

An experiment on temptation and attitude towards paternalism

Michał Krawczyk[†]

University of Warsaw

Łukasz Woźny

Warsaw School of Economics

Abstract: In this project we investigate experimentally the link between self-control and attitude towards paternalism in a principal-agent framework. This allows us to distinguish between models of costly self-control and models of time inconsistency, which often make identical predictions in other contexts. We invite our subjects for a free lunch: a burger or a turkey. We verify in a pre-test that the burger is considered (much) more tasty and tempting, while the turkey is seen as healthier. In the experiment proper we observe incentivized choices of four types: what menus (must eat burger; must eat turkey; your choice: burger or turkey) subjects assign to another; how they reward each of these menu choices yet another participant made for them; which of the two dishes they pick on the spot (if given the choice); whether they want to pre-commit to a choice of dish for a future session. Similarly to some recent experimental results we find a significant fraction of subjects willing to self-commit. We also observe non-trivial sets of individuals who reward highly a restricted choice and paternalistically restrict other's choice. Moreover, there is a strong link between these three tendencies, suggesting a common thread underlying the use of commitment devices and paternalistic behavior as well as approval thereof in environments involving temptations. These findings are consistent with the models of costly self-control rather than time inconsistency.

“I generally avoid temptation unless I can't resist it.” May West

1. Introduction

Self-control problems (SCP) represent a significant divergence from standard economic models. They allow for a distinction between what is good for the agent (what she *should* do) and what she ends up doing (what she *wants* to do). The *should* choices, such as learning, working, saving, exercising and dieting, tend to involve immediate costs and larger long-run benefits. The *want* choices, such as leisure, overspending, overeating, drinking and smoking, offer immediate gratification and often substantial deferred costs. Many people tend to go for the *wants* more often than they had planned to and more than they agree they should and feel comfortable with. They are also prone to regretting them. As a result they may be willing to restrict their future choices, even at a cost, to the *should* option (self-commitment). Important consequences for welfare analysis and policy recommendations arise. In particular, whereas in the standard model more choice is always weakly

[†] Corresponding author, mkrawczyk@wne.uw.edu.pl. Helpful comments from participants of seminars at the University of California Santa Barbara and the Warsaw School of Economics, as well as the 6th Hurwicz Workshop in Warsaw are gratefully acknowledged. Aleksandra Bednarczuk, Natalia Starzykowska and Marta Sylwestrzak provided research assistance. The project was funded by DEC-2013/11/D/HS4/03813. Usual disclaimers apply.

preferable, having some options removed from the menu may be beneficial for an agent experiencing SCP.

If so, various forms of *paternalism* may be advocated, whereby the “sinful” choices are made less salient, discouraged, punished, or altogether removed from the menu. Interestingly, however, there is very limited evidence that support for paternalism tends to be triggered by SCP and associated with willingness to self-commit. Indeed, but a few studies have tried to identify a link between the two and the results seem mixed. Yet, given that limited rationality and self-control problems continue to be presented as a major factor justifying paternalism, confirming or disproving the association between the attitudes towards choice limitation being imposed by self and others appears important. Indeed, many bans, restrictions and taxes aimed at reducing consumption of certain goods seem to be largely driven by belief in SCP. For example, *externalities* associated with drinking alcohol are probably smaller in 12 year olds (who, even when drunk, less often drive cars or carry guns and knives etc.) than 22 year olds; if the latter are allowed to drink while the former are not, it is chiefly due to (perhaps controversial) belief that 22 year olds can better recognize what is good for them. Likewise, smoking marijuana is illegal in many countries, although its externalities are probably much more benign than those of alcohol or even cigarettes. Typically, those who oppose legalization justify it by mentioning marijuana’s addictive potential and harm to the user, not other people (Pew Research Center, 2015). These individuals thus seem to appreciate that the authorities restrict consumption possibilities, by excluding products that are tempting but possibly harmful.

On the other hand, many temperance movements, such as Poland’s Crusade for the Liberation of Man call for voluntary abstinence but not law-enforced restricted access (except for minors). Clearly, the proposition that it matters little whether the set of available options is restricted by self or another cannot be taken for granted; an individual may take pride in self-committing, say, not to drink any alcohol but find it annoying if a nanny state limits access to it. Support for paternalism and specifically for taking some options away from the menu may arise from other considerations and psychological tendencies than SCP. Externalities aside, the belief that the patron may be better informed, the cost of acquiring information, the aversion to choice and the sense of belonging and submission to authority are but a few examples.

Studying menu restrictions imposed by others may also be helpful in distinguishing between the main theoretical frameworks that have been offered to account for self-commitment. The main question is how an individual will react to a restriction. Assuming some minimum reciprocity (which is typically found in most individuals, Fehr and Gächter, 1998), we can infer from such a reaction how she feels about the restriction. If the models of *time-inconsistent preferences* (Strotz, 1956) are correct, adding a tempting object to the menu may lead to it being chosen by the myopic *wanting* self, against the individual’s long-run interests. If thus a choice restriction is made (synchronously) by another, it probably frustrates the myopic self, so a negative reaction is expected. By contrast, models of *costly self-control* (Gul and Pesendorfer, 2001) propose that even if a temptation is resisted, it may come with a non-trivial psychic cost. Thus, it is desirable that it be removed and a positive reaction to such a restriction is predicted.

One could wonder why this indirect method of inferring attitude towards a choice set would be necessary. The main reason is that if no other player is involved, it is difficult to make a clear distinction between the choice *of* the menu and the choice *from* the menu. In fact, why would

resisting the temptation to include the “sinful” object in the menu be easier than not picking this object once it is already on the menu? In a typical experimental study, this separation is of temporal nature: the choice of the menu takes place immediately, while the choice from the menu is postponed. In such a case, the models of temporal inconsistency give very similar predictions to the models of costly self-control. Moreover, one may wonder if such a separation (which, for practical reasons often only means a few minutes of waiting) should matter at all.

In this project we run an incentivized experiment with natural stimuli (food) in which we give the same participants the opportunity to self-commit to future consumption, restrict others’ choices, as well as react to such restrictions. This allows us to separate preferences over menus from preferences over items in a menu in an incentive compatible way. We find substantial heterogeneity, with non-negligible minorities that restrict own choices, restrict others’ choices and reward restrictions imposed on their own choices. Moreover, these three groups largely overlap. These findings are most consistent with the models of costly self-control. As an additional manipulation, we investigate the impact of visceral state (hungry vs. satiated) on willingness to self-commit and find no difference.

2. Previous literature

SCP were addressed in the empirical economic literature when the so called “preference reversals” were observed. The term pertains to the situation, in which a decision maker plans to choose an alternative a over alternative b in the future, but reverses her preferences (chooses b over a) when the choice actually is to be made¹. Since the seminal contribution of Strotz (1955-1956), such observations have been modelled taking preference change as a primitive. In particular, researchers have distinguished between current (period 1) preference (denoted by \succeq_1) and future (period 2) preference (\succeq_2). Then the “reversal” observations were explained by time-inconsistent preferences: these of period 1 self, say $a \succ_1 b$, and these of period 2 self: $b \succeq_2 a$. A typical representation involves models of hyperbolic discounters, i.e. decision makers whose discount factor (between two future dates) changes over time. In such a non-rational preference relation, one distinguishes *sophisticated* from *naïve* decision makers. Specifically, a sophisticated decision maker anticipates her future change of heart, and is compelled to choose *the best plan but only among those that would be actually followed* by her period-2 self. Hence, a sophisticated decision maker will not choose a plan that the period 2 self would like to change. If given a game-theoretic interpretation (Phelps and Pollak, 1968; Peleg and Yaari, 1973), this plan is analogous to a subgame-perfect Nash equilibrium of the two-stage game between self 1 (planner) and self 2 (doer); but see Caplin and Leahy (2001) for a detailed comparison. It can be shown that the sophisticated decision maker would like to pre-commit (if possible) to some decision plan, by e.g. investing today to restrict her own second-period options. Later O’Donoghue and Rabin (1999) also considered the case of a naïve decision maker, who optimizes consecutively, unaware of her future preference change².

¹ Importantly, it is the case even if no new information was acquired. Moreover, as documented, many subjects do not seem to learn (and make consistent choices) even if the choice situation repeats, hence the “preference reversal” is somehow persistent.

² Interestingly, Heidhues and Koszegi (2009) argued that sophistication does not need to generate higher welfare than naivette.

