice between unhealthy, g, and healthy, h, foods. The price of each food item (healthy and unhealthy) is $1. With a fixed budget of

\$10, this establishes the choice set, $(g,h) \in X = \{(0,10),(1,9),...,(9,1),(10,0)\}$.[13] We study choices made at two points in time. First, we examine the choices made in advance of food delivery, which we refer to as advance decisions. Second, we study choices made at the time of food delivery, which we call immediate decisions. Finally, we explore the extent of commitment demand exhibited as an advance preference for smaller choice sets.

We analyze behavior through the lens of temptation based models (Gul and Pesendorfer, 2001; Fudenberg and Levine, 2006).[14] In applying temptation preferences, we make two assumptions. The first is that temptation is felt only at the time of food delivery, and the second is that advance decisions are generated under the guise of commitment.[15] Under these assumptions, for a given choice set, A, and immediate decisions, the Gul and Pesendorfer (2001); Fudenberg and Levine (2006) formulation delivers a decisionmaker who maximizes

$$max_{(g,h)\in A}\ u(g,h) + v(g,h) - v(g^t,h^t)$$

The index $v(g,h)$ is a temptation ranking, capturing the impulses felt by the decisionmaker at the time of food delivery. The maximum of this index for the choice set, A, is the most tempting choice available at the time of food delivery, $v(g^t,h^t) \equiv max_{(g,h)\in A} v(g,h)$. The difference $v(g,h) - v(g^t,h^t)$ represents the foregone utility of choosing the bundle (g,h) and not the most tempting item in the choice set, $v(g^t,h^t)$. These are self-control costs potentially born in immediate decisions that prevent the subject from giving in fully to temptation.

The index $u(g,h)$ is a commitment ranking, capturing the decisionmaker's preferences if he had been able to commit to a single bundle in advance. Consider a subject facing a singleton choice set in an immediate decision, having committed ex-ante to a particular bundle. In such a case, $v(g,h) = v(g^t,h^t)$, no self-control costs are borne, and the decisionmaker receives utility $u(g,h)$. Hence, $u(g,h)$ is a commitment ranking in the sense that it provides the decisionmaker's advance ranking over singleton choice

[13] For ease of explication we abstract from the preferences within food categories and investigate only the number of healthy and unhealthy items chosen.

[14] In particular, we apply the Gul and Pesendorfer (2001) formulation of self-control preference which corresponds to the class of Fudenberg and Levine (2006) formulations that restrict to linear cost functions.

[15] The first of these assumptions is effectively what is assumed in (Gul and Pesendorfer, 2001), restricting the feelings of temptation to second period choice. This assumption is testable by comparing advance and immediate choices to examine the extent and direction of inconsistencies. Relating observed inconsistencies to food rankings provides a second basis for investigation. The second assumption is defensible given our design of blinding subjects to reallocation opportunities in the first week of the experiment.

sets.

The interplay of temptation and commitment rankings determines immediate choice.[16] Tension between $u(g,h)$ and $v(g,h)$ provides the potential for dynamic inconsistency in the sense that advance rankings and immediate choice need not coincide. Note, however, that the decisionmaker may not give in entirely to temptation and choose (g^t, h^t).

Our setting is particularly amenable to the application of temptation based models as we have pre-determined the initial choice set X; identify the ranking $u(g,h)$ from advance decisions; identify the combination of $u(g,h) + v(g,h)$ from immediate decisions; and identify the preference for smaller choice sets from commitment choice. There is a natural set of alternative formulations for dynamically inconsistent choice related to quasi-hyperbolic time preferences (Laibson, 1997; O'Donoghue and Rabin, 1999b). In our setting, applying such models would require a set of assumptions for the discounted utilities of healthy and unhealthy items and their intertemporal organization of costs and benefits. This, in turn, might require additional information on storage opportunities. In our discussion, we attempt to port the insights gained from our study back to these models as well.

3.1 Advance Decisions

Our advance decisions provide the decisionmaker with the choice set $(g,h) \in X = \{(0,10),(1,9),...,(9,1),(10,0)\}$ and ask her to choose a bundle (g^A, h^A) for delivery one week later. That is, the decisionmaker is asked to choose a singleton choice set under the guise of commitment. Under our model of temptation preferences, the decisionmaker chooses according to the commitment ranking, $u(g,h)$, such that

$$(g^A, h^A) = argmax_{(g,h)\in X} u(g,h).$$

We identify (g^A, h^A) for each subject.

3.2 Immediate Decisions

Importantly, the advance decisions are made only under the guise of commitment. The bundle (g^A, h^A) actually serves as an anchor point for the next element of the design. Upon delivery, we provide subjects with the opportunity to exchange up to four items

[16] In immediate choice, the decisionmaker maximizes $u(g,h) + v(g,h)$. Gul and Pesendorfer (2001) note that maximizing this combination represents the optimal tradeoff between commitment and temptation. We ignore the case where no tension exists, $u(g,h) = v(g,h)$.

from their bundle with items in a box of 2 healthy items and 2 unhealthy items. Hence, in immediate decisions, we offer four bundles in addition to (g^A, h^A), generating the choice set $X' = \{(g^A, h^A), (g^A + 1, h^A - 1), (g^A + 2, h^A - 2), (g^A - 1, h^A + 1), (g^A - 2, h^A + 2)\}$. We ignore the elements of X' that substitute g for g and h for h, leaving unchanged the number of healthy and unhealthy items. Note that $X' \subset X$ such that every bundle available in X' was also available and not chosen from X.[17]

Choice from the set X' is governed by the interplay of temptation and commitment rankings. Hence,

$$(g^I, h^I) = argmax_{(g,h)\in X'} u(g, h) + v(g, h) - v(g^t, h^t),$$

where $v(g^t, h^t) = max_{(g,h)\in X'} v(g, h)$. We identify (g^I, h^I) for each subject and identify dynamic inconsistencies as revealed preference violations,

$$(g^I, h^I) \neq (g^A, h^A).$$

Given that immediate choice is determined by the interplay between temptation and commitment ranking, while advance choice is determined by commitment ranking alone, dynamic inconsistencies will correspond to increasing the temptation value of the chosen items. Different decisionmakers may find different items tempting, such that any deviation between (g^I, h^I) and (g^A, h^A) could be taken as evidence in support of temptation preferences. However, a natural intuition is that unhealthy foods are tempting, such that (g^I, h^I) will involve greater levels of g than (g^A, h^A).

3.3 Commitment Demand

In the model of temptation preferences, commitment demand is exhibited as a preference for smaller choice sets. In particular, we investigate the preference for singleton choice sets. In Week 2.5, after having made the allocation (g^A, h^A), subjects are asked whether they prefer the delivery driver to bring only (g^A, h^A) or to deliver the set $X' = \{(g^A, h^A), (g^A + 1, h^A - 1), (g^A + 2, h^A - 2), (g^A - 1, h^A + 1), (g^A - 2, h^A + 2)\}$. We assume that this decision is made without any tempting impulses.

