

Revealing temptation through menu choice: field evidence

Séverine Toussaert

February 10, 2019

Motivation (1)

1. Commitment demand is a key behavioral implication of economic theories of temptation.
 - ▶ See Laibson (1997), Gul and Pesendorfer (2001), Dekel, Lipman and Rustichini (2009), Fudenberg and Levine (2006, 2012)
2. Implication widely tested both in the lab and in the field
 - ▶ Lab: Augenblick et al. (2015), Houser et al. (2018)
 - ▶ Field: Ashraf et al. (2006), Giné et al. (2010), Royer et al. (2015), Kaur et al. (2015)

Motivation (2)

- ▶ What is in a typical dataset?

Motivation (2)

- ▶ What is in a typical dataset?
 - ▶ 2 options a and b , where b construed as the temptation.
 - ▶ DM asked to choose between $\{a\}$ and $\{a, b\}$.
 - ▶ DM who demands commitment by choosing $\{a\}$ reveals to have a self-control problem with b .

Motivation (2)

- ▶ What is in a typical dataset?
 - ▶ 2 options a and b , where b construed as the temptation.
 - ▶ DM asked to choose between $\{a\}$ and $\{a, b\}$.
 - ▶ DM who demands commitment by choosing $\{a\}$ reveals to have a self-control problem with b .
- ▶ But it is not always clear what the source of temptation is: why not offering $\{b\}$?
 - ▶ Ameriks et al. (2007): problem of underconsumption, especially towards the end of the life cycle.

Motivation (2)

- ▶ What is in a typical dataset?
 - ▶ 2 options a and b , where b construed as the temptation.
 - ▶ DM asked to choose between $\{a\}$ and $\{a, b\}$.
 - ▶ DM who demands commitment by choosing $\{a\}$ reveals to have a self-control problem with b .
- ▶ But it is not always clear what the source of temptation is: why not offering $\{b\}$?
 - ▶ Ameriks et al. (2007): problem of underconsumption, especially towards the end of the life cycle.



Motivation (3)

- ▶ Little evidence of commitment demand driven by temptation:
 - ▶ Low take-up rates in range 10%-35%.
- ▶ Why this low demand?
 - ▶ One possible reason: commitment opportunities are too limited.
 - ▶ Maybe $\{a, b\} \succ \{a\}, \{b\}$ but $\{a, b\} \succ \{a, b, c\}$ for some option c
- ▶ Studying temptation requires a richer language:
 - ▶ Approach here: study \succeq over a fairly rich set of “menus”
 - ▶ Using information on the entire ordering, develop a language of temptation.
 - ▶ Test how much can be learned by using this language.

Using menu choice to reveal temptation (1)

- ▶ In Toussaert (2018), I conduct a lab experiment in which:
 1. I elicit preferences over a set of menus $\{\{a\}, \{b\}, \{a, b\}\}$
 2. Then observe a choice from $\{a, b\}$ for some people
 3. Build a typology of agents based on this data
 4. Find about 25% of “self-control types” who:
 - (i) express $\{a\} \succ \{a, b\} \succ \{b\}$
 - (ii) choose a from $\{a, b\}$
- ▶ Here: use menu choice to study “real” temptations in the field.
- ▶ Do not observe choices from menus (Stage 2).
- ▶ Instead, focus on Stage 1: enrich the set of menus and explore further the notion of “type”.

Using menu choice to reveal temptation (2)

- ▶ Conduct a field study with a selected population: participants in a weight loss challenge.
- ▶ Study temptation to eat unhealthy by eliciting preferences over lunch reimbursement options differing in their food coverage.
- ▶ Using data on the entire ordering, develop revealed preference measures of temptation and validate them with survey data.
- ▶ Test whether those measures can predict self-control problems during the challenge:
 - ▶ Demand for and default on a goal setting contract
 - ▶ Likelihood of completing (a) the challenge and (b) the study
 - ▶ Likelihood of claiming reimbursement

What we are (not) going to learn

- ▶ What I will NOT show:
 - ▶ Restricting choice sets helps people lose weight.
 - ▶ Restricting choice sets helps people eat better.
- ▶ What I am hoping to demonstrate:
 - ▶ Language of menus rich enough to capture multiple facets of temptation.
 - ▶ Can talk about *source*, *strength* and *structure* of temptation.
 - ▶ Can predict self-control problems in at least related domains
 - ▶ e.g., revealed temptation to eat unhealthy = + 10-35 ppts more likely to default on contract.