More recently, Gul and Pesendorfer (2001, GP henceforth) and Dekel, Lipman and Rustichini (2001) rationalized preference reversals in a two-period model involving \succeq preferences defined over menus (subsets) of the decision/consumption sets, from which a consumption choice is subsequently made. Preference reversal is obtained if $\{a\} \succ \{a,b\} \sim \{b\}$, indicating that a decision maker prefers singleton menu $\{a\}$ over $\{b\}$, but if facing a choice of an item from the menu $\{a,b\}$ would succumb to temptation and choose b . This is similar to implications of the Strotz model, and sometimes referred to as an overwhelming temptation case, as a decision maker always succumbs to temptation. However, GP model also allows for preference ordering $\{a\} \succ \{a,b\} \succ \{b\}$, corresponding to a decision maker that exhibits some self-control. That is, when facing a menu $\{a,b\}$, he manages to resist and choose alternative a , but at the same time incurring (self-control) cost associated with foregoing the tempting item b . GP obtain their utility representations result (on the set of menus over lotteries) assuming the so-called *set betweenness* axiom³.

More generally, the models specifying preferences over menus also allow to rationalize various other preferences, where set-betweenness axiom is not satisfied. For example, Kreps (1979) studied preferences for flexibility, which allow for relations $\{a,b\} \succeq \{a\}$ and $\{a,b\} \succeq \{b\}$. These could be interpreted as preferences of a decision maker, who is not certain about his future preferences and as a result would not like to pre-commit to any specific choice. By contrast, Sarver (2008) and Kopylov (2012) considered preferences allowing for anticipation of regret or guilt. Using our notation, these can be exemplified by the following ordering: $\{a\} \succeq \{b\} \succeq \{a,b\}$. In such a case fewer options are preferred, as a decision maker may experience guilt when choosing the tempting item b , or may feel regret if his choice is ex-post inferior. Finally, the standard rational decision maker can be represented by $\{a,b\} \sim \{a\} \succeq \{b\}$, that is, any menu is as good as its best alternative.

Recently, some papers have addressed the question of temptations and costly self-control in a principal-agent framework, which is relevant for our design. These include models of the principal, who chooses the optimal contract to screen the naïve agents displaying dynamically inconsistent preferences (Eliaz and Spiegel 2006). Heidhues and Koszegi (2010) model credit markets, in which borrowers have a taste for immediate gratification. Della Vigna and Malmendier (2004) characterize the optimal contract design for (partially) naïve agents with time-inconsistent; while Esteban and Miyagawa (2006) characterize the optimal menu pricing, when consumers face temptation. Finally, the papers by Gilpatric (2008), Yilmaz (2013) and Woźny (2015) characterize the optimal incentive scheme for naïve and sophisticated agents with time-inconsistent or GP preferences.

Economists have also sought obtain to testable implications of models of intertemporal choice (or revealed time-preferences). Specifically, Montiel Olea and Strzalecki (2014), Dzielwski (2015) as well as Echenique, Imai and Saito (2016) provided axiomatic characterizations of discounted utility models including exponential and hyperbolic discounting as special cases, as well as design experimental procedures to test them. These models, while allowing for a hyperbolic discounting specification, also assume that decision maker is time-consistent. Hence, the problem of preference reversal

³ More recently, generalizations of *set-betweenness* axiom have been proposed allowing, among other, for random temptations (Dekel, Lipman, Rustichini 2009); dynamic temptations (Gul and Pesendorfer 2004, Noor 2007); or choice-dependent temptation cost (Olszewski, 2011). More on links between time (in)consistency models (such as Strotz 1956) and models of costly self-control (such as GP) can be found in Dekel and Lipman (2012).

(involving both: nonstationarity and time-inconsistency) cannot be tested (or identified) using their setups⁴.

Deriving testable implications of the discounted utility models allowing for preference-reversal is problematic (see Janssens, Kramer and Swart (2015) for a discussion). However, a number of papers tried to investigate this phenomenon in an experimental or empirical setting (see Bucciol, 2012). Some of these studies focused on the prevalence of self-commitment, typically finding up to one-third of participants willing to utilize such devices (Ashraf et al. 2006, a field experiment in a Philippine bank; see also Gine et al. (2016); Casari 2009, a lab experiment on intertemporal choice). Others, including Augenblick, Niederle, and Sprenger (2015) documenting results of an experiment concerning demand for commitment in real effort tasks.⁵ See also Bitterly (2014) for a more comprehensive literature overview.

As mentioned before, empirically distinguishing between the two main different psychological mechanisms considered in theoretical literature (time-inconsistency vs. GP), both of which could drive demand for commitment, is particularly difficult. One attempt has been recently made by Toussaert (2014), who used a partly random allocation of menus, so that she could observe predictions and actual choices of those, who did not want to have a choice in the first place. In line with models of costly self-control, she found that even out of individuals who correctly expected to resist the temptation, many preferred not to face it at all. Very recently, isolated studies started investigating how individuals facing possible temptation react to having their menu restricted *by somebody else*. Kataria et al. (2014) investigated subjects' reactions to restrictions imposed on their menu of alternative monetary gambles. Overall, they punished such restrictions *unless they knew that the gamble failed* (hindsight bias).

Most closely related to our approach are studies that investigate both paternalism and self-commitment. Laboratory subjects of Le Lec and Tarrow (2015) were asked to make a series of binary comparisons between menus (involving one or more websites that could be explored during the session). They did so first for themselves and subsequently – for another participant, knowing his or her preference. The authors found that in the latter task, most subjects were willing to offer larger menus, even if this did not improve the utility of the most-preferred item on the menu. By contrast, when choosing for self, they typically preferred having unwanted objects removed from the menu. The study used between-subject design, so little can be said of the correlation between the two types of menu choices. Uhl (2013) paid his subjects for showing up for a session early in the morning. Interestingly, while the proportions of subjects willing to impose a commitment on self and others were similar, there was no correlation between the two tendencies. By contrast, in a survey of Danish students, Pendersen et al. (2014) found some link between self-control and support for strong (such as choice restrictions) but not weak (such as nudges) forms of paternalism.

Finally, as related to our experimental manipulation of hunger, studies found that the more caloric, less healthy *want* food is more likely to be chosen for future consumption when subjects are currently hungry (Read and Van Leeuwen, 1998) but we are not aware of a study on the impact of hunger on demand for commitment.

⁴ See also Adams, et al. (2014) studying time-consistency problem taking into account a collective level of household decisions.

⁵ Here working over-time resulted from dynamic inconsistency.

3. Theoretical framework

In this section we propose a theoretical framework for our experiment. Following GP, consider consumer i with preferences over menus (say M) given by the utility:

$$U_i(M) = \max_{x \in M} \{u_i(x) + v_i(x)\} - \max_{t \in M} v_i(t),$$

In this notation, u_i stands for commitment utility and v_i for temptation utility. When choosing an item from the menu, the agent uses preferences $u_i + v_i$, and when evaluating menus, he uses U_i . The preferences over items in the menu and preferences over menus are hence directly linked. This implies that “current” preferences (over items) and “future” preferences (over menus) are consistent. The term $\max_{t \in M} v_i(t) - v_i(x)$ is called the cost of self-control. A limiting example, where the cost of self-control is overwhelming, is given by the Strotz formulation. Here a game-theoretical or time-inconsistency interpretation of the Strotz preferences would follow with the first period (planner) agent having preferences u_i (over tomorrow's alternatives) and the second (doer) preferences v_i . That is, at any point in time the agent has “current” preferences given by v_i and “planning” preferences given by u_i . A sophisticated agent foresees his preference change and adopts only such plans that are time-consistent, i.e. will actually be followed by the second period agent. For this reason, at any point in time the agent chooses items accordingly to preferences v_i but when earlier evaluating such future plans or menus, he solves the following problem:

$$\max_{x \in M} u_i(x) \text{ st. } v_i(x) \geq v_i(y) \text{ for all } y \in M.$$

Here, the decision maker maximizes utility u_i but only on such elements that will be followed by agent possessing preferences v_i . This implies that current preferences over menus are given by $U_i(M) = \max_{x \in M} v_i(x)$, while preferences over future menus are evaluated using $\max_{x \in M} u_i(x) \text{ st. } v_i(x) \geq v_i(y) \text{ for all } y \in M$. Hence, there is the inconsistency between the current preferences (over items and menus) and the future ones (over items and menus).

This implies among other that, in the GP model agent may or may not succumb to temptation, while in the Strotz model agent always succumbs to tempting (second period) alternatives, if tempting alternatives are in the menu. Next, if a tempting object decreases utility of an agent as proposed by GP, having it removed by the principal should make the agent grateful. If the agent displays some positive reciprocity, she should be willing to reward such restrictions, even it involves a (sufficiently low) monetary cost. By contrast, agents suffering from dynamic inconsistency may be willing to reward the principal who restricts their future, but not immediate options. Studying choices in a principal-agent setting allows disentangling preferences for menu restriction (done by the others) from preferences and actual choices from the menu, that was actually implemented.