[17] The set X' may induce some censoring. Subjects wishing to consume a lot of g in the immediate choices may not be able to do so as they are only offered $g^A + 2$. Further it may be that the items included in X' are not sufficiently tempting. Hence, we may be identifying a lower bound on dynamic inconsistency.

A subject will choose to commit if

$$u(g^A, h^A) > u(g^I, h^I) + v(g^I, h^I) - v(g^t, h^t),$$

where again $v(g^t, h^t) = max_{(g,h)\in X'} v(g, h)$. Because (g^A, h^A) is the maximizer of $u(g, h)$ and the choice set X' always generates negative utility associated with self-control costs, this inequality will always hold. That is, if subjects suffer from temptation at all, they will value commitment. We identify the preference for making immediate decisions from the singleton choice set (g^A, h^A) over X' from each subject's commitment choice.

Note that a subject need not succumb to temptation in order to demand commitment. A subject who succumbs to temptation, and hence delivers dynamic inconsistencies, satisfies

$$u(g^I, h^I) + v(g^I, h^I) - v(g^t, h^t) > u(g^A, h^A) + v(g^A, h^A) - v(g^t, h^t).$$

The commitment inequality above holds with or without this restriction.[18]

3.4 Sophistication

The temptation based models of Gul and Pesendorfer (2001); Fudenberg and Levine (2006) feature complete sophistication in the sense of O'Donoghue and Rabin (1999b). That is, the decisionmaker is fully aware of her potential to feel temptation when making immediate decisions. This sophistication leads her to exhibit an advance preference for commitment, preferring smaller choice sets with limited opportunity for temptation. A natural extension of these models would account for partial sophistication, wherein subjects are not fully aware of their future temptation potential. O'Donoghue and Rabin (2001) provide such an extension in the domain of quasi-hyperbolic discounting by including an additional behavioral parameter that captures the decisionmaker's beliefs about her future present bias.

Following a similar path, one could accommodate partial sophistication in temptation based models by endowing the advance decisionmaker with beliefs about imme-

[18] A subject will commit and be dynamically inconsistent if

$$u(g^I, h^I) + v(g^I, h^I) - v(g^t, h^t) < u(g^A, h^A) < u(g^I, h^I) + v(g^I, h^I) - v(g^A, h^A).$$

A subject who is dynamically consistent will exhibit $(g^A, h^A) = (g^I, h^I)$ and will still satisfy the commitment equation

$$u(g^A, h^A) + v(g^A, h^A) - v(g^t, h^t) < u(g^A, h^A).$$

diate choice.[19] As in models of partially sophisticated quasi-hyperbolic discounting, it is the distribution of preferences and beliefs and the correlation between the two that determines the pattern of results.

To date, relatively little is known about the relationship between temptation and sophistication. Ashraf et al. (2006) show a limited, but significant, correlation between diminishing impatience in monetary choices and take-up of a savings commitment device for women. Augenblick et al. (2013) show a small, but significant correlation between dynamic inconsistency and commitment demand for real effort tasks. Kaur et al. (2010) document a relationship between take-up of a dominated wage contract and increased effort close to paydays. The above studies suggest that subjects who have self-control problems are at least partially aware of it. However, in each exercise the proportion of explained variance in commitment choice is small, indicating potentially limited correlation between temptation and sophistication.

4 Results

We present the results in four sub-sections. The first two sub-sections discuss advance and immediate choices of food bundles in Weeks 1 and 2 of the experiment, documenting substantial evidence of dynamic inconsistency. The third sub-section provides evidence of commitment demand and links previously measured behavior to commitment choice. Finally, the fourth sub-section provides robustness tests analyzing alternate measures of dynamic inconsistency, incorporating elicited information of food preferences, and investigating stability in choice throughout the experiment.

4.1 Advance Choice

In Week 1 of the study, the 218 subjects made advance decisions for healthy and unhealthy foods. Each subject had the opportunity to choose 10 items, yielding a potential total of 2180 advance food choices. One subject chose only 7 items; thus, we have 2177 total advance food choices in our dataset.

[19] For example, the decisionmaker could envisage maximizing

$$max_{(g,h)\in A}\ u(g,h) + \alpha[v(g,h) - v(g^t,h^t)],$$

where the parameter $\alpha \in [0,\infty]$ captures the decisionmaker's distorted forecast of her future temptation and hence future temptation costs. If $\alpha = 1$, no belief distortions are present, and the subject behaves exactly as in Gul and Pesendorfer (2001). If $\alpha \in [0,1)$, the subject believes she will be less tempted than she truly will be. If $\alpha = 0$, the subject is completely naive such that she will potentially succumb to temptation and find no value in commitment. If $\alpha > 1$, the subject believes she will be more tempted than she truly will be.

Figure 2 presents a summary of the 2177 foods chosen in advance, presented as dashed black bars. Following our prior definitions we term healthy items, h, as being from the set {bananas, cucumbers, Granny Smith apples, green peppers, oranges, pears, plums, Red Delicious apples, red peppers, tomatoes}, and unhealthy items g as being from the set {Cheetos, Cheez-Its, Doritos, fudge brownies, Honey Buns, potato chips, Nutter Butters, Oreo cookies, PayDay bars, Snickers bars}. A vertical black line in Figure 2 separates healthy and unhealthy items. Out of 2177 foods chosen, 1175 are healthy foods, while 1002 are unhealthy foods.