Plan of the talk

1. Description of the dataset and experimental design
2. Revealed temptation in the reimbursement program
3. Survey validation of revealed preference measures
4. Predictive power of revealed preference measures

The subject pool

- ▶ Participants in a weight loss challenge conducted at NYU
- ▶ Data concerns the 2014 edition (4th edition)
- ▶ Only faculty and staff members eligible to participate
- ▶ 113 enrolled in study (out of 193)
- ▶ 35 y.o. and 79% female
- ▶ 31% of returning participants
- ▶ Large majority overweight:
 - ▶ Mean weight of 204 lbs (male) and 172 (female)
 - ▶ US average: 196 lbs (male) and 166 lbs (female)
- ▶ Average weight loss goal of 14.4 lbs

Rules of the challenge

- ▶ 8-week challenge (March - April 2014)
- ▶ In the spirit of the “Biggest Loser” TV show: winner is the one who loses highest % of body mass over the challenge.
- ▶ Monitoring: 4 weigh-ins, bi-monthly (March 4th, March 25th, April 15th, April 29th)
- ▶ Small prizes for losing highest number of pounds between 2 weigh-ins.
- ▶ Extra support:
 - ▶ Free gym pass for the month of March (private gym)
 - ▶ Four fitness and nutrition classes organized by NYU

Structure of the study and timeline

- ▶ Participants recruited at first weigh-in for study on improving health through exercise and nutrition.
- ▶ \$20 gift card for completing a two-part online study.

Online Survey	Completion Period	Survey Content
Survey 1 $N = 113$	March 4th through March 11th, 2014	Part 1: Basic socio-demographics Questions about participation Part 2: Goal setting contract Part 3: Reimbursement program
Survey 2 $N = 87$	April 29th through May 6th, 2014	Feedback questions about challenge and study Intertemporal choice tasks Self-control measures of Ameriks et al. (2007)

Goal setting contract

- ▶ Offer commitment contract to achieve self-set attendance goals.
- ▶ Participants could commit to goals in 1, 2 or 3 categories:
 - ▶ Gym visits (over one month)
 - ▶ Follow-up weigh-ins (out of 3)
 - ▶ Wellness events (out of 4)
- ▶ Lost their study payment (\$20) for not achieving them.
- ▶ Free-form entry initially - regressions will include dummy for completion date.

Description of the reimbursement program (1)

- ▶ I elicit participants' temptation to eat unhealthy by studying their preferred coverage in a lunch reimbursement program.
- ▶ Lunch reimbursement program over one month:
 - ▶ 10% of participants drawn at random at the end of the challenge.
 - ▶ Up to \$300 reimbursed for meals taken in April.
 - ▶ Had to bring their receipts to be reimbursed.
- ▶ Three food categories:
 - ▶ G = salads, soups, fruits, yogurts + water
 - ▶ O = hot and cold sandwiches, cereal bars + juice
 - ▶ R = burgers, pizza, fried foods, pastries + soda

Description of the reimbursement program (2)

- ▶ Participants asked to rank 7 reimbursements options:

$$\mathcal{M} := \{G, O, R, GO, GR, OR, GOR\}$$

- ▶ Elicitation of weak order \succeq on \mathcal{M} :
 - ▶ Participants assigned a rank number 1-7 to each option.
 - ▶ Could assign the same rank to multiple options to allow for indifferences.
- ▶ Incentive compatible elicitation procedure:
 - ▶ Probabilistic implementation with higher odds of receiving an option ranked higher.
 - ▶ Indifferences made easier to report.
 - ▶ Learned selected option after completing Survey 1.

Description of the reimbursement program (3)

- ▶ Participants asked to rate each food item on a 1-7 scale:
 - ▶ Health value (Survey 1, $N = 113$)
 - ▶ Temptation value (Survey 1, $N = 113$)

- ▶ For each food item, also asked to evaluate on 0-100 scale:
 - ▶ Actual consumption (Survey 1, $N = 113$)
 - ▶ Ideal “*should*” consumption (Survey 2, $N = 87$)
 - ▶ Unrestricted “*want*” consumption (Survey 2, $N = 87$)

Summary of dataset

For each participant, observe:

- ▶ Ranking \succeq of options in $\mathcal{M} := \{G, O, R, GO, GR, OR, GOR\}$
- ▶ Subjective ratings of food items and consumption ($N = 87$)
- ▶ Decision to enter goal setting contract
- ▶ Attendance of weigh-ins and wellness events
- ▶ Gym attendance data from two sources:
 - ▶ Electronic records from badge scans g_1 ($N = 69$)
 - ▶ Self-reports g_2 ($N = 83$)
 - ▶ Final measure $G = \min(g_1, g_2)$ ($N = 112$)
- ▶ Whether returned receipts for reimbursement

Revealed Temptation: Theory (1)

- ▶ A standard DM should weakly prefer GOR .
- ▶ In contrast, a DM who is tempted by a food category may prefer to eliminate it from the coverage.
- ▶ Idea of using menu choice to model temptation first formalized by Gul and Pesendorfer (2001) through Set Betweenness axiom

$$M \succeq M' \text{ implies } M \succeq M \cup M' \succeq M'$$

- ▶ Assume $G \succ R$. Three interesting cases allowed by the model:

$$(STD) \quad G \sim GR \succ R$$

$$(SC) \quad G \succ GR \succ R$$

$$(NSC) \quad G \succ GR \sim R$$

Revealed Temptation: Theory (2)

- ▶ Dekel, Lipman and Rustichini (2009) argue that SB is too restrictive.
- ▶ Example: DM with $G \succ GO \succ GR \succ GOR$
- ▶ Interpretation 1: Stochastic temptation
- ▶ Allowed if SB relaxed to Weak Set Betweenness (WSB)

