JENA ECONOMIC RESEARCH PAPERS

2012 – 050

Paternalism With Hindsight Do protégés react consequentialistically to paternalism?

by

**Mitesh Kataria
M. Vittoria Levati
Matthias Uhl**

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

© by the author.

Paternalism With Hindsight

Do protégés react consequentialistically to paternalism?

Mitesh Kataria^a, M. Vittoria Levati^{a,b,*}, and Matthias Uhl^c

^a*Max Planck Institute of Economics, Kahlaische Str. 10, 07745 Jena, Germany*

^b*University of Verona, DSE, Via dell'Artigliere 19, 37129 Verona, Italy*

^c*Technical University of Munich, Lothstrasse 17, 80335 Munich*

Abstract

We investigate experimentally whether the protégés' reaction to paternalism depends on the consequences of the paternalistic action to their well-being. We find that protégés punish a paternalist restricting their freedom of choice. Yet, this negative reaction is not based on principled grounds because, with hindsight, protégés punish the paternalist only if the restriction makes them worse off. Conversely, if the restriction makes them better off, the protégés on average do not punish and, sometimes, they even reward the paternalist. This suggests that protégés take a consequentialist stand on paternalism. Controlling for intentions ascribed to the patron does not alter our finding.

JEL classification: C92, D6, P16

Keywords: Paternalism, Consequentialism, Value of freedom, Experiment

We thank Dominique Cappelletti, Alon Harel, Elina Lampi, Matteo Ploner, and Birendra Rai for helpful suggestions. The authors bear the responsibility for any errors and omissions.

*Corresponding author. Max Planck Institute of Economics, Strategic Interaction Group, Kahlaische Str. 10, 07745 Jena, Germany. Tel.: +49 3641 686629; fax: +49 3641 686667. Email: levati@econ.mpg.de; vittoria.levati@univr.it

1 Introduction

Thomas Schelling cites the anecdote of a politically radical son in his twenties who is furious with his conservative father for “putting a sum of money in trust that the son may use for political contributions only when he reaches the conservative age of forty” (Schelling 1984, p. 6). Can these negative feelings toward the paternalistic act change and become positive when the son reaches forty? A condition under which an affirmative answer may be given to this question is that the son actually converts to conservatism when he turns forty. He might then appreciate the mere consequences of his father’s paternalism, i.e., the prevention of supporting political purposes that, with the wisdom of hindsight, he disapproves.

Governments in most Western societies are increasingly embracing paternalistic policies especially in areas where individual and societal costs of certain types of behavior (such as alcohol abuse, smoking, gambling, consumption of unhealthy food) are widely discussed.¹ In the United States, paternalistic regulations are not limited to one side of the political spectrum. The Democratic US president Barack Obama has selected Cass Sunstein (one of the thought leaders of “libertarian paternalism”) to be his regulatory czar and to run the Office of Information and Regulatory Affairs. The current Republican mayor of New York City, Michael Bloomberg, has proposed to ban the sale of sodas larger than 16 ounces in bars and shops of his government district as part of a larger campaign against unhealthy food. In Australia, the Productivity Commission has recommended the use of a poker machine commitment system under which “people with gambling problems could be required to set a limit on the amount of money they wish to spend on playing poker machines in a given period (...) or set a time limit on their use of poker machines” (Thomas and Buckmaster 2010, p. 23).²

Given these widespread paternalistic state interventions, the question can be raised as to whether or not citizens judge a paternalistic government, which introduces coercive measures, on the basis of the consequences of these measures. The answer to such a question may have important implications

¹Following Dworkin (2010), we take paternalism to mean “the interference of a state or an individual with another person, against their will, and defended or motivated by a claim that the person interfered with will be better off or protected from harm.”

²Other examples of paternalistic regulations include banning narcotics, warnings on cigarettes, public health advertising, and the obligations to wear a motorcycle helmet or to fasten seat-belts.

concerning the timing of the paternalistic interventions by a government that wants to be re-elected. Specifically, if the consequences are what matter, then the paternalistic interventions should be pursued well before the election year so that their effect can be visible. Otherwise, the interventions might be postponed until around election time.