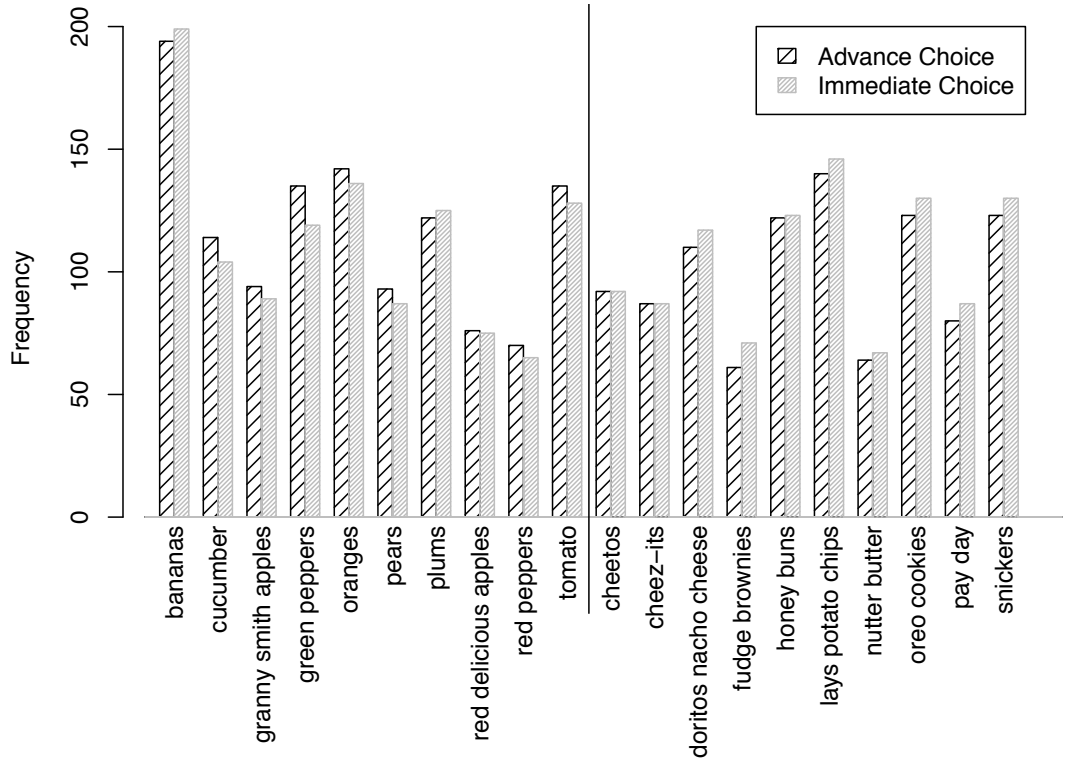

Figure 2: Delivery 1 Choices

Foods exhibit varying popularity both within and across the groups of items. Bananas were the most chosen element of g and potato chips were the most chosen element of h. Within bundles, repeat items are relatively rare. 140 of 218 bundles have no repeated items and 207 bundles consist of seven or more unique items.

Advance choices identify the bundle (g^A, h^A). These bundles appear to be diverse in terms of the proportion of healthy and unhealthy items and widely heterogeneous across subjects. The average advance choice bundle consists of 54.0% (s.d.=20.7%) healthy

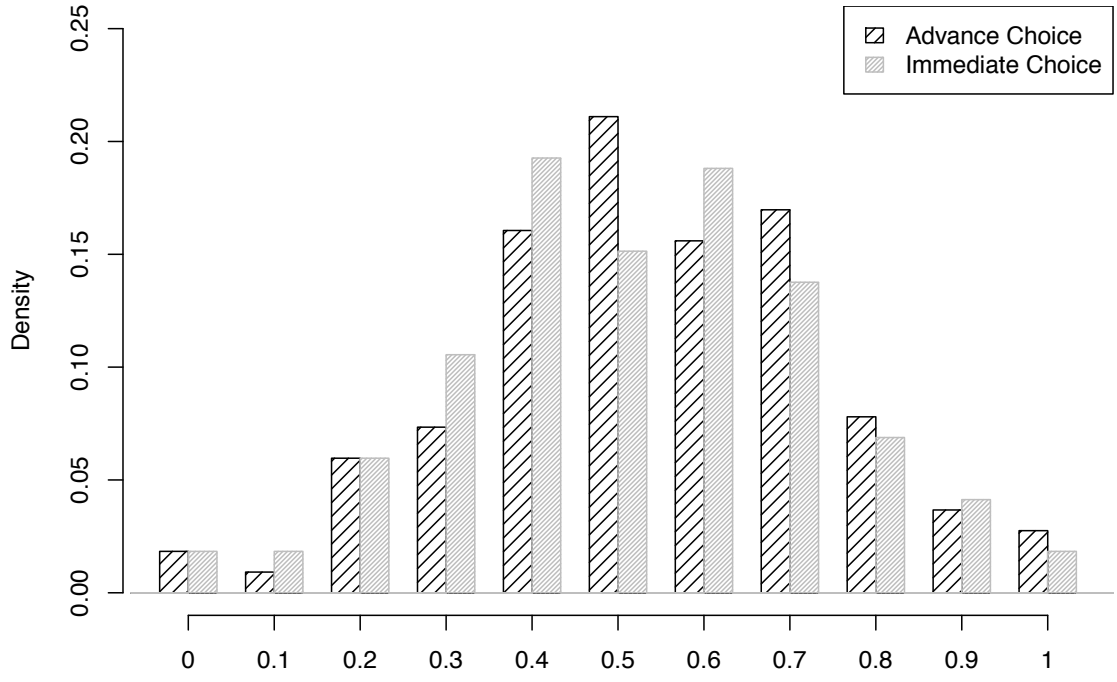

Figure 3: Proportion of Healthy Bundle Items, h

items.[20] Figure 3 presents the distribution of healthy items in each bundle generated at the moment of advance choice as dashed black bars, demonstrating substantial heterogeneity.[21] One important feature of Figure 3 is the relative infrequency of bundles completely concentrated on healthy items or on unhealthy items. Only 10 of 218 bundles are concentrated this way, suggesting that the subjects are not employing the arbitrage strategies that would lead them to bundle concentration noted in Section 2.1.

4.2 Immediate Choice and Dynamic Inconsistency

We now turn to immediate choices made at the time of food delivery in Week 2 of the study. Upon delivery, subjects were given the opportunity to make up to four substitutions from a customized selection of two healthy and two unhealthy items. That is, subjects were given the opportunity to choose from the set $X' = \{(g^A, h^A), (g^A +$

[20] We present bundle proportions instead number of healthy items to account for the one subject who did not make 10 choices.

[21] Appendix Figure A1 provides the empirical cumulative distribution function.

$1, h^A - 1), (g^A + 2, h^A - 2), (g^A - 1, h^A + 1), (g^A - 2, h^A + 2)\}$.

Figure 2 reports the aggregate immediate choices made when faced with choice set X' as filled gray bars. Notable from Figure 2 is substantial evidence of dynamic inconsistency with a clear direction of change. Out of 2177 foods chosen, 1127 are healthy foods, while 1050 are unhealthy foods. That is, the number of healthy foods chosen decreases by 48, while the number of unhealthy foods chosen increases by 48. The inconsistency is observable across food items. Eight of ten healthy foods are chosen less frequently in immediate choice relative to advance choice, while eight of ten unhealthy foods are chosen more frequently. This pattern of aggregate inconsistency is in line with the intuition that unhealthy foods have higher temptation value.

Immediate choices identify the bundle (g^I, h^I). These bundles continue to vary substantially across subjects. The average immediate choice bundle consists of 51.8% (s.d.= 20.9%) healthy items. Figure 3 presents the distribution of bundle proportions generated at the moment of immediate choice as filled gray bars, continuing to demonstrate substantial heterogeneity.[22]

We compare (g^A, h^A) to (g^I, h^I) with simple tests for paired data. Given our large sample size, even relatively small inconsistencies in choice are potentially detectable. In a paired t-test, the estimated difference in mean healthy proportions of 54.0%-51.8% = 2.2% is statistically different from zero, $t = 6.55$, $(p < 0.01)$. A Wilcoxon signed rank test also rejects the null hypothesis for equal distributions in healthy food proportions between advance and immediate choice, $z = 6.17$, $(p < 0.01)$.[23] Of our 218 subjects, 172 exhibit $(g^I, h^I) = (g^A, h^A)$. Of the 46 dynamically inconsistent subjects, 2 choose the more healthy bundle $(g^I, h^I) = (g^A - 1, h^A + 1)$, 38 choose the less healthy bundle $(g^I, h^I) = (g^A + 1, h^A - 1)$, and 6 choose the less healthy bundle $(g^I, h^I) = (g^A + 2, h^A - 2)$.[24]

The contrast between advance and immediate bundles is compelling. On average, subjects are dynamically inconsistent, preferring less healthy items when making immediate choices than when deciding in advance. Revealed preference violations of

[22]Appendix Figure A1 provides the empirical cumulative distribution function.