To assess the willingness to reward a restricted choice set by some other subject we refer to the recent models of kindness and reciprocity. Cox, Friedman and Gjerstad (2007) defined⁶ reciprocal preferences using a parameter measuring the marginal rate of substitution between own and others wealth. This parameter depends among others on reciprocity, as defined by the difference between maximal payoff the player can guarantee himself and some other, say benchmark, payoff. Hence, the model of Cox et al. gives a simple definition of reciprocal preferences based on actually obtained

⁶ Falk and Fischbacher (2006) propose an equilibrium model of reciprocity based on psychological games. A model of Cox, Friedman and Gjerstad (2007) departs from game theoretical explanations, hence is not based on the higher order (equilibrium) beliefs.

outcomes scaled by reciprocal parameter. As such, in our framework, it implies that if agents differentiate their willingness to reward a given choice set, then their reciprocity parameter must vary sufficiently between obtained menus.

Concentrating on a specific functional form we now embed temptation-driven preferences into models of reciprocity. Towards this end, we follow Cox et al. (2007) and consider preferences with reciprocity given by utility F :

$$F(m, y) = my^\theta,$$

where $\theta \in (-1, 1)$ is a reciprocity parameter, while m and y (assumed to be positive) are self and other subject's utility. Parameter θ measures the strength with which the other agent's utility enters my preferences. We refer to θ as a reciprocity parameter as (in the moment) we will let it depend on the choice of the other agent. Specifically, we set m and y as quasi-linear utility over menus and money:

$$m(M, w_m) = U_m(M) + w_m \text{ and } y(M, w_y) = U_y(M) + w_y.$$

Suppose that (as is actually the case in our experiment) each agent is endowed with some (nonnegative) income w . Conditional on the obtained menu M_m we let the agent split his income between nonnegative w_m and nonnegative w_y , i.e. part of income for herself and for the other subject. There is a multiplier effect, i.e. positive parameter α measuring the slope of exchange between w_m and w_y . This implies that the budget constraint is given by $w_y = \alpha(w - w_m)$. We let agent i have GP or Strotz preferences (as specified above), but if he has standard, rational preferences without temptations, we would take $U_i(M) = \max_{x \in M} u_i(x)$.

We now derive the optimal choice in our model. Clearly, if $\theta = 0$, then the optimal choice is $w_m = w$, as the agent is not reciprocal. Now assume that reciprocity parameter is specified by the function:

$$\theta(M_m) = \frac{U_m(M_m) - u_0}{U_m^{max} - U_m^{min}},$$

where u_0 is the benchmark utility level and U_m^{max} is the utility from the best menu. Similarly, we denote the worst case by U_m^{min} . Our benchmark will be the utility from the worst menu. Hence in our case $\theta \in (0, 1)$. Then, the optimal choice of reward w_m is computed from the first order condition, yielding:

$$w_m = \frac{U_y(M_y) + \alpha(w + U_m(M_m))}{\alpha(1 + \theta)} - U_m(M_m),$$

whenever interior, and we set $w_m = w$ or $w_m = 0$ in the corner cases.

Denoting by C (commitment) the item preferred according to preferences u_m and by T (temptation) the item preferred according to v_m , we obtain that for GP preferences or preferences with no temptations: $U_m^{max} - U_m^{min} = u_m(C) - u_m(T)$, but if agent has Strotz preferences then $U_m^{max} - U_m^{min} = v_m(T) - v_m(C)$.

Clearly, θ is monotone in $U_m(M_m)$ and so is interior w_y (for $\theta \in (0, 1)$).

To evaluate the willingness to pay for the menu restriction we now define a bonus for free choice, i.e. a difference between payments w_y^{FC} rewarding free choice and w_y^C rewarding the singleton menu with the commitment choice only:

$$w_y^{FC} - w_y^C = \alpha(w_m^C - w_m^{FC}) = \alpha \left[\frac{\theta(FC)U_m(FC)}{1+\theta(FC)} - \frac{\theta(C)U_m(C)}{1+\theta(C)} \right] - (U_y(M_y) + \alpha w) \left[\frac{1}{1+\theta(FC)} - \frac{1}{1+\theta(C)} \right].$$

Observe that, if $U_m(FC) = U_m(C)$ then bonus for free choice is zero, while if $U_m(C) > U_m(FC)$ (the GP case) then bonus for free choice is positive⁷. Finally, for the Strotz case letting $U_m(M) = \max_{x \in M} v_m(x)$ we obtain that $U_m(FC) > U_m(C)$ and so the bonus for free choice is negative.

Thus, the bonus for free choice depends, among other things, on the difference of utilities between menus FC and C while the amount of transfer w_y depends on the difference between utility of obtained menu and some benchmark utility.

Each decision maker also plays a role of the menu setter for some other agent. Specifically, in our model he chooses M_y . If $\theta > 0$, then optimal choice of the menu should maximize the other agent's utility. This requires knowledge of U_y . In the case U_y is expected to be identical with U_m , the menu setter may choose a menu that is best for him, i.e. FC or C. If, on the other hand, no prior knowledge about U_y is delivered, then FC may be considered as the best alternative.

4. Pre-test

In order to verify that our intended stimuli were appropriate (that is, that there is indeed room for temptation), we run an online pre-test with 135 individuals registered in the subject pool of the University of Warsaw Experimental Economics Lab (UWEEL). The English transcript is provided in Appendix B. Figure 1 shows that while subjects were roughly equally split between the dishes when choosing for immediate consumption, they strongly preferred Turkey in the long run as well as for a loved one. The differences between the distribution of responses to the "to eat now" question and those for any of the two remaining questions are obviously highly significant.

⁷ Here with a slight abuse of notation we use C for a single item menu of the commitment alternative and T a single item menu with temptation only.

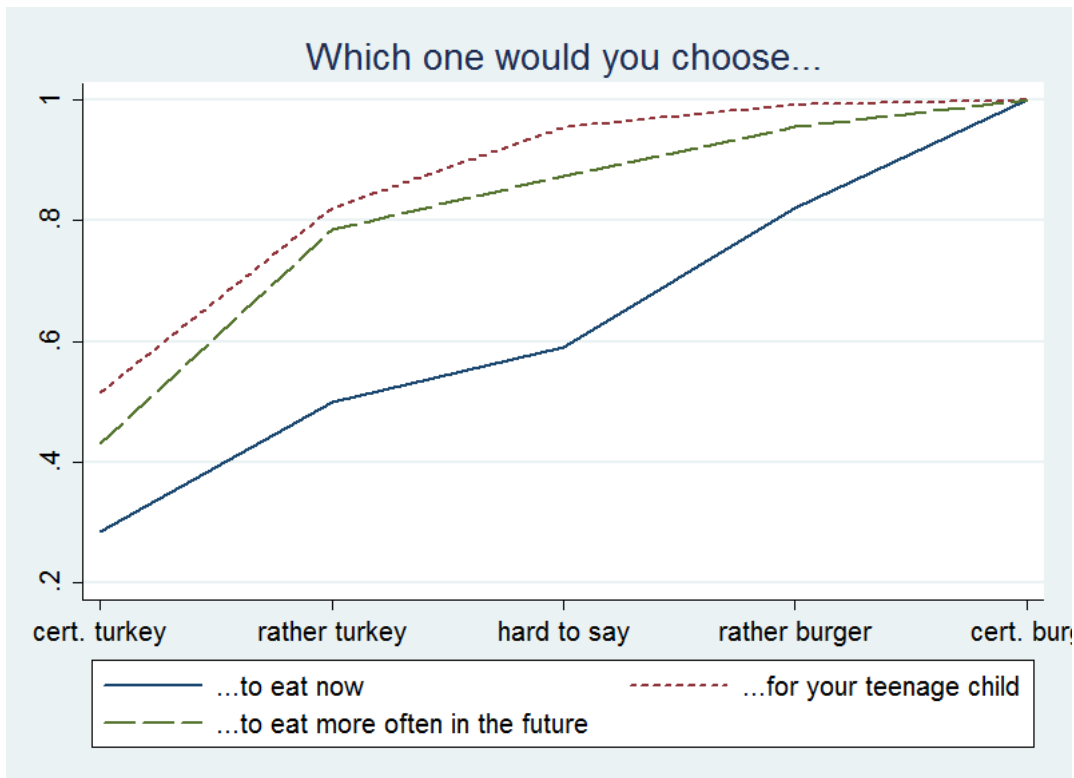


Figure 1. Hypothetical choices in the pre-test: cumulative distribution functions

Pre-test participants were also asked to compare the two dishes on several dimensions. Table 1 shows summary statistics and Figure 2 provides distribution of our key variable: “Which dish is more tempting?”. Clearly, the Burger is the unhealthy, “sinful”, tempting option.