If $\{x\} \succeq \{y\}$ for all $x \in M, y \in M'$ then $M \succeq M \cup M' \succeq M'$

- ▶ Representation

$$V(M) = \sum_{v \in V} p(v) \left\{ \max_{x \in M} [u(x) + v(x)] - \max_{y \in M} v(y) \right\}$$

- ▶ Self-control cost $c(x, M) = \max_{y \in M} v(y) - v(x)$ with prob $p(v)$
- ▶ GP 2001 case: $|V| = 1$

Revealed Temptation: Theory (3)

- ▶ Example: DM with $G \succ GO \succ GR \succ GOR$
- ▶ Interpretation 2: Cumulative temptation
- ▶ Allowed if SB relaxed to Positive Set Betweenness (PSB)

$$M \succeq M' \text{ implies } M \succeq M \cup M'$$

- ▶ Representation

$$V(M) = \max_{x \in M} [u(x) + \sum_{j \in J} v_j(x)] - \sum_{j \in J} \max_{y \in M} v_j(y)$$

- ▶ \neq self-control cost for temptation j , $c_j(x, M) = \max_{y \in M} v_j(y) - v_j(x)$
- ▶ GP 2001 case: $|J| = 1$

Revealed Temptation: Theory (4)

- ▶ The two models are non-nested, although one needs choices from menus to fully distinguish them.
- ▶ How well do they rationalize preferences in my dataset?
- ▶ How much can we separate them by solely relying on menu choice data?
- ▶ Too permissive? Or is Strict Set Betweenness too restrictive?

Revealed Temptation: Measurement (1)

Construct revealed preference measures of temptation to study *source*, *strength* and *structure* of temptation.

- ▶ *source*: What options do people eliminate from their choice set?
- ▶ *strength*: How systematically do they eliminate such options?
- ▶ *structure*: What temptation model is most consistent with their commitment choices?

Revealed Temptation: Measurement (2)

Test for the presence of temptation by looking at:

- ▶ The top choice: preference for a restricted coverage?
- ▶ Pairwise comparisons between 2 nested options:
 - ▶ *Global Temptation index* for G, O and R

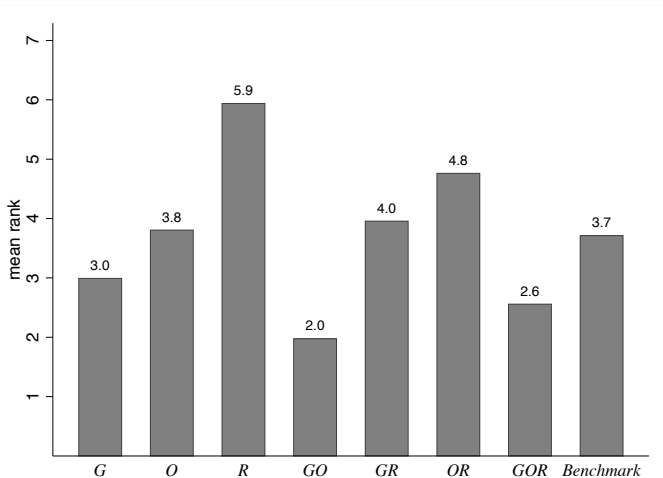
$$GT_{-R} = \sum_{\mathcal{M}_R} 1_{\{M \setminus \{R\} \succ M\}} \text{ where } \mathcal{M}_R \in \{GR, OR, GOR\}$$

- ▶ R is *globally tempting* if $GT_{-R} = 3$.
- ▶ Pairwise comparisons of non-nested options to test Set Betweenness and its relaxations.

Revealed Temptation: Findings

- ▶ 82.3% of strict orderings
- ▶ Will contrast findings with benchmark:
 - ▶ 1,000 random permutations of ranks for each individual
 - ▶ Allows to preserve the distribution of indifferences

Revealed Temptation: Mean rank



Revealed Temptation: Top choices

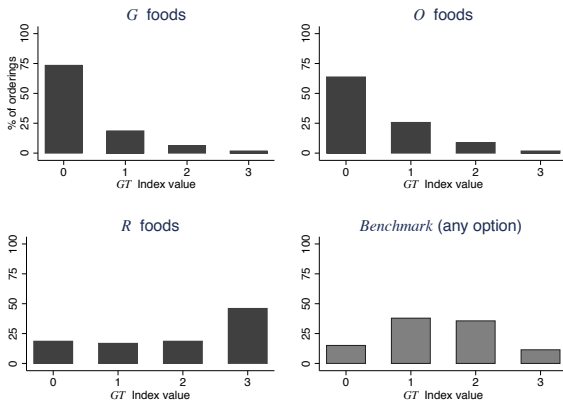
Table: Distribution of top choices

Top option(s)	Actual sample % (N)	Benchmark %	p -value
Option G	15.0 (17)	12.3	0.388
Option GO	32.7 (37)	12.3	< 0.001
Option GOR	31.9 (36)	12.3	< 0.001
Other option	6.2 (7)	48.9	< 0.001
No unique top	14.2 (16)	14.2	1.000
Total	100 (113)	100.0	100.0

Notes: “No unique top” if assigned rank 1 to several options; p -values from binomial tests that the observed frequency is equal to the benchmark frequency.