[23]It should be noted that the most pronounced the dynamic inconsistency could be for each subject would be to make two exchanges for less healthy items, yielding $(g^I, h^I) = (g^A + 2, h^A - 2)$, subject to censoring. This would generate a mean healthy proportion for immediate choice of 34.4% (s.d. = 19.7%) and an estimated difference of 19.6%. As such, our level of observed inconsistency is about 10% the total possible observable value.

[24]Recall that this analysis ignores any inconsistencies that leave the number of healthy and unhealthy items unchanged. Though the majority of inconsistencies lead to different numbers of healthy and unhealthy foods, 36 of 178 subjects with $(g^I, h^I) = (g^A, h^A)$ exchanged either h for h or g for g when making their immediate choices. In section 4.4, we examine our results with alternative measures for dynamic inconsistency.

$(g^A, h^A) \neq (g^I, h^I)$ are documented for roughly 21% of subjects, and 96% of violations are towards less healthy bundles in immediate choice.

4.3 Commitment Demand

In Week 2.5, before the delivery of their items in Week 3, subjects were offered the opportunity to commit to their advance bundle (g^A, h^A) and not face the set $X' = \{(g^A, h^A), (g^A + 1, h^A - 1), (g^A + 2, h^A - 2), (g^A - 1, h^A + 1), (g^A - 2, h^A + 2)\}$ when making their immediate choice. Seventy-three of 218 (33%) subjects preferred not to face the expanded choice set X' when making immediate decisions. Commitment demand is a critical prediction of temptation-based models, and our data show that a sizable fraction of subjects are willing to restrict themselves to a singleton choice set.

Linkages between commitment demand and patterns in behavior are critical for validating and distinguishing between models of dynamic inconsistency. We investigate the correlates of commitment demand by linking commitment to previously measured behavior and collected socio-demographic data. Figure 4 presents initial results. Subjects who demand commitment construct advance bundles that are 57.1% (s.d. =19.6%) healthy items while subjects who do not demand commitment construct advance bundles that are 52.4% (21.1%) healthy items, $t = 1.60,\ (p = 0.11)$. Subjects who demand commitment also construct immediate bundles that are more healthy than subjects who do not demand commitment, 56.3% (19.8%) versus 49.5% (21.2%), $t = 2..29,\ (p < 0.05)$. When comparing advance and immediate decisions in the first delivery, committing subjects' bundles grow 0.8% (4.0%) less healthy while non-committing subjects' bundles grow 2.9% (5.3%) less healthy. The difference between committing and non-committing subjects is significant at all conventional levels, $t = 2.97,\ (p < 0.01)$.[25] In general, committing subjects are substantially less likely to be dynamically inconsistent. Nine of 46 (20%) dynamically inconsistent subjects demand commitment, while 64 of 172 (37%) dynamically consistent subjects do so.[26] Subjects who demand commitment are exhibiting more healthy choices and are less dynamically inconsistent than their non-committing counterparts.

Table 3 provides corresponding regression analysis with linear probability models for take-up of commitment and previously measured behavior. The measure of dynamic inconsistency is the difference in healthy food proportions between advance

[25] A Mann-Whitney rank sum test for difference in distributions yields a similar conclusion, $z = 3.02,\ (p < 0.01)$.

[26] Interestingly, both subjects whose bundles grew more healthy such that $(g^I, h^I) = (g^A - 1, h^A + 1)$ demanded commitment.

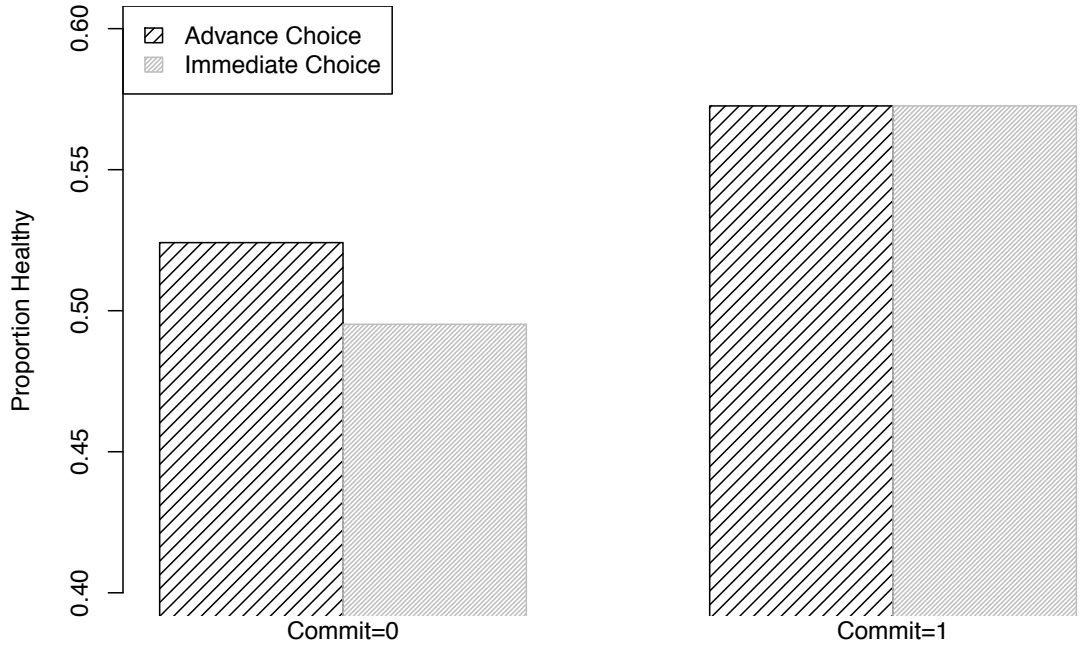

Figure 4: Delivery 1 Choice and Commitment Demand

and immediate bundles, *(Advance - Immediate) Proportion Healthy*. The measure of healthiness in advance choice is the proportion of healthy foods, *Advance Proportion Healthy*. In columns (1) through (3), we document that dynamic inconsistency and less healthy advance decisions are significantly correlated with lower take-up of commitment. A difference of 10%-age points in the proportion of healthy foods between advance and immediate choice is associated with around a 19%-age point decrease in the probability of commitment. In columns (4) and (5), we restrict attention to a subsample of 192 subjects for whom we have complete socio-demographic information. Socio-demographics seem virtually uncorrelated with commitment choice in our sample. Importantly, across specifications we find very limited predictive power in total for commitment. Including all observable characteristics and prior behavior in column (5), we explain only around 6% of the variation in commitment choice.