Table 1. Pre-test: Which dish is more... (1=definitely turkey, ... , 5=definitely burger)

	healthy	eco-friendly	tasty	attractive	satiating	tempting
Mean	1.33	2.01	3.33	3.31	3.67	3.62
Median	1	2	3	4	4	4

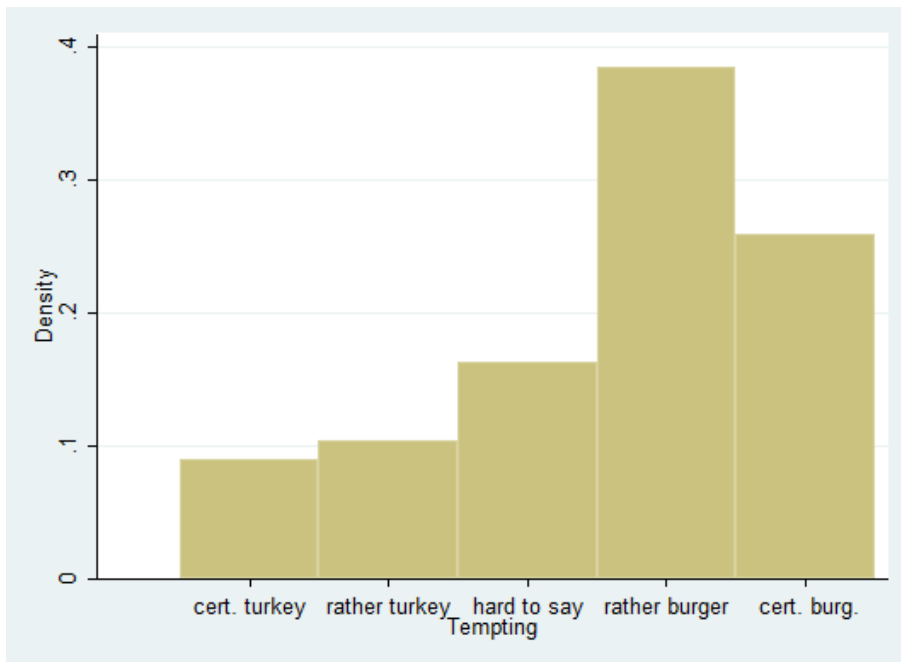


Figure 2. Pre-test: burger is more tempting

5. Design and procedures

The sessions of the experiment proper were conducted at various university buildings in Warsaw, Poland, always at lunch time (broadly speaking). Participants were recruited by posters, flyers, student Facebook groups and e-mails sent out to the subject pool of UWEEL (excluding individuals asked to participate in the pre-test). They were informed they would receive a free lunch and earn a few zloty, provided they fill in a questionnaire.

In the end, 199 individuals took part, of which 52% were male. Mean age was 23.8 years. About 90% were students, of which about three quarters studied economics.

Upon arrival, subjects were invited to inspect the two dishes and read descriptions identical with those used in the pre-test. Photographs thereof were also projected on a screen. Subjects were informed (see Appendix C for complete instructions) that there were three *menus*: 1. {Burger}, 2. {Turkey} and 3. {Your choice: Burger or Turkey}. The first two will henceforth be referred to as *restricted menus*. They were asked to choose one of the menus for another participant (to which we shall refer as “choice of menu” by a “menu setter” for a “chooser”). They were also told to reward each of these three choices made for *them* by yet another subject prior to learning which one she actually chose (strategy method). For this purpose, they were endowed with 5 zloty (ca 1 euro) and told a reward could be anything between 0 and 5 zloty and would be tripled for the recipient. For example, if a reward of 2 zloty was envisioned for menu {Burger} and this menu was indeed given, the chooser would pay 2 zloty out of her endowment of 5 and the menu setter would receive $3 \cdot 2 = 6$ zloty. The high multiplier was used to encourage non-zero gifts, thereby facilitating comparisons between menus. We will refer to these decisions as “choice of rewards”. In total, they could thus earn up to $15 + 5 = 20$ zloty (when their choice was maximally rewarded while they did not reward at all). They were also asked to predict what they would choose if they had a chance to choose, i.e. received Menu 3 {Your choice: Burger or Turkey}. We call it “prediction of choice”. Finally, if they actually received this menu, they were asked to choose Burger or Turkey (“choice from the menu”).

Because any subject was matched with their peers coming shortly before or after them, they would typically find out very soon which dish was obtained (that would normally be consumed on the spot) and how much was earned. Subjects were also told that a similar session would take place soon and asked if they wanted to pre-commit to any specific dish. Finally, demographic questions and questions related to dieting habits, drinking and smoking were distributed. Subjects would indicate how much they enjoyed their meal, take their earnings and leave.

Sixteen different version of the questionnaire were used in a 2 (order of dishes) X 2 (order of choices) X 2 (numbers in the example) X 2 (commitment before or after consumption) design. First, some subjects always had the Burger listed first, the others – the Turkey. Second, some subjects started with the choice of menu, the others with the choices of rewards. Third, the un-restricted menu was highly rewarded in an example provided to some subjects and it was poorly rewarded in an example provided to remaining subjects; in any case, subjects were told the specific values in the example were arbitrary and illustrative only. Finally, some subjects were asked if they wanted to commit to consumption at a future session before they ate their lunch and some were asked thereafter. Manipulations on the first three dimensions were of minor importance and were not expected to make a substantial behavioral difference. The last manipulation allowed investigating if willingness to commit and the choice depend on visceral factors such as being hungry vs satiated.

To comply with the promise given to the subjects and make their commitments meaningful, after a five-month break, we sent every participant of the original session (except for those who said there were not interested in that) an invitation to a new one. For various reasons, only twenty-seven showed up. They were identified using their e-mails or mobile phone numbers and again asked if they wanted a Turkey or a Burger. However, those who had pre-committed to one of the options during the first session, were at this point reminded of that. All subjects received the appropriate meal and filled in personality questionnaires: the brief self-control scale (Tangney, et al., 2004) as used by Błachnio and Przepiórka (2014) and the Polish version of the abbreviated Big Five inventory (Bąk et al, 2014). These additional measures were not significantly related to our variables of interest, possibly partly due to small sample size and self-selection, and will not be investigated in detail here.

6. Results

The choice of self-commitment

Our first dimension of interest is willingness to self-commit to a specific dish in the future. Overall, 20.6% of subjects preferred to immediately make a binding choice of the dish to be consumed during the new session. This was associated with abstaining in last parliamentary elections ($p = .022$) and voting the right-wing Law and Justice (PiS) ($p = .019$ when non-voters disregarded, $p = .002$ when included).

To concisely represent preference for freedom of choice vs. restriction of choice on individual level, the variable *bonus* was defined as difference between the reward the subject was willing to give for Menu 3 and the largest of the reward she was willing to give for the restricted menus. For example, if the rewards for menus 1-3 were 3, 4 and 2 respectively, then *bonus* was equal to minus 2. Given construction of the variable, it should take value of zero assuming each menu is as good as the best

item on the menu and reward for the menu is a function of the realized utility of the menu. If the reward is an increasing function of the expected utility of the menu and there is some uncertainty about others' preferences, a positive value of *bonus* is expected.

Not surprisingly in view of these results, there were also strong, negative, links of self-commitment with bonus ($p = .019$ in a Mann-Whitney test) and with letting the other party choose ($p = .004$ in a chi-square test), see Appendix A. The logistic regression in the appendix also shows that this link cannot be simply explained by the fact that individuals showing strong preference for one of the dishes were willing to both self-commit (because they knew they would not change their mind) and give another participants a restricted menu (because they thought they knew what would be good for him or her). When we control for strong preference (subjects responding "definitely burger" or "definitely turkey" to the question "Now let's assume that you receive the Menu: {Your choice: Burger or Turkey}. Which of these two are you going to choose?"), the variable *gave choice* is still a significant predictor of commitment.

The choice of rewards

Cumulative distribution of rewards that subjects gave for each of the three possible menu choices made for them is represented in Figure 3. Substantial heterogeneity can be observed, with intermediate values being most common. On average, subjects rewarded Menu 3 {Your choice: Burger or Turkey} most generously. Next, the distribution of bonus variable is shown in Figure 4. As it turns out, some subjects rewarded one of the restricted menus higher than Menu 3. The frequencies of all different orderings of the three menus are provided in Appendix A. It shows in particular that there was substantial heterogeneity and that only 15.6% of subjects made no distinctions between menus in their rewards (which typically meant always giving 5 or always giving 0).