Revealed Temptation: Global Temptation Index

Figure: Temptation value of G , O and R foods



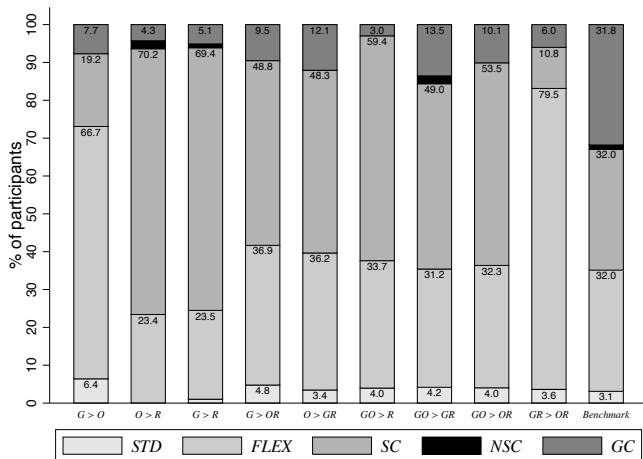
Notes: For each category $m \in \{G, O, R\}$, “Index value” refers to the value of the Global Temptation Index $GT_{-m} = \sum_{\mathcal{M}_m} \mathbb{1}_{\{M \setminus \{m\} \succ M\}} \in \{0, 1, 2, 3\}$.

Revealed Temptation: Structure of Commitment (1)

- ▶ *Standard (STD)*: $M \succ M'$ implies $M \sim M \cup M' \succ M'$
- ▶ *Flexibility (FLEX)*: $M \succ M'$ implies $M \cup M' \succ M \succ M'$
- ▶ *No Self-Control (NSC)*: $M \succ M'$ implies $M \succ M \cup M' \sim M'$
- ▶ *Self-Control (SC)*: $M \succ M'$ implies $M \succ M \cup M' \succ M'$
- ▶ *Global Commitment (GC)*: $M \succ M'$ implies $M \succ M' \succ M \cup M'$

Revealed Temptation: Structure of Commitment (2)

Figure: Distribution of orderings for most popular comparisons



Revealed Temptation: Structure of Commitment (3)

Comparing any two non-nested menus M and M' such that $M \succ M'$:

1. If $R \in M' - M \Rightarrow$ mostly *SC*

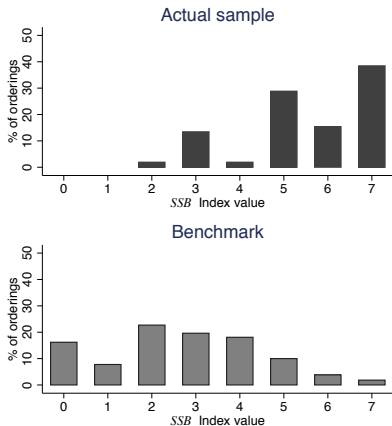
▶ e.g., (G, R) or (GO, R)

2. If $R \notin M \cup M'$ or $R \in M \cap M' \Rightarrow$ mostly *FLEX*

▶ e.g., (G, O) or (GR, OR)

Revealed Temptation: Structure of Commitment (4)

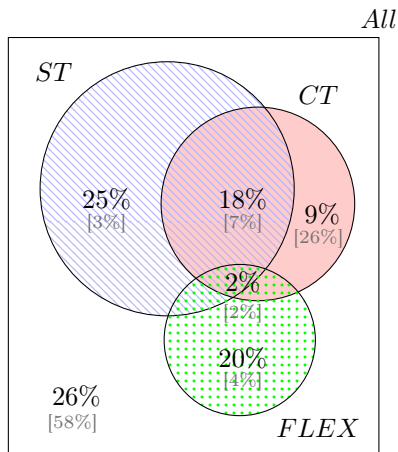
Figure: Strict Set Betweenness when R globally tempting



Notes: Index $SSB_{-R} = \sum_{\mathcal{P}_R} \mathbb{1}_{\{M \succ M \cup M' \succ M'\}} \in \{0, 1, \dots, 7\}$ where $\mathcal{P}_R = \{(G, R), (O, R), (G, OR), (GO, R), (O, GR), (GO, GR), (GO, OR)\}$.

Revealed Temptation: Structure of Commitment (5)

Figure: % classifiable with at most one violation



Notes: *Weak Set Betweenness (WSB)*: Stochastic Temptation; *Positive Set Betweenness (PSB)*: Cumulative Temptation