Table 3: Commitment Demand

	Dependent Variable: Commit = 1				
	(1)	(2)	(3)	(4)	(5)
(Advance - Immediate) Proportion Healthy	-1.885***		-1.969***	-1.797***	-1.769***
	(0.557)		(0.586)	(0.634)	(0.647)
Advance Proportion Healthy		0.248	0.281*	0.316**	0.327**
		(0.150)	(0.150)	(0.155)	(0.154)
Age					-0.001
					(0.003)
Income					-0.000
					(0.000)
Male (=1)					-0.036
					(0.074)
Household Size					-0.001
					(0.023)
Constant	0.376***	0.201**	0.226***	0.229***	0.331
	(0.036)	(0.084)	(0.083)	(0.086)	(0.202)
# Observations	218	218	218	192	192
R-Squared	0.039	0.012	0.054	0.051	0.063

Notes: Ordinary least squares regression. Robust standard errors in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01.

4.4 Robustness Tests

The results to here demonstrate three key facts. First, subjects exhibit dynamic inconsistency in food choices, transitioning from more healthy advance bundles to less healthy immediate bundles on average. Second, commitment demand is prevalent with roughly one-third of subjects preferring to avoid a tailored box of tempting alternatives. Third, commitment demand is linked with more healthy and less dynamically inconsistent choice. Before interpreting these findings, we provide a series of robustness tests. These tests focus on alternate measures of dynamic inconsistency, the incorporation of subject food-rating data, and the stability in behavior throughout the experiment.

4.4.1 Alternate Measures of Dynamic Inconsistency

Our measure of dynamic inconsistency focuses on Week 2 exchanges that alter the proportion of healthy and unhealthy items in a subject's bundle. Forty-six of 218 (21%) subjects exhibit such inconsistencies, with 44 of 46 (96%) of inconsistencies being towards less healthy bundles upon immediate choice. Here we examine the robustness of our findings to considering two further forms of inconsistency. First, we examine any dynamic inconsistency, including inconsistencies that leave unchanged the proportion of healthy and unhealthy items. We define the variable *Any Inconsistency* as an indi-

cator for whether any exchanges were made by a subject in Week 2. Eighty-two of 218 (37.6%) subjects have *Any Inconsistency* equal to one, indicating that the majority of inconsistencies (46 of 82) involve changes that alter the proportion of healthy and unhealthy items in the bundle. The correlation between *Any Inconsistency* and *(Advance - Immediate) Proportion Healthy* is $\rho = 0.57,\ (p < 0.01)$. Table 4 reproduces the analysis of Table 3 with *Any Inconsistency* as the measure of dynamic inconsistency. The results are qualitatively unchanged. Across specification, inconsistencies in choice are linked to subjects being less likely to demand commitment.

Table 4: Commitment Demand with Any Inconsistency

	Dependent Variable: Commit = 1				
	(1)	(2)	(3)	(4)	(5)
Any Inconsistency	-0.185***		-0.191***	-0.199***	-0.200***
	(0.062)		(0.063)	(0.069)	(0.069)
Advance Proportion Healthy		0.248	0.273*	0.318**	0.331**
		(0.150)	(0.149)	(0.154)	(0.154)
Age					-0.001
					(0.003)
Income					-0.000
					(0.000)
Male (=1)					-0.038
					(0.073)
Household Size					-0.001
					(0.022)
Constant	0.404***	0.201**	0.259***	0.259***	0.361*
	(0.042)	(0.084)	(0.086)	(0.089)	(0.200)
# Observations	218	218	218	192	192
R-Squared	0.036	0.012	0.050	0.057	0.069

Notes: Ordinary least squares regression. Robust standard errors in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01.

Second, we examine the nutritional composition of each bundle. Following the nutrition information in Table 1, we construct the total calories, fat, saturated fat, added sugar, carbohydrates, protein, natural sugar, and fiber for both advance and immediate bundles. Figure 5 provides the relationship along each of these dimensions for advanced and immediate bundles along with t-tests comparing means. It is clear from Figure 5 that the average level of inconsistency moving from healthy to unhealthy foods is echoed in nutritional composition. Immediate bundles contain substantially and statistically significantly more calories, fat, saturated fat, added sugar, and carbohydrates, and slightly more protein than advance bundles. Advance bundles contain significantly more natural sugar and fiber. To summarize the inconsistency in food choice, we perform principal components analysis on the eight nutritional differences

contained in Figure 5. We extract the first principal component as a *Food Consistency Factor*.[27] The *Food Consistency Factor* provides a summary for the consistency in choice at the nutritional level. The correlation between the *Food Consistency Factor* and *(Advance - Immediate) Proportion Healthy* is $\rho = -0.76$, $(p < 0.01)$, indicating that subjects with higher food consistency factors exhibit fewer dynamic inconsistencies from healthy to unhealthy items.[28] Table 5 re-conducts the analysis of Table 3 with *Food Consistency Factor* as the measure of dynamic inconsistency. The results are qualitatively unchanged. Across specification, inconsistencies in nutrition are linked to subjects being less likely to demand commitment.

Table 5: Commitment Demand with Principal Component of Food Inconsistency

	Dependent Variable: Commit = 1				
	(1)	(2)	(3)	(4)	(5)
Food Consistency Factor	0.071**		0.073**	0.073**	0.071**
	(0.030)		(0.031)	(0.033)	(0.032)
Advance Proportion Healthy		0.248	0.259*	0.301*	0.317**
		(0.150)	(0.151)	(0.156)	(0.155)
Age					-0.001
					(0.003)
Income					-0.000
					(0.000)
Male (=1)					-0.025
					(0.074)
Household Size					-0.000
					(0.023)
Constant	0.336***	0.201**	0.196**	0.199**	0.309
	(0.032)	(0.084)	(0.085)	(0.088)	(0.202)
# Observations	218	218	218	192	192
R-Squared	0.023	0.012	0.036	0.042	0.053

Notes: Ordinary least squares regression. Robust standard errors in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01.

Taken together, these results demonstrate robustness of our central findings to alternate measures of dynamic inconsistency.

[27]The first principal component has an eigenvalue of 4.65 and explains 58.2% of the variance in the eight nutritional differences. One more factor had an eigenvalue in excess of 1, 2.03, and explained an additional 25% of the variance. In unreported results, this second factor was observed to provide limited additional explanatory power in the specifications of Table 5. The *Food Consistency Factor* is positively correlated with (i.e., loads positively on) the difference in calories, fat, saturated fat, carbohydrates, added sugar and protein between advance and immediate choice; and is negatively correlated with (i.e., loads negatively on the difference in natural sugar and fiber).

[28]The correlation between the *Food Consistency Factor* and *Any Inconsistency* is $\rho = -0.48$, $(p < 0.01)$.