Figure 3. On average, freedom of choice (Menu 3) was rewarded most

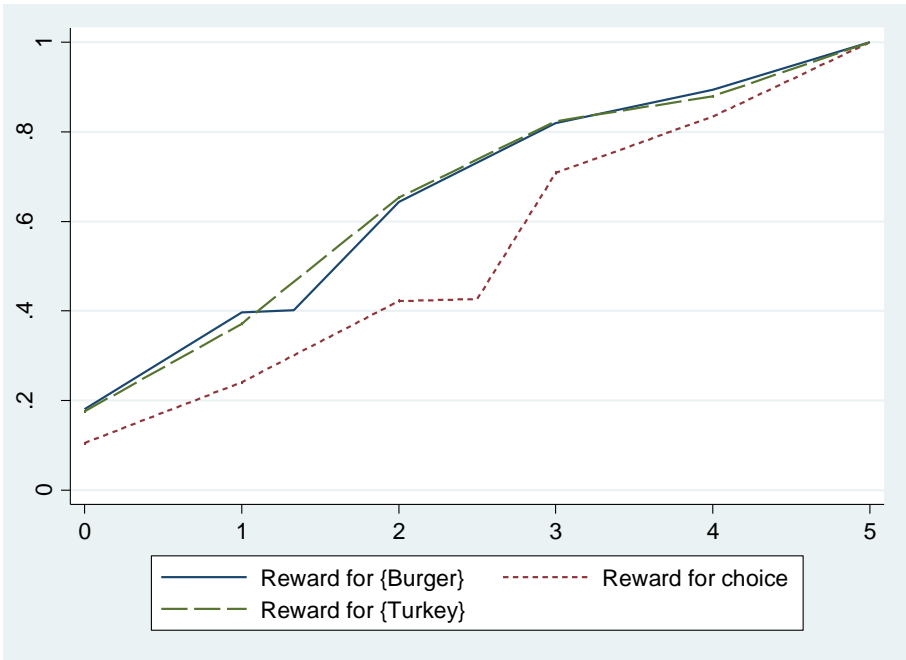
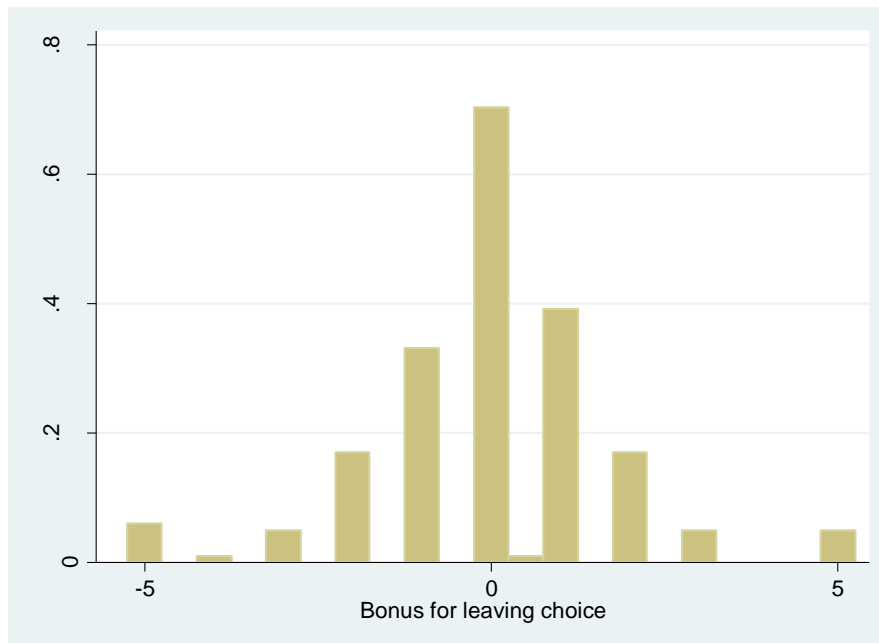


Figure 4. Compared to the most-awarded restricted menu, freedom of choice was rewarded by some, punished by others



We now seek to verify *who* was willing to reward choice (compared to the most-rewarded restricted menu) and who was willing to punish it. Again, there were few predictive demographic variables: not voting ($p = .032$) and voting for PiS was associated with lower *bonus* for the freedom of choice, while voting for the liberal Nowoczesna, with higher *bonus* ($p = .034$ and $p = .050$ resp. when non-voters are discarded and $p = .143$ and $p = .011$, respectively, when included, all in Mann-Whitney tests). Again, the signs are robust, but not all results are significant when a regression is run.

The fact that the pattern is somewhat similar to the choice of menu is not a coincidence, as the tendency to make a paternalistic choice and the tendency to relatively generously reward a paternalistic choice were strongly correlated: among the 65 subjects who made paternalistic choices, mean bonus was -1 , while it was $.47$ among 134 subjects who gave the choice of dish to their matched partners. The difference is strongly significant in a rank-sum test ($p < .001$). Again, all the tests and regressions can be found in Appendix A.

The choice of menu

Overall, two-thirds of subjects let their matched partner choose their own dish (coded as gave choice=1), 20.1% forced them to have the turkey and 12.6% to have a burger. No demographic variables significantly correlated with these choices, except that individuals who reported having voted for PiS party were more likely to choose one of the restricted menus ($p = .039$ in a chi-square test when non-voters are disregarded and $p = .061$ when they are included). Analogous results were obtained when the “paternalistic” option restricting menu specifically to Turkey was the focal variable, while there were no significant determinants of the less common case of restricting menu to Burger. Same results showed up, although less significant, when various sets of controls were included in a logistic regression, see Appendix A. These regressions resulted in some mixed effects of self-reported alcohol and cigarette consumption. The finding for PiS is consistent with the idea that supporters of this conservative party are more likely to behave in a paternalistic way.

As already mentioned, there is a strong link between self-commitment and paternalistic choices (see table 2).

Table 2. Frequency of population that self-commit and choose paternalistic menus.

commitment	gave choice		Total
	No	Yes	
No	22%	57%	79%
Yes	11%	10%	21%
Total	33%	67%	100%

Finally, those of 21% self-committing subjects end up with an average bonus of -0.6, while 79% not committing with mean 0.14 bonus. Those 11% that self-commit and gave paternalistic choice end up with average bonus amounting to -1.62.

Summary of the findings in view of considered theories

A few general observations must be made. First, the predictions of the standard theory, i.e. that no rewards would be given, are not confirmed. In fact, almost all subjects choose at least one non-zero reward. Second, vast majority of subjects do differentiate between menus, giving different rewards, apparently because they like some menu choices better than others. This allows testing our key predictions. We thus observe, third, that a non-trivial fraction of subjects restricts others' as well as own future choices. This is consistent with all main theories allowing for preference reversals associated with self-control issues. Crucially, however, we observe a significant fraction of subjects *rewarding* a restriction imposed by others. Moreover, this tendency is correlated with self-commitment and restricting others' options. These findings support theories of costly self-control but not theories of temporal inconsistency.

7. Discussion and conclusion

This is one of very few experimental studies that permit disentangling predictions of the main theories allowing for preference reversals due to self-control issues (and finding the costly self-control models superior). We use naturalistic (yet pretested) stimuli, which might increase external validity of our findings. We observe a non-trivial but plausibly low, comparable to previous studies, fractions of commitments and of paternalistic choices. Moreover, and as contrary to some literature, we document a non-trivial fraction of population that self-commits and also rewards such restrictions in cash.

To be sure, several limitations must be acknowledged. First, we cannot be completely sure that more highly rewarded menus are indeed those that subjects were happier with; we must rely on past (admittedly, voluminous) literature providing evidence of prevalence of reciprocal behaviour in this respect. Second, our setting made it impractical to extensively elicit additional psychological measures from our subjects. This was partially done in a follow-up study; however, no clear patterns could be found due to small sample size as mentioned before. Third, it may be difficult to distinguish between theories of choice reversals and some general preference for or against freedom of choice. In particular, when commenting upon their choices, quite a few subjects simply said that it is best when everybody can choose for him- or her-self. It cannot be excluded that some subjects had

exactly the opposite views. If someone preferred the choice to be restricted per se, not because they perceived any specific temptation, they could restrict their future choice and another's choice, and be fine with their choice being restricted.⁸ We believe that such pure anti-freedom preferences are very rare, particularly in the European culture. Finally and related to this, it must be noted that some subjects restricted their own or others' choices to Burger, presumably the tempting choice, which may be consistent with the letter but not the spirit of costly self-control models. On the other hand, from the pre-test we know that a (small) fraction of subjects pointed at turkey as the more tempting option.

On the whole however, we believe our study represents a useful step towards empirically uncovering the link between paternalistic choices and self-control problems. In particular, it shows that naturalistic stimuli can be conveniently combined with stringent laboratory control, possibly the optimal design choice for the study of temptation. If the main finding, favouring models of costly self-control, is confirmed in future research, important implications follow. Consider, for example, restrictions on advertisements of alcohol which are in force in many countries. Overall, there appears to be no evidence that they reduce total consumption (Nelson, 2010). Instead, one brand or one type of alcohol is substituted with another. It would seem therefore that advertising bans are hardly justifiable. However, to the extent that observing ads of alcoholic beverages increases the cost of exerting (ultimately successful) self-control, there could be substantial welfare effects even if total alcohol consumption is not changed.