Revealed Temptation: Summary

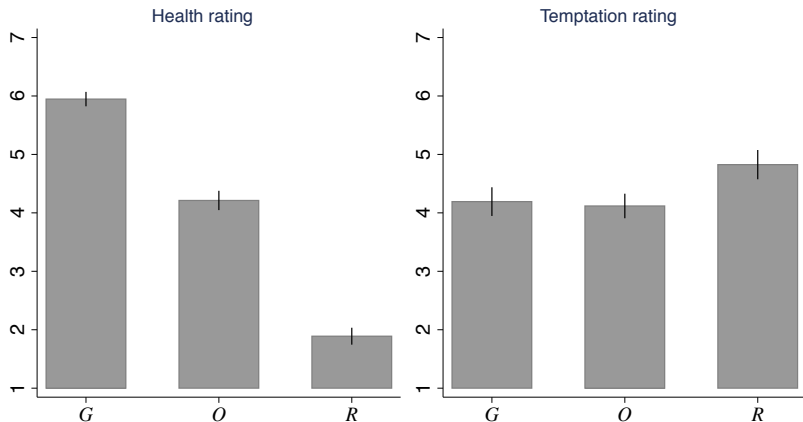
- ▶ Substantial demand for restricted coverage: only 39% of respondents assigned rank 1 to *GOR*.
- ▶ *R* is a robust temptation, but not *O*.
- ▶ Temptation by *R* takes the form of *SSB*.
- ▶ About 45% of respondents have temptation preferences à la DLR 2009 (*WSB* and *PSB*).

Survey validation of temptation measures (1)

- ▶ How well do those measures match respondents' perceptions of the food items?
- ▶ Are those who prefer to remove R from the coverage really tempted?

Survey validation of temptation measures (2)

Figure: Mean health and temptation scores by food category



Survey validation of temptation measures (3)

Figure: Temptation scores by food category and GT_R score

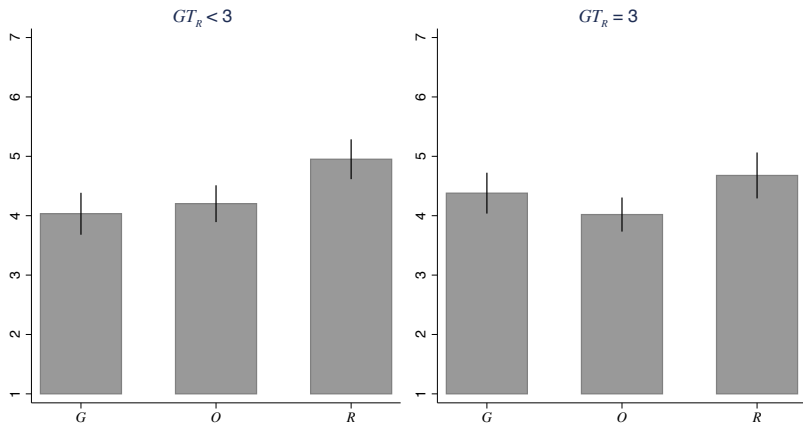
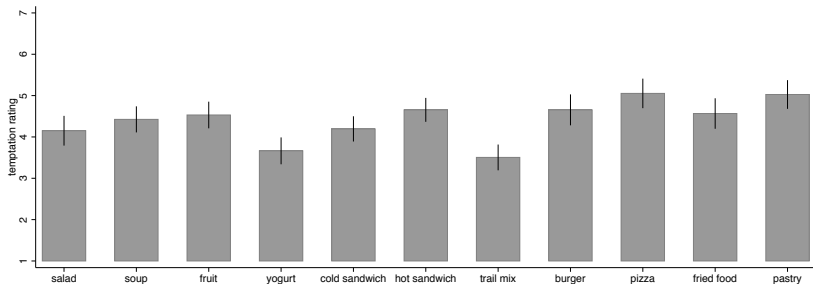


Figure: Temptation rating by food item

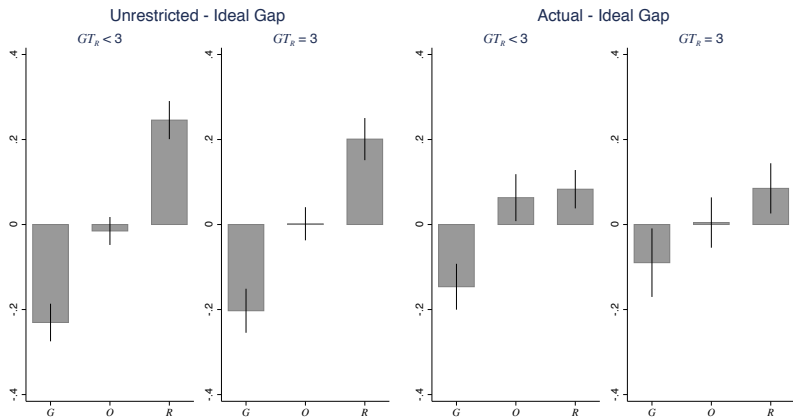


Survey validation of temptation measures (4)

- ▶ Asked 3 questions about food consumption (0-100 scale):
 - ▶ Actual: “*Since the beginning of the year, how often did you have each of the following options for lunch?*”
 - ▶ Ideal: “*Ideally, indicate how frequently you think you should consume each of the following food items*”
 - ▶ Unrestricted: “*Suppose you could eat anything you want without gaining a single pound and without any consequences for your health. How frequently would you eat each of the following food items?*”
- ▶ Construct indices of relative consumption frequency:
$$s_j(G) = \frac{f_j(G)}{f_j(G)+f_j(O)+f_j(R)} \text{ for } j \in \{A, I, T\}.$$
- ▶ Actual - Ideal gap = $s_A - s_I$ (in GP terms, $s_{u+v} - s_u$)
- ▶ Unrestricted - Ideal gap = $s_U - s_I$ (in GP terms, $s_v - s_u$)