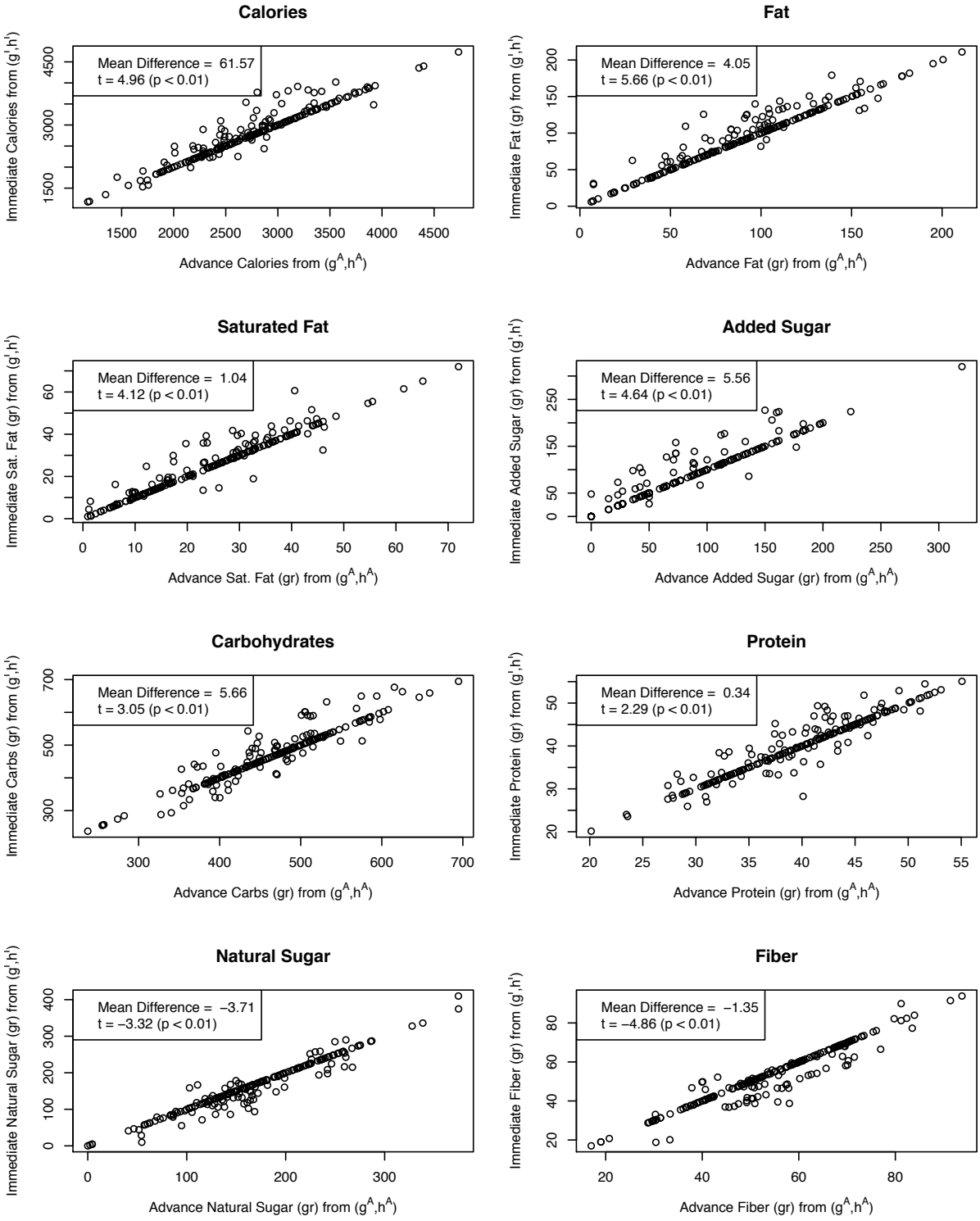

Figure 5: Nutritional Information

4.4.2 Incorporating Food Ratings and Personal Restriction

Critical to our analysis are the initial food ratings provided by subjects. Based upon these food ratings, we construct customized boxes of potentially tempting alternatives.

Two issues are potentially relevant. First, subjects may make relatively few ratings, limiting our ability to ensure that the additional items are indeed tempting. Second, subjects may rate many of the goods in their initial bundle highly, such that they are potentially unlikely to be tempted by the provided alternatives.

Table 6 reproduces the analysis of Table 3, column(3) with several sample restrictions. In column (1), we reproduce Table 3, column(3), while in column (2) we restrict the sample to those subjects who make 15 or more food ratings. The results are virtually unchanged when restricting the sample to those subjects for whom we can determine what foods they may or may not actually find tempting. In column (3), we make an additional sample restriction. We examine subjects whose initial bundles contain less than 90% maximally rated items (the sample median for the subjects who make 15 or more food ratings). Though less precise, we again find a qualitatively similar pattern of results. Importantly, in these two restricted samples we find that the proportion of dynamically inconsistent subjects and proportion of committing subjects is virtually unchanged when compared to the full sample. Across sample restrictions, we find consistent results: subjects exhibit dynamic inconsistency and commitment and a significant negative correlation between the two.

Table 6: Incorporating Ratings Data

	Dependent Variable: Commit = 1					
	(1) Full Sample	(2) $\geq$ 15 Ratings	(3) $\geq$ 15 Ratings & $<$ 0.9 Max	(4) Full Sample	(5) $\geq$ 15 Ratings	(6) $\geq$ 15 Ratings & $<$ 0.9 Max
(Advance - Immediate) Proportion Healthy	-1.969***	-2.045***	-1.707*	-2.026***	-2.169***	-1.988**
	(0.586)	(0.644)	(0.934)	(0.573)	(0.617)	(0.934)
Advance Proportion Healthy	0.281*	0.299*	0.286	0.272*	0.270*	0.306
	(0.150)	(0.158)	(0.291)	(0.146)	(0.153)	(0.280)
Advance Ratings Max Proportion				-0.192*	-0.252**	-0.442**
				(0.102)	(0.119)	(0.222)
Constant	0.226***	0.216**	0.248	0.369***	0.424***	0.472**
	(0.083)	(0.086)	(0.166)	(0.116)	(0.135)	(0.199)
Proportion Dynamically Inconsistent	0.21	0.22	0.23	0.21	0.22	0.23
Proportion Committing	0.33	0.33	0.37	0.33	0.33	0.37
# Observations	218	191	93	218	191	93
R-Squared	0.054	0.059	0.044	0.070	0.082	0.085

Notes: Ordinary least squares regression. Robust standard errors in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01.

The food ratings data also provide for additional analyses. One clear possibility is that subjects who are dynamically consistent have already chosen their favorite foods. Providing them with additional items provides no benefit and, given some small costs of making the choice, they opt to commit to avoid having to choose from the additional items again. Columns (4) through (6) of Table 6 examines this possibility by including as a regressor the proportion of advance bundle items that are maximally rated. An

interesting correlation is observed across specifications. Subjects who demand commitment construct advance bundles that contain fewer maximally rated items. Hence, subjects who demand commitment are not necessarily already satisfied with their initial bundles.[29]

An additional implication of Table 6 columns (4) through (6) is that subjects who demand commitment may have already been seeking to restrict themselves. Because bundles can contain repeated items, not choosing all maximally rated items is tantamount to self-restriction. That is, it entails choosing less preferred items. This self-restriction correlates significantly with commitment demand across specifications. This correlation is important as it shows that subjects who restrict themselves to the singleton of their advance choice in Week 3 were potentially exercising self-control when making advance choices in Week 1.

4.4.3 Stability of Choice

The observed dynamic inconsistency is identified from comparing the bundle (g^A, h^A) constructed in Week 1 to the bundle (g^I, h^I) constructed in Week 2. In addition to temptation preferences, inconsistencies in choice could have a number of alternate sources, including random utility shocks and the resolution of uncertainty between Week 1 and Week 2. Here we investigate two dimensions of stability that may speak to the empirical relevance of these sources for inconsistency.