References

- Adams, A., L. Cherchye, B. De Rock, and Verriest, E. (2014): Consume now or later? Time inconsistency, collective choice, and revealed preference, *American Economic Review*, 104, 4147–4183.
- Ashraf, N., Karlan, D., and Yin, W. (2006): Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines, *Quarterly Journal of Economics*, 121(2), 635-672.
- Augenblick, N., Niederle, M. and Sprenger, C. (2015): Working over time: dynamic inconsistency in real effort tasks, *Quarterly Journal of Economics*, 130, 1067–1115.
- Bąk, W., Mielniczuk, E., Oleś, P. K., Purc, E., and Łaguna, M. (2014): Short measure of personality TIPI-P in a Polish sample, *Roczniki Psychologiczne*, 17(2), 421-437.
- Bitterly, T. B., Mislavsky, R., Dai, H., and Milkman, K. L. (2014): Dueling with desire: a synthesis of past research on want/should conflict, In W. Hofmann and L. Nordgren (eds.) *Psychology of Desire*.
- Błachnio, A., and Przepiorka, A. (2015): Dysfunction of self-regulation and self-control in Facebook addiction. *Psychiatric Quarterly*, 1-8.
- Buccioli, A. (2012): Measuring self-control problems: a structural estimation, *Journal of the European Economic Association*, 10(5), 1084-1115.

⁸Note that this is not psychological choice-aversion: an individual who finds it difficult to make decisions could be happy to have one's own menu restricted but would not want to choose for somebody else.

- Caplin, A., and Leahy, J. (2006): The recursive approach to time inconsistency, *Journal of Economic Theory*, 131(1), 134-156.
- Casari, M., and Luini, L. (2009): Cooperation under alternative punishment institutions: an experiment, *Journal of Economic Behavior & Organization*, 71(2), 273-282.
- Cox, J.C., D. Friedman, and Gjerstad, S. (2007): A tractable model of reciprocity and fairness, *Games and Economic Behavior*, 59(1), 17-45.
- Dekel, E., and Lipman, B. L. (2012): Costly self-control and random self-indulgence, *Econometrica*, 80(3), 1271-1302.
- Dekel, E., B. L. Lipman, and Rustichini, A. (2001): Representing preferences with a unique subjective state space, *Econometrica*, 69(4), 891-934.
- Dekel, E., B. L. Lipman, and Rustichini, A. (2009): Temptation-driven preferences, *Review of Economic Studies*, 76(3), 937-971.
- Della Vigna, S., and Malmendier, U., (2004): Contract design and self-control: theory and evidence, *Quarterly Journal of Economics*, 119 (2), 353–402.
- Dziewulski, P. (2015): Revealed time-preference, Department of Economics Discussion Paper Series, University of Oxford, No. 748.
- Echenique, F., T. Imai, and Saito, K. (2016): Testable implications of quasi-hyperbolic and exponential discounting, Mimeo.
- Eliaz, K., and Spiegler, R., (2006): Contracting with diversely naive agents, *Review of Economic Studies*, 73 (3), 689–714.
- Esteban, S., and Miyagawa, E., (2006): Temptation, self-control, and competitive nonlinear pricing, *Economic Letters*, 90 (3), 348–355.
- Falk, A., and Fischbacher, U. (2006): A theory of reciprocity, *Games and Economic Behavior*, 54(2), 293-315.
- Fehr, E., and Gächter, S. (1998): Reciprocity and economics: the economic implications of homo reciprocans, *European economic review*, 42(3), 845-859.
- Gilpatric, S.M., (2008): Present-biased preferences, self-awareness and shirking, *Journal of Economic Behavior & Organization*, 67 (3–4), 735–754.
- Gine, X., Goldberg, J. Silverman, D. and Yang, D. (2016): Revising commitments: field evidence on the adjustment of prior choices, *Economic Journal*, DOI: 10.1111/ecoj.12378.
- Gul, F., and Pesendorfer, W. (2001): Temptation and self-control, *Econometrica*, 69(6), 1403-1435.
- Gul, F., and Pesendorfer, W. (2004): Self-control and the theory of consumption, *Econometrica*, 72(1), 119-158.
- Heidhues, P., and Koszegi, B., (2009): Futile Attempts at Self-Control, *Journal of the European Economic Association*, 7(2-3), 423-434.
- Heidhues, P., and Koszegi, B., (2010): Exploiting naivete about self-control in the credit market, *American Economic Review*, 100 (5), 2279–2303.
- Janssens, W., Kramer, B. and Swart, L. (2015): Be patient when measuring hyperbolic discounting: stationarity, time consistency and time invariance in a field experiment, unpublished manuscript.
- Kataria, M., Levati, M. V., and Uhl, M. (2014): Paternalism with hindsight: do protégés react consequentialistically to paternalism? *Social Choice and Welfare*, 43(3), 731-746.

- Kopylov, I. (2012): Perfectionism and choice, *Econometrica*, 80(5), 1819-1843.
- Le Lec, F., and Tarrow, B. (2015): Is more choice always preferred? Experimental evidence. Working Paper.
- Kreps, D. M. (1979): A representation theorem for "preference for flexibility", *Econometrica*, 47(3), 565-577.
- Montiel Olea, J. L. and Strzalecki, T. (2014): Axiomatization and measurement of quasi-hyperbolic discounting, *Quarterly Journal of Economics*, 129, 1449–1499.
- Nelson, J. P. (2010): Alcohol advertising bans, consumption and control policies in seventeen OECD countries, 1975–2000, *Applied Economics*, 42(7), 803-823.
- Noor, J. (2007): Commitment and self-control, *Journal of Economic Theory*, 135(1), 1-34.
- O'Donoghue, T., and Rabin, M. (1999): Doing it now or later, *American Economic Review*, 89(1), 103-124.
- Olszewski, W. (2011): A model of consumption-dependent temptation, *Theory and Decisions*, 70(1), 83-93.
- Peleg, B., and Yaari, M. E. (1973): On the existence of a consistent course of action when tastes are changing, *Review of Economic Studies*, 40(3), 391-401.
- Pedersen, S. K., Koch, A. K., and Nafziger, J. (2014): Who wants paternalism? *Bulletin of Economic Research*, 66(S1), S147-S166.
- Pew Research Center (2015): In debate over legalizing marijuana, disagreement over drug's dangers. <http://www.people-press.org/2015/04/14/in-debate-over-legalizing-marijuana-disagreement-over-drugs-dangers/>
- Phelps, E., and Pollak, R. (1968): On second best national savings and game equilibrium growth, *Review of Economic Studies*, 35, 195-199.
- Read, D., and Van Leeuwen, B. (1998): Predicting hunger: the effects of appetite and delay on choice, *Organizational Behavior and Human Decision Processes*, 76(2), 189-205.
- Sarver, T. (2008): Anticipating regret: why fewer options may be better, *Econometrica*, 76(2), 263-305.
- Strotz, R. H. (1955-1956): Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies*, 23(3), 165-180.
- Tangney, J. P., R. F. Baumeister, et al. (2004): High self-control predicts good adjustment, less pathology, better grades, and interpersonal success, *Journal of Personality*, 72(2), 271-324.
- Toussaert, S. (2014): Connecting commitment to self-control: evidence from a field experiment with participants in a weight loss challenge. Working Paper.
- Toussaert, S. (2015): Eliciting temptation and self-control through menu choices: a lab experiment. Working Paper.
- Woźny, Ł. (2015): On incentives, temptation and self-control, *Mathematical Social Sciences*, 74, 60–67.
- Yilmaz, M., (2013): Repeated moral hazard with a time-inconsistent agent, *Journal of Economic Behavior & Organization*, 95 (C), 70–89.

Appendix A: Additional tables

Non self-explicable variables:

pref for burg=1 if subject predicts to definitely choose turkey when given a choice, 2=rather turkey, 3=hard to say, 4=rather burger 5=definitely burger

strong pref = 1 if pref for burger =1 or pref for burger=5

PiS and nowoczesna are the two political parties most popular in our sample, being conservative and liberal resp.

BMI, the body mass index, is calculated as weight in kg/height in meters

alco>1/month is a dummy variable indicating that subjects self-reports consuming alcohol more often than about once per month

econ indicates a major in economic sciences

business_sch is a dummy indicating a session run at the Warsaw School of Economics rather than the University of Warsaw