Survey validation of temptation measures (6)

Figure: Consumption gaps by value of the GT_R score



Survey validation of temptation measures: Summary

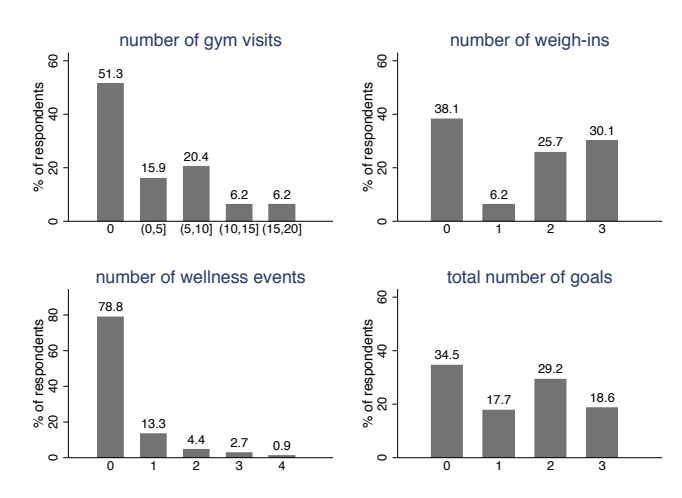
- ▶ Subjective ratings well aligned with commitment preference:
 - ▶ Clear ordering of G , O and R on health dimension.
 - ▶ R more tempting than G and O ; O not more tempting than G .
- ▶ Temptingness of R conflicts with consumption goals:
 - ▶ Unrestricted - Ideal Gap is large and positive for R .
 - ▶ Actual - Ideal Gap generally smaller, consistent with self-restraint.

Predictive power of temptation measures

- ▶ Built different types of menu choice measures:
 1. *Source*: Top choice (*G top*, *GO top*, *GOR top*, *Other*)
 2. *Strength*: GT_{-R} index
 3. *Structure*: SSB_{-R} index
- ▶ Commitment appears to reflect temptation concerns.
- ▶ Do those measures predict other behaviors likely symptomatic of self-control problems?
 1. Goal setting contract take-up and success
 2. Completion of challenge and study
 3. Reimbursement claims

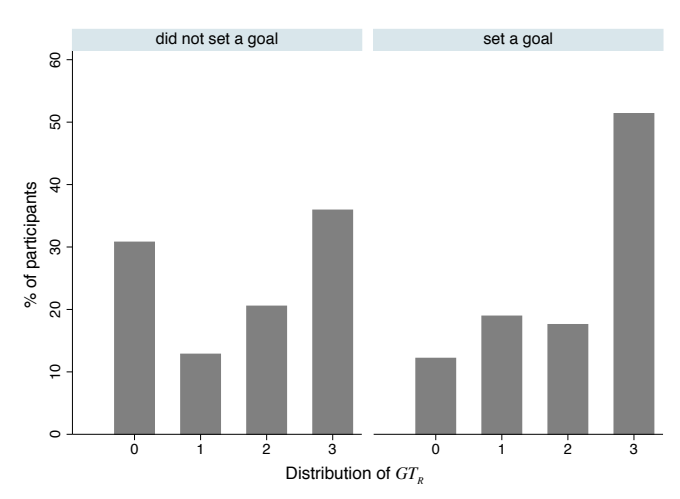
Goal setting (1)

Figure: Distribution of goals



Goal setting (2)