First, we examine the correlations in choice at the subject level for advance bundles. The proportion of healthy foods in Week 1 advance bundles and in Week 2 advance bundles correlates highly, $\rho = 0.66,\ (p < 0.01)$. Interestingly, for subjects who are dynamically inconsistent, the correlation is substantially lower, $\rho = 0.37,\ (p < 0.05)$.[30] This low correlation is driven almost entirely by the 9 of 46 inconsistent subjects who choose to commit. For dynamically inconsistent, non-committing subjects the correlation is $\rho = 0.54,\ (p < 0.01)$ while for dynamically inconsistent, committing subjects the correlation is $\rho = -0.09,\ (p = 0.81)$.[31] This result indicates that dynamic inconsistency may be driven by shocks, as inconsistent subjects construct quite different advance bundles in Week 2 than they constructed in Week 1. However, if shocks played a large role in generating choice, one would not expect the greatest inconsistencies in

[29]Further, the correlation between the proportion of maximally rated items and *(Advance - Immediate) Proportion Healthy* is quite limited, $\rho = -0.05,\ (p = 0.46)$. Hence, the coefficients for the change in health proportions in columns (4) through (6) of Table 6 look similar to those of columns (1) through (3).

[30]For subjects with *Any Inconsistency* the correlation is $\rho = 0.61,\ (p < 0.01)$.

[31]For dynamically consistent non-committing subjects, the correlation is $\rho = 0.73,\ (p < 0.01)$ while for dynamically consistent, committing subjects the correlation is $\rho = 0.73,\ (p < 0.01)$.

choice to come from those who choose to restrict themselves via commitment.

Second, we examine the stability of dynamic inconsistency for subjects who choose not to commit. Non-committing subjects construct advance bundles in Week 2 that are 49.2% (21.0%) healthy items and construct immediate bundles in Week 3 that are 46.3% (21.9%) healthy items. In a paired t-test, the average estimated difference in mean healthy proportions of 49.0%-46.1% = 2.9% is statistically different from zero, $t = 5.70,\ (p < 0.01)$. A Wilcoxon signed rank test also rejects the null hypothesis for equal distributions in healthy food proportions between advance and immediate choice $z = 5.31$ $(p < 0.01)$. Thirty-nine of 145 non-committing subjects are dynamically inconsistent between Weeks 2 and 3 of the study, which compares favorably to the 37 of 145 non-committing subjects who are dynamically inconsistent between Weeks 1 and 2. Absent commitment, dynamically inconsistent subjects remain dynamically inconsistent on average. At the subject level, regularities are also observed. Though the correlation in *(Advance - Immediate) Proportion Healthy* falls outside of the range of significance, $\rho = 0.13,\ (p = 0.13)$, the correlation in exhibiting *Any Inconsistency* is significant, $\rho = 0.30,\ (p < 0.01)$.[32] Taken together, these results indicate that subjects who do not commit are dynamically inconsistent on average in the second round of the study, and that there is potential stability in succumbing to temptation at the subject level over time.

5 Discussion and Conclusion

We provide field experimental evidence on dynamic inconsistency and commitment demand in food choice. Taken together, our findings show that dynamic inconsistencies are prevalent, with 20% of subjects altering their proportion of healthy foods chosen between advance and immediate bundles, and roughly 40% of subjects exhibiting any inconsistency in choice. The direction of inconsistency is generally towards less healthy bundles in immediate choice. Around one-third of subjects demand a binding commitment device, restricting themselves to their advance choice.

Subjects who demand commitment are more likely to be dynamically *consistent* when compared to non-committing subjects. Hence, it is subjects who overcome temptation, rather than those who succumb to it, that demand commitment. Investigation of food rating data demonstrate that these committing subjects appear to be restricting themselves in their advance choices, choosing their most preferred items relatively

[32] Additionally a Pearson's Chi-Squared test for independence in *Any Inconsistency* comparing the two rounds is rejected, $\chi^2(1) = 12.72,\ p < 0.01$ and a Fisher's exact test also rejects independence $(p < 0.01)$.

less frequently. Hence, it is disproportionately subjects who exhibit self-control in their advance choices that take advantage of the binding restriction provided by our commitment device.

We analyze our results through the lens of temptation models such as Gul and Pesendorfer (2001) and Fudenberg and Levine (2006). The results are consistent with two central implications of these models: first, the existence of dynamic inconsistencies, borne out of the tension between advance and immediate preferences; and second, the preference for smaller choice sets achieved through binding commitment. These two central results are also consistent with alternative models such as quasi-hyperbolic discounting (Laibson, 1997; O'Donoghue and Rabin, 1999a, 2001), which also predict both dynamic inconsistency and commitment demand.[33]

Though the models referenced predict dynamic inconsistencies and commitment demand, they are differentiated by the relationship between the two. In Gul and Pesendorfer (2001) and Fudenberg and Levine (2006), decisionmakers always value commitment, yielding zero correlation between the extent of inconsistency and the extent of commitment demand. Under the perfectly sophisticated model of O'Donoghue and Rabin (2001), all dynamically inconsistent individuals demand commitment, while all dynamically consistent individuals do not, yielding a positive correlation between the extent of inconsistency and the extent of commitment demand. Under partial sophistication, the correlation depends upon the relationship between beliefs about dynamic inconsistency and the inconsistency itself. In all prior empirical exercises, it appears that more dynamically inconsistent individuals are weakly more aware of their inconsistency, yielding the observed positive correlation between the extent of inconsistency and the extent of commitment demand (Ashraf et al., 2006; Augenblick et al., 2013; Kaur et al., 2010). Our result is in contrast to these findings. It appears that more dynamically inconsistent individuals are *less* likely to be aware of their inconsistency. It should be noted that across exercises, much of the variance in commitment choice remains unexplained. In our study, the proportion of explained variance is around six percent, similar to that of Augenblick et al. (2013).[34] Ultimately, the relationship between temptation and sophistication deserves substantial attention in future research. Empirical data in this vein will add value in distinguishing between models of dynamic inconsistency and pointing the way towards new extensions. Though it is beyond the scope of this paper to articulate such extensions, one is sketched briefly in footnote 19.

[33]We focus primarily on the temptation models, wherein all model elements are directly observable to the researcher, to remain agnostic on timing issues.

[34]Augenblick et al. (2013) also note that many dynamically consistent individuals demand commitment in their study, potentially indicative of some misspecification of their estimated model.