The choice of menu

```
. tab gave_choice PIS, chi
```

Gave choice	PIS		Total
	0	1	
0	50	15	65
1	117	17	134
Total	167	32	199

Pearson chi2(1) = 3.5016 Pr = 0.061

```
. tab gave_choice PIS if non_v==0, chi // non-voters disregarded
```

gave choice	PIS		Total
	0	1	
0	30	15	45
1	79	17	96
Total	109	32	141

Pearson chi2(1) = 4.2635 Pr = 0.039

```
. tab gave_turkey PIS, chi
```

gave turkey	PIS		Total
	0	1	
0	139	20	159
1	28	12	40
Total	167	32	199

Pearson chi2(1) = 7.1880 Pr = 0.007

. tab gave_turkey PIS if non_v==0, chi // non-voters disregarded

gave turkey\	PIS		Total
	0	1	
0	92	20	112
1	17	12	29
Total	109	32	141

Pearson chi2(1) = 7.2646 Pr = 0.007

Logistic regression models

explained var\	gave choice		gave turkey	
PIS	-0.7270*	-0.6993	1.2845**	1.1953**
nowoczesna	0.3431	0.4505	0.2821	0.3343
non_voter	-0.2136	-0.0615	0.3455	0.1659
male	0.1037	0.0894	-0.0761	0.3941
BMI		0.0849		-0.0985
smoker		-0.9755**		0.7523
alco>1/month		0.6357*		-0.7567*
pref for burg		-0.0451		-0.2553*
strong pref		0.2171		0.2668
_cons	0.8004***	-1.4315	-1.7575***	1.2266
N	198	197	198	197

legend: * p<.1; ** p<.05; *** p<.01

The choice of rewards

Distribution of orderings of the menus

reward_orde ring	Freq.	of which: constrained	
		Percent	Freq. Percent
b>c>t	29	14.57	1 0.50
c=b>t	18	9.05	1 0.50
c>b>t	20	10.05	0 0
c>t=b	26	13.07	7 3.52
c>t>b	21	10.55	0 0
t=b=c	31	15.58	25 12.56
t=b>c	4	2.01	0 0
t=c>b	21	10.55	0 0
t>b=c	9	4.52	4 2.01
t>b>c	6	3.02	0 0
t>c>b	14	7.04	0 0
Total	199	100.00	38 19.01

t={turkey}, b={burger}, c=choice. In the rightmost columns the number of observations is provided, for which the ordering could have been affected by the [0,5] constraints on rewards, because the median reward was equal to 0 or 5. For example, if rewards for {t,b,c} are {5,5,4}, it is classified as t=b>c but it cannot be ruled out that the subject was in fact willing to reward t or b even higher, so that her real preference was t>b>c or b>t>c. This corresponds to corner solution in our theoretical model in section 3. Frequencies of such cases are provided in the rightmost columns

to give a more complete view of distribution of preference. By constructions, they are zero for orderings with no indifferences.

```
. ranksum bonus, by(non_voter)
Two-sample Wilcoxon rank-sum (Mann-Whitney) test
non_voter |      obs      rank sum      expected
-----+-----
          0 |      141      14863.5      14100
          1 |       58       5036.5       5800
-----+-----
combined |      199      19900      19900
unadjusted variance    136300.00
adjustment for ties      -7756.88
-----
adjusted variance      128543.12
Ho: bonus_~e(non_vo~r==0) = bonus_~e(non_vo~r==1)
      z =      2.130
      Prob > |z| =      0.0332
```

```
. ranksum bonus if non_v==0, by(PIS)
Two-sample Wilcoxon rank-sum (Mann-Whitney) test
PIS |      obs      rank sum      expected
-----+-----
          0 |      109      8155.5       7739
          1 |       32      1855.5       2272
-----+-----
combined |      141      10011      10011
unadjusted variance    41274.67
adjustment for ties      -2497.88
-----
adjusted variance      38776.79
Ho: bonus_~e(PIS==0) = bonus_~e(PIS==1)
      z =      2.115
      Prob > |z| =      0.0344
```

```
. ranksum bonus, by(PIS)
Two-sample Wilcoxon rank-sum (Mann-Whitney) test
PIS |      obs      rank sum      expected
-----+-----
          0 |      167      17124.5      16700
          1 |       32      2775.5       3200
-----+-----
combined |      199      19900      19900
unadjusted variance    89066.67
adjustment for ties      -5068.81
-----
adjusted variance      83997.86
Ho: bonus_~e(PIS==0) = bonus_~e(PIS==1)
      z =      1.465
      Prob > |z| =      0.1430
```

```

. ranksum bonus if non_v==0, by(nowoczesna)
Two-sample Wilcoxon rank-sum (Mann-Whitney) test
nowoczesna |      obs      rank sum      expected
-----+-----
          0 |      102       6830       7242
          1 |       39       3181       2769
-----+-----
combined |      141      10011      10011
unadjusted variance    47073.00
adjustment for ties    -2848.78
-----
adjusted variance      44224.22
Ho: bonus_~e(nowocz~a==0) = bonus_~e(nowocz~a==1)
      z =  -1.959
      Prob > |z| =  0.0501

```

```

. ranksum bonus, by(nowoczesna)
Two-sample Wilcoxon rank-sum (Mann-Whitney) test
nowoczesna |      obs      rank sum      expected
-----+-----
          0 |      160      15208      16000
          1 |       39       4692       3900
-----+-----
combined |      199      19900      19900
unadjusted variance   104000.00
adjustment for ties    -5918.67
-----
adjusted variance      98081.33
Ho: bonus_~e(nowocz~a==0) = bonus_~e(nowocz~a==1)
      z =  -2.529
      Prob > |z| =  0.0114

```

Regression models

```

-----
          bonus |          (1)          (2)
-----+-----
          PIS | -0.6635*          -0.4281
nowoczesna |  0.3655           0.2752
non_voter | -0.5928*          -0.5599*
      male |  0.0436           0.0173
gave choice |                   1.4635***
          BMI |                   -0.0221
      smoker |                   0.3001
alco>1/month |                   -0.3504
pref for burg |                   0.0381
strong pref |                   -0.1586
          _cons |  0.1729           -0.1779
-----+-----
          N |          199          197
          r2 |  0.0530           0.2029
-----

```

legend: * p<.1; ** p<.05; *** p<.01
POST means that commitment choice was made after having consumed the lunch.

The choice of self-commitment

```
. tab commitment PIS, chi
```

commitment	PIS		Total
	0	1	
0	139	19	158
1	28	13	41
Total	167	32	199

Pearson chi2(1) = 9.3447 Pr = 0.002

```
. tab commitment PIS if non_v==0, chi // non-voters disregarded
```

commitment	PIS		Total
	0	1	
0	87	19	106
1	22	13	35
Total	109	32	141

Pearson chi2(1) = 5.5392 Pr = 0.019

```
. tab commitment non_v, chi
```

commitment	non_voter		Total
	0	1	
0	106	52	158
1	35	6	41
Total	141	58	199

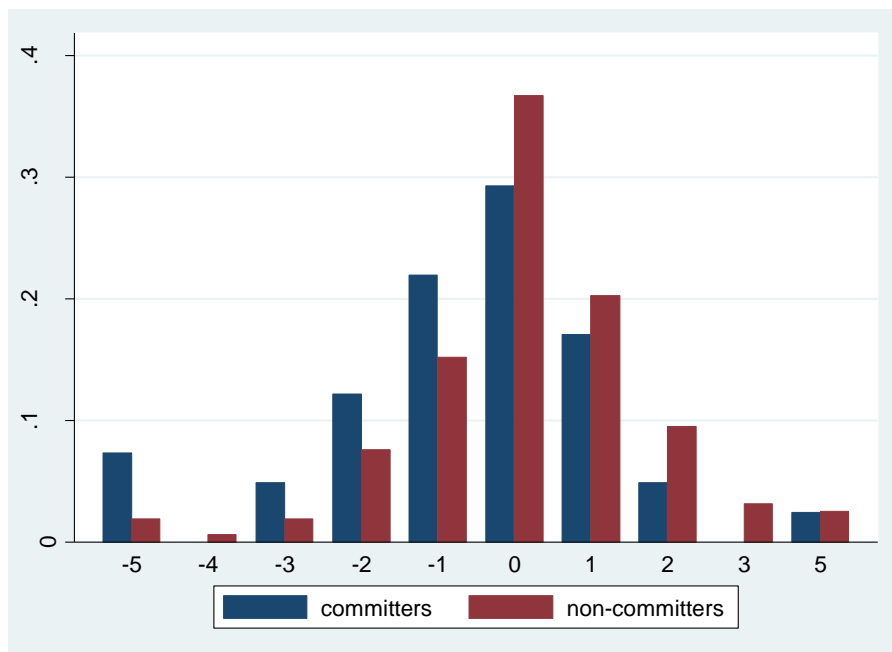
Pearson chi2(1) = 5.2658 Pr = 0.022

Logistic regression models

commitment	(1)	(2)
POST	0.1703	0.2170
left_choice	-0.8928**	-1.2307***
male	-0.4550	-0.6027
smoker	0.4013	0.0753
PIS	0.8408*	0.8832
nowoczesna	-0.1062	-0.0123
non_voter	-0.8967*	-0.8313
pref for burg		-0.0099
BMI		-0.0147
alco>1/month		0.2707
strong pref		1.7888***
econ		0.4149
business_sch		-0.5363
_cons	-0.6739	-1.4151
N	198	197
r2		

legend: * p<.1; ** p<.05; *** p<.01

Histograms: *bonus* by commitment



Appendix B: Pre-test

This survey is conducted for scientific purposes only. A prize of PLN 50 will be awarded to one of the responders, to be selected at random. If you want to participate in the drawing, you must leave your e-mail address at the end, so that we will be able to contact you. All the answers will be confidential.

There are no correct or incorrect answers in this survey; we are interested in your own opinion. Do not think very long about the answers, just pick the one that comes to your mind.

Please carefully read the descriptions and of dishes provided below and have a look at the corresponding pictures.