Figure: Revealed temptation and goal setting



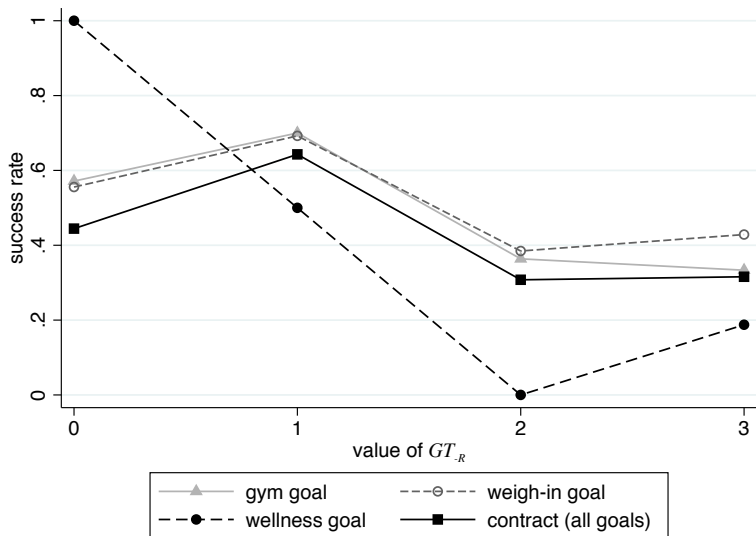
Determinants of contract take-up

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>G top</i>	0.031 (0.136)	0.049 (0.136)						
<i>GO top</i>	0.191* (0.109)	0.232** (0.107)						
<i>Other top</i>	0.057 (0.124)	0.039 (0.121)						
<i>GT_{-R}</i>			0.075** (0.037)	0.099*** (0.037)	0.103*** (0.038)			
<i>GT_{-G}</i>					0.034 (0.070)			
<i>GT_{-O}</i>					-0.014 (0.070)			
<i>SSB_{-R}</i>						0.026 (0.017)	0.038** (0.017)	0.050*** (0.018)
<i>SSB_{-G}</i>								-0.034 (0.041)
<i>SSB_{-O}</i>								-0.060 (0.044)
<i>female</i>		-0.094 (0.106)		-0.126 (0.105)	-0.122 (0.107)		-0.119 (0.107)	-0.153 (0.109)
<i>single</i>		0.051 (0.091)		0.087 (0.088)	0.089 (0.089)		0.079 (0.089)	0.071 (0.089)
<i>age</i>		0.007 (0.005)		0.008* (0.005)	0.007 (0.005)		0.008* (0.005)	0.009** (0.005)
<i>years of educ</i>		-0.092*** (0.029)		-0.097*** (0.029)	-0.098*** (0.029)		-0.097*** (0.029)	-0.100*** (0.029)
<i>prior participant</i>		-0.092 (0.091)		-0.073 (0.089)	-0.071 (0.090)		-0.082 (0.090)	-0.089 (0.090)
<i>weight loss goal</i>		0.005 (0.004)		0.004 (0.004)	0.005 (0.004)		0.004 (0.004)	0.003 (0.004)
<i>goal confidence</i>		0.421** (0.197)		0.385** (0.193)	0.386* (0.195)		0.417** (0.194)	0.421** (0.193)
<i>(goal confidence)²</i>		-0.048** (0.020)		-0.045** (0.020)	-0.046** (0.020)		-0.048** (0.020)	-0.048** (0.020)
<i>diets attempted</i>		0.006 (0.009)		0.005 (0.009)	0.005 (0.009)		0.006 (0.009)	0.006 (0.009)
<i>Day 1 decision</i>	-0.252*** (0.093)	-0.213** (0.088)	-0.270*** (0.091)	-0.232*** (0.085)	-0.230** (0.088)	-0.259*** (0.092)	-0.215** (0.087)	-0.193** (0.088)
<i>N</i>	113	113	113	113	113	113	113	113
<i>adj. R²</i>	0.069	0.184	0.092	0.213	0.199	0.078	0.196	0.201

Goal success

- ▶ High rate of contract default: 60.3% (44/73)
- ▶ Success rates by goal category:
 - ▶ Exercise goal: 46.3% (25/54)
 - ▶ Weight loss goal: 48.6% (34/70)
 - ▶ Wellness goal: 20.8% (5/24)
- ▶ But goal setters have higher attendance overall:
 - ▶ Gym visits: 7.5 vs. 4.5, $p = 0.017$
 - ▶ Weigh-ins: 1.6 vs 1.0 (out of 3), $p = 0.005$
 - ▶ Wellness events: 38% vs. 6% attend at least one ($p < 0.001$)

Success rates by value of GT_R



Determinants of goal success

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>G_top</i>	-0.341** (0.143)	-0.353*** (0.130)						
<i>GO_top</i>	-0.165 (0.135)	-0.243* (0.133)						
<i>Other_top</i>	-0.154 (0.149)	-0.112 (0.129)						
<i>GT_R</i>			-0.098** (0.048)	-0.120** (0.050)	-0.116** (0.050)			
<i>GT_G</i>					0.025 (0.080)			
<i>GT_O</i>					-0.044 (0.093)			
<i>SSB_R</i>						-0.029 (0.019)	-0.036* (0.018)	-0.059** (0.022)
<i>SSB_G</i>								0.038 (0.046)
<i>SSB_O</i>								0.111 (0.083)
<i>female</i>		-0.038 (0.104)		-0.005 (0.112)	-0.007 (0.113)		-0.000 (0.121)	0.044 (0.124)
<i>single</i>		0.054 (0.100)		0.042 (0.098)	0.042 (0.100)		0.067 (0.097)	0.079 (0.100)
<i>age</i>		-0.005 (0.005)		-0.008* (0.005)	-0.007 (0.005)		-0.009* (0.005)	-0.010** (0.004)
<i>years of educ</i>		0.036 (0.031)		0.048 (0.031)	0.045 (0.033)		0.034 (0.030)	0.053* (0.030)
<i>prior participant</i>		0.196* (0.105)		0.178* (0.103)	0.179* (0.103)		0.177* (0.105)	0.213** (0.105)
<i>weight loss goal</i>		-0.004 (0.003)		-0.003 (0.003)	-0.003 (0.003)		-0.002 (0.003)	-0.002 (0.003)
<i>goal confidence</i>		0.053 (0.278)		0.050 (0.274)	0.049 (0.274)		0.041 (0.256)	0.014 (0.266)
<i>(goal confidence)²</i>		-0.002 (0.029)		-0.001 (0.029)	-0.001 (0.029)		-0.002 (0.027)	0.001 (0.028)
<i>diets attempted</i>		-0.027** (0.011)		-0.026** (0.011)	-0.025** (0.010)		-0.025** (0.011)	-0.028** (0.012)
<i>Day 1 decision</i>	0.165 (0.125)	0.180 (0.113)	0.185 (0.127)	0.204* (0.115)	0.211* (0.116)	0.162 (0.127)	0.176 (0.118)	0.145 (0.130)
<i>N</i>	149	149	149	149	149	149	149	149
<i>adj. R²</i>	0.090	0.147	0.103	0.162	0.151	0.083	0.138	0.153