Our results also give insights to innovations in food policy. Advance bundles in our experiment are arguably healthier along a variety of nutritional dimensions. A simple, and perhaps effective, policy for altering food choices in this direction would be mandated advance choice. Such a policy may even be implementable at scale given the existence of government transfer programs targeted towards nutrition. The findings in our low-income urban study population are particularly valuable as they are likely externally valid for low-income populations of interest. At the least, policy-makers should consider allowing individuals enrolled in nutrition assistance programs to make advance decisions by allowing pre-payment for foods in grocery stores or as part of community supported agriculture (CSA) programs. The Supplemental Nutrition Assistance Program (SNAP) already allows pre-payment of up to 14 days, yet this option is not widely advertised. Our findings also indicate the potential promise of policies related to the provision of commitment devices for changing observable nutrition outcomes. Our committing individuals make healthier and less inconsistent choices without commitment. Given the stability in choice, commitment likely did not alter substantially the patterns of consumption in our sample. Hence, even with substantial commitment demand, a policy providing commitment devices may not alter much the healthiness of consumed foods.

Our findings of dynamic inconsistency relate to an exciting body of research investigating temptation models outside of the standard experimental domain of time-dated monetary choices. Relatively little experimental work has investigated dynamic inconsistencies between advance and immediate choice, and even less has investigated such behaviors in the field. Further research investigating temptation models in field settings is necessary to more deeply understand the relevance of the behavioral insights for a broad spectrum of behavior.

References

Ainslie, George and Varda Haendel, "The Motives of the Will," in E. Gottheil, K. Durley, T. Skodola, and H. Waxman, eds., *Etiologic Aspects of Alcohol and Drug Abuse*, Charles C. Thomas, 1983.

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, Muriel Niederle, and Charles Sprenger, "Working over time: Dynamic inconsistency in real effort tasks," Technical Report, National Bureau of Economic Research 2013.

Brown, Alexander L, Zhikang Eric Chua, and Colin F Camerer, "Learning and visceral temptation in dynamic saving experiments," *Quarterly Journal of Economics*, 2009, *124* (1), 197–231.

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.

Cubitt, Robin P. and Daniel Read, "Can Intertemporal Choice Experiments Elicit Preferences for Consumption," *Experimental Economics*, 2007, *10* (4), 369–389.

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.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue, "Time discounting and time preference: A critical review," *Journal of economic literature*, 2002, pp. 351–401.

Fudenberg, Drew and David K. Levine, "A dual self model of impulse control," *American Economic Review*, 2006, p. Forthcoming.

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.

Gine, Xavier, Jessica Goldberg, Dan Silverman, and Dean Yang, "Revising Commitments: Time Preference and Time-Inconsistency in the Field," *Working Paper*, 2010.

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

Halevy, Yoram, "Time Consistency: Stationarity and Time Invariance," *Working Paper*, 2012.

Jia, Haomiao and Erica I Lubetkin, "The impact of obesity on health-related quality-of-life in the general adult US population," *Journal of public health*, 2005, *27* (2), 156–164.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-Control and the Development of Work Arrangements," *American Economic Review, Papers and Proceedings*, 2010, *100* (2), 624–628.

Kumanyika, Shiriki Kinika and Sonya Grier, "Targeting interventions for ethnic minority and low-income populations," *The Future of Children*, 2006, *16* (1), 187–207.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

McClure, Samuel M, David I Laibson, George Loewenstein, and Jonathan D Cohen, "Separate Neural Systems Value Immediate and Delayed Monetary Rewards," *Science*, oct 2004, *306* (5695), 503–507.

Mokdad, Ali H, Earl S Ford, Barbara A Bowman, William H Dietz, Frank Vinicor, Virginia S Bales, and James S Marks, "Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001," *Jama*, 2003, *289* (1), 76–79.

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

_ and _, "Doing it Now or Later," *American Economic Review*, 1999, *89*, 103–124.

_ and _, "Choice and Procrastination," *The Quarterly Journal of Economics*, 2001, *116* (1), 121–160.

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.

_ , George Loewenstein, and Shobana Kalyanaraman, "Combining the Immediacy Effect and the Diversification Heuristic," *Journal of Behavioral Decision Making*, 1999, *12*, 257–273.

Royer, Heather, Mark F Stehr, and Justin R Sydnor, "Incentives, Commitments and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company," Technical Report, National Bureau of Economic Research 2012.

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

Solnick, Jay V., Catherine H. Kannenberg, David E. Eckerman, and Marcus B. Waller, "An Experimental Analysis of Impulsivity and Impulse Control in Humans," *Learning and Motivation*, 1980, *11*, 61–77.

Appendix: Not For Publication

A Additional Figures

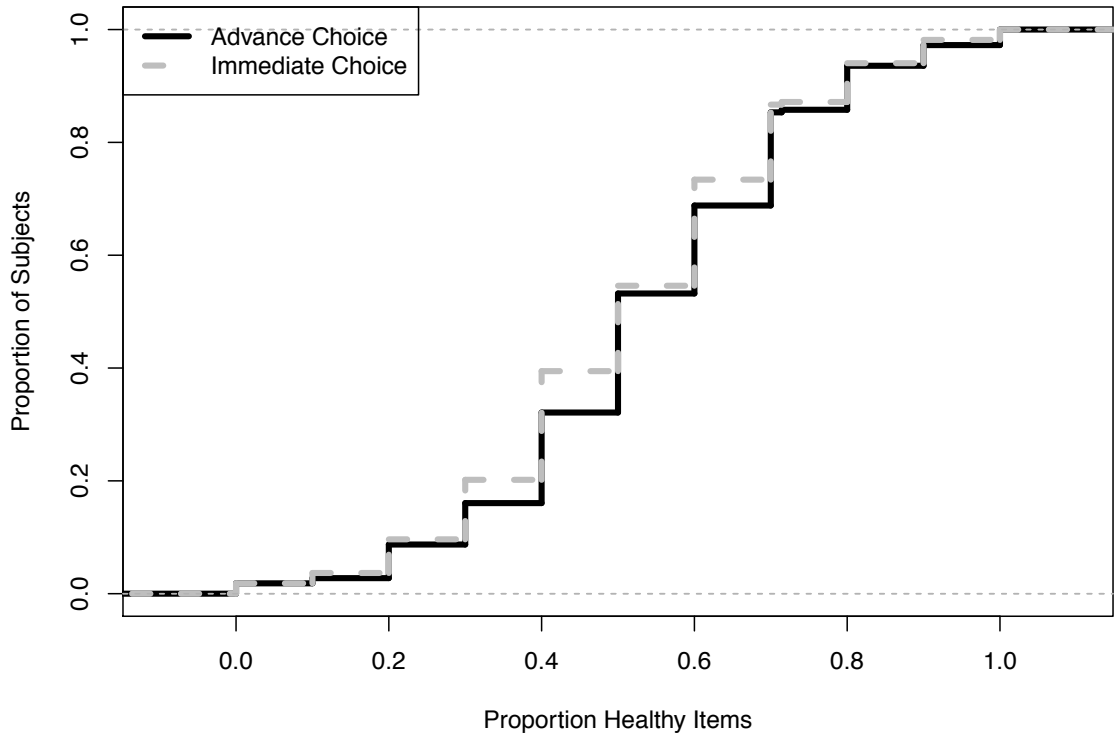

Figure A1: Empirical Cumulative Distribution Functions

B Experiment Script

B.1 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.

B.2 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].

B.3 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.

B.4 Week 2 Delivery and Item Selection for Week 3

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.

B.5 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?

B.6 Week 3 Delivery

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.