DESCRIPTION OF THE DISHES

Beef burger with barbecue sauce, roasted potatoes with herbs, salad



Description: this dish is based on products which provide necessary energy. Here is the nutritious grilled beef Burger spiced with herbs served with traditional American barbecue sauce. A satiating and filling meal, it is a real temptation for any connoisseur.

Nutritional values:

kcal	Proteins	Fat	Saturated fat	Monounsaturated fat	Polyunsaturated FAT	Cholesterol	Total carbohydrate	Detary fiber
795	49,0	31,6	7,85	15,70	5,87	126	85,1	7,6

Turkey with pineapple curry, bulgur and steamed vegetables



Description: this dish was created by dieticians for a good balance of proteins and carbohydrates. Fat was only used in small amounts, to prepare that meal. The dish was prepared with the use of grilling, steaming and roasting. A great option for people who want to keep fit and stay in shape.

Nutritional values:

kcal	Proteins	Fat	Saturated fat	Monounsaturated fat	Polyunsaturated FAT	Cholesterol	Total carbohydrate	Detary fiber
398	48,4	9,0	5,33	2,95	1,70	76	63,2	8,2

Which one would you rather have if both were offered for the same, reasonable price?

definitely dish no. 1/rather dish no. 1/hard to say/rather dish no. 2/definitely dish no. 2 //none of them

If you were to decide on your diet for the next year, would you rather like to see more dishes similar to dish no. 1 (burger) or similar to dish no. 2 (turkey) there?

[definitely similar to dish no. 1 ...]

If you had a teenage child, which dish would you rather order for him or her?

[definitely dish no. 1 ... none of them]

You had an opportunity to become familiar with two dishes: beef burger with roasted potatoes and turkey with bulgur and steamed vegetables. Now please answer these questions about them. Which of these two dishes is... [more healthy, more eco-friendly, more tasty, more attractive, more satiating, more tempting] [definitely burger etc.]

Please mark all the statements that apply to your dieting experiences with an "X"

- I feel fine with my body weight. I have never or almost never tried to reduce it.
- I have been dieting in the past and it was usually successful
- I have tried dieting in the past, with mixed success

I have often tried to reduce my body weight, but to no avail.

I try to lead a healthy lifestyle. I care a lot about what I eat.

I do not have any time for planning my meals. I just eat what I feel like.

[demographic data, additional remarks]

Appendix C: Instructions

Enter the number that you have received from the organizers: _____

Card no. 1

We kindly ask you to keep silence.

Every participant will receive a free lunch. There are two kinds of dishes: a Burger and a Turkey. Each participant will receive only one of these two. Which dish you receive may depend on decisions that you and possibly another participant make. Similarly, your decisions may affect which dish yet another participant receives. Please read the description of the dishes (enclosed on a separate sheet) now and look at the photos displayed on the screen. You can also have a look directly at the exhibited dishes.

***** continue reading after you have finished reading the attached descriptions*****

You will be matched to a randomly picked participant. You will not get to know each other's identities. You will have to decide now, whether the other participant receives a Burger, a Turkey or whether (s)he will choose by her/himself. In other words, you can choose one of the three *Menus* for him or her:

{Burger}

{Turkey}

His/her choice: Burger or Turkey

This participant will not have an influence on what you get. However, (s)he can reward you with money for the Menu that you choose for him/her. It is thus in your best interest to pick a good menu.

Which menu do you want to give to the other participant? (mark one of the options with an "X").

- {Burger}
- {Turkey}
- His/her choice: Burger or Turkey

PLEASE RAISE YOUR HAND AND GIVE THIS CARD BACK TO THE ORGANIZERS. Keep the description of the dishes. Fill in Card no. 2

[please see Appendix B for the description of the dishes]

Go back to Card no.1 (the location marked with stars) and read on.

Enter the number that you have received from the organizers: _____ Card no.2

Answer all the questions on this card and then return it to the organizers.

Just as you have chosen the Menu for another participant, yet another participant (thus not the same for whom you were choosing the Menu) is going to choose a Menu for you. You will not get to know each other's identities. You can reward him/her for his or her choice. You receive 5 PLN just for your participation in this part. For each of the possible Menus, you decide how much out of this money (perhaps the entire amount or nothing at all) you would like to give to the other participant as a reward. Depending on which Menu you actually get, you will pay the relevant amount of money and the rewarded participant will receive THREE TIMES that amount.

Example: you want to reward a choice of a Menu for you in the following way: for choosing {Burger}: 2 PLN, for choosing {Turkey}: 3 PLN, for (Your choice: Burger or Turkey): 1 PLN. Let's assume that you actually receive the {Burger} Menu. You eat the Burger and pay 2 PLN from your endowment of 5 PLN and the participant who chose that Menu for you receives $2 \cdot 3 = 6$ PLN and likewise for the {Turkey} Menu. If you receive (Your choice: Burger or Turkey) Menu, you choose what to eat, pay 1 PLN and the other participant receives 3 PLN. Of course, this is just an example. The specific numbers were chosen arbitrarily and should not influence your own decisions.

The money that you keep out of your endowment of 5 PLN (in the example above, if you receive a {Turkey} Menu, $5 - 3 = 2$ PLN) will be added to your earnings from the previous part.

When choosing the menu for you, the other participant will not know what rewards you have specified for each specific Menu.

How much would you like to pay as a reward for each of possible Menus? (Write three amounts from 0 to 5 each).

Reward for choosing a {Burger} for you: _____ PLN.

Reward for choosing a {Turkey dish} for you: _____ PLN.

Reward for choosing (Your choice: Burger or Turkey) for you: _____ PLN.

Now let's assume that you receive the Menu: (Your choice: Burger or Turkey). Which of these two are you going to choose? (mark one of the options with an "X").

- Definitely Burger
- Rather Burger
- Hard to say
- Rather Turkey
- Definitely Turkey

GIVE THIS CARD BACK TO THE organizers and start filling in Card no. 3

Enter the number that you have received from the organizers: _____

Card no.3

Briefly answer the following questions.

How did you decide which Menu to offer to another participant?

How did you decide how to reward the choice of Menu made for you by another participant?

***** Please wait until you know which menu you have received.*****

You might have to wait for a while.

Which Menu you have received (filled by the organizers)

{Burger} {Turkey} (Your choice: Burger or Turkey)

CONTINUE WHEN YOU KNOW WHICH MENU YOU HAVE RECEIVED

IF YOU HAVE RECEIVED (Your choice: Burger or Turkey), which of these two do you choose for yourself?

(Ignore this question if you have received another menu)

Burger Turkey

GIVE THIS CARD BACK TO THE ORGANIZERS and pick up the appropriate dish. Fill in Card no. 4

Enter the number that you have received from the organizers: _____

Card no. 4

Please answer the following questions. You can do it before, during, or after your lunch. As has been mentioned before, all of your answers will remain confidential.

Do you smoke cigarettes? Please mark one of the options with an "X"

Yes No

If so, how many do you smoke on a typical day? _____

If not, have you ever smoked cigarettes?

Yes No

How often do you drink beverages containing alcohol?

- never
- about once a month or less often
- two to four times a month
- two to three times a week
- four or more times a week

Please mark all the statements that apply to your dieting experiences with an "X" [as in the pre-test]

Please flip the page

Your gender

Card no. 4 – continued

Female Male

How old are you? _____

How tall are you? _____

What is your body weight? _____

What do you study?

Economics or business sth else: _____ I don't study

How would you locate your political views on a scale from 1=definitely left-wing to 10=definitely right wing?

left 1 2 3 4 5 6 7 8 9 10 right

Which party did you support in last year's parliamentary elections (if you voted at all)

I did not vote (or I casted an invalid ballot) KORWiN Kukiz'15 .Nowoczesna PiS PO PSL Razem Zjedn. Lewica else: _

How did you learn about this experiment

I'm on the ORSEE e-mail list I have seen a poster some other way: _____

Have you participated in laboratory economic experiments in the past?

Yes No

Please answer these questions after having eaten at least some of the lunch:

How satisfied are you with the lunch?

not at all 1 2 3 4 5 6 7 8 9 10
completely

We are planning to run another, similar session before the end of the semester, perhaps with somewhat higher cash earnings. Are you potentially interested in participation

Yes, please e-mail or text me. I do not promise to register. No

During this new session, we will have the same dishes. You can choose between Burger and Turkey right now, after having eaten today's lunch (and your choice will be binding for the new session) or you can choose during the new session.

I want to choose now I want to choose during the new session

IF YOU WANT TO CHOOSE NOW, which dish would you like to choose? (please ignore this question if you indicated you wanted to choose at the new session.

Burger Turkey

If you were to choose during the new session, which one do you expect to choose?

Definitely Burger

Rather Burger

Hard to say

Rather Turkey

Definitely Turkey

Please write down your e-mail (or telephone number if you prefer to be texted about the next session)

Many thanks for participation. If you have any further thoughts or comments, please write them down or approach the organizers.