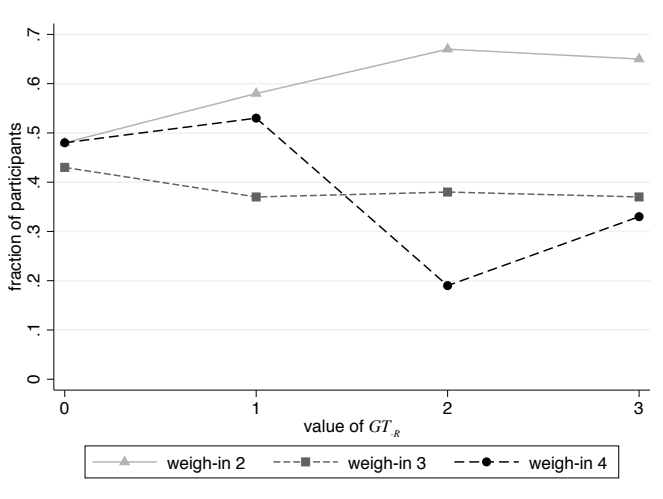
Link between menu preferences and other outcomes

Those tempted by R foods are also:

- ▶ Less likely to complete the challenge.
- ▶ Less likely to respond to Survey 2.
- ▶ Less likely to return receipts for reimbursement.

Weigh-in attendance over time

Figure: Attendance of the weigh-ins by value of the GT_{-R} index



Survey 2 completion

- ▶ 87 of the 113 study participants completed Survey 2.
- ▶ Attrition clearly non random as non respondents were:
 - ▶ more likely to have set goals (85% vs. 60%, $p = 0.019$)
 - ▶ less likely to attend 3 follow-up weigh-ins ($p < 0.05$ for all 3)
 - ▶ less likely to attend the gym (2.7 vs. 6.7 visits, $p < 0.01$)
- ▶ But not necessarily less motivated ex ante: similar intentions to attend gym, weigh-ins and wellness events.
- ▶ Menu preferences predict likelihood of responding, controlling for goal setting and attendance of last weigh-in:
 - ▶ Ranking G at top: ≈ 20 pp less likely to respond
 - ▶ Extra point on GT_{-R} score = ≈ 7 pp less likely to respond

Claiming reimbursement

- ▶ Only 17% (19) submitted receipts
- ▶ Reasons:
 - ▶ Lost receipts, forgot to ask, could not get itemized: 41%
 - ▶ Not worth effort given low chance of winning: 34%
 - ▶ Usually brings own lunch: 26%
 - ▶ Option did not cover foods ordered: 10%
- ▶ Correlated with showing up to last weigh-in and option assigned.
- ▶ Menu preferences predict likelihood of responding, controlling for option assigned and attendance of last weigh-in:
 - ▶ Ranking *GO* at top: \approx 20pp less likely to submit receipts
 - ▶ Extra point on $GT-R$ score = \approx 6-7pp less likely to submit receipts

Conclusion

- ▶ Find strong evidence of commitment demand driven by temptation as measured through menu preferences.
- ▶ Revealed preference approach, structural, more agnostic and comprehensive than in previous studies.
- ▶ Related to take-up and default on goal setting contract.
- ▶ Menu preference measures of temptation offer a promising venue to measure self-control problems.

Discussion (1)

To what extent does menu choice “reveal” temptation? Identifying temptation with preference for commitment might be

- ▶ **Too weak** i.e., there can be commitment without temptation:
 - ▶ Distribution of GT_{-R} almost unchanged with tighter definition.
 - ▶ Most findings remain in subsample with $G, O \succ R$.
 - ▶ Restriction on singletons might be too strict. Example:
 $GO \succ GOR \succ R \succ G$.

- ▶ **Too strong** i.e., there can be temptation without commitment:
 - ▶ Commitment requires sophistication from tempted DM.
 - ▶ But fairly sophisticated subjects: entered weight loss challenge.
 - ▶ Full naiveté unlikely when it comes to food cravings.

Discussion (2)

- ▶ Awareness of self-control problems on the extensive margin.
- ▶ But misunderstanding of the intensive margin of self-control:
 - ▶ Those who preferred to avoid R were more likely to take up the goal setting contract.
 - ▶ But they were also more likely to fail to reach their goals.
- ▶ Correlation of commitment demand across domains:
 - ▶ (Beliefs about) self-control may have some domain generality.
 - ▶ Other candidate explanations: signaling/experimenter demand.
- ▶ Subject pool and choice environment quite specific: more research needed to test robustness of findings to other settings.