Ascertaining whether people's attitudes and behavior toward paternalistic acts are influenced by their consequences is therefore of great importance for policy. Yet, we are not aware of any empirical effort to examine this issue.³ In this paper we study experimentally protégés' reaction to paternalism when the consequences of the paternalistic actions have proven to be unambiguously beneficial or harmful.⁴ A subject can act either as a protégé or as a patron. The protégés may gamble their experimental money in a lottery with zero expected return, while the patron can restrict their gambling choices.⁵ Then, in the hindsight treatment, the protégés are informed about the realization of the lottery and are given the opportunity to punish or reward the patron. Although imposing restrictions on liberty may be perceived by the protégé as undesirable for it challenges the principle of an individual's sovereignty over his own choices (see, e.g., Sen 1987), once the consequences of the restrictions are evident, the protégé's judgment may change. Consequentialist reasoning (e.g., Honderich 1995) suggests that if the protégé realizes that the paternalistic act has made him better off, he should value it.

By looking at the reaction of the protégés when the patron's restriction makes them better off and when the restriction makes them worse off, we can provide support for the view that behavior and feelings toward paternalism are in line with consequentialism. By comparing the results of the hindsight treatment with those of a foresight treatment, where the consequences of the restriction are unknown, we can show that knowledge of the consequences influences behavior.

³Johansson-Stenman (2012) is a notable exception. Based on survey evidence, he studies whether people see an action as ethically bad because of its overall consequences or on some other grounds. His results suggest that most people are consequentialists in their ethical perceptions. While Johansson-Stenman focuses on people's normative views, we look at people's actual behavior.

⁴The laboratory provides an ideal environment for studying unambiguous outcomes, whereas real-life outcomes can be more or less ambiguous depending on the context.

⁵The choice of a lottery seems to be the appropriate starting point for an experimental analysis because paternalistic policies often deal with choices whose outcomes are risky/uncertain.

Recently, philosophy is opening up to empirical work that takes “folk intuitions” into account (see, e.g., Appiah 2008). While traditional armchair philosophers tell us how moral agents *ought* to be, contemporary empirical philosophers maintain that any normative theory should keep in mind that “ought implies can” (e.g., Knobe 2003; Greene et al. 2009). In this vein, policy makers are well-advised to take people’s systematic moral intuitions into account when designing their policies.

Orthodox economists are notoriously opposed to paternalism.⁶ Developments in behavioral economics have, however, brought to light a large variety of decision biases and decision-making errors that may serve to inform paternalistic policies.⁷ As O’Donoghue and Rabin (2003, p. 186) write: “humans are humanly rational rather than superhumanly rational”, implying that people do not always act in their own best interest. Notably, the violation of the rationality assumption is a key point in any argument for paternalistic policies. Thaler and Sunstein (2003) have promoted the concept of “libertarian paternalism” (also termed “asymmetric paternalism” by Camerer et al. 2003) according to which, policy makers should exploit well-established psychological biases in order to “nudge” (Thaler and Sunstein 2008) individual behavior toward self-interest without limiting choice. Gruber and Köszegi (2008), on the other hand, use the behavioral approach (specifically, a model of hyperbolic discounting) to justify a sin tax that makes smokers internalize the long-term cost of their bias for immediate gratification.

While the discussion in the literature generally focuses on the *justifiability* of paternalism from an ethical or welfare perspective, our research will be concerned with the *perception* of paternalism by those who are directly affected by it. This will help us understand how common people, not educated philosophers or economists, feel about paternalism.

The remainder of the paper is organized as follows. In the next section, we lay out our experimental design and procedures. After detailing our research hypotheses (Section 3), we present the experimental findings (Section 4). We summarize the main points of our study and offer concluding

⁶Some liberals, following Immanuel Kant, believe that liberty has an intrinsic value. Other liberals, following John Stuart Mill, believe that liberty has merely an instrumental value for it helps to maximize utility or aggregate happiness.

⁷See, e.g., Camerer et al. (2003) for a description of cases in which paternalism may be useful.

remarks in Section 5.

2 Experimental design

2.1 Phases and treatments

The experiment consisted of two successive phases. The first phase was designed to elicit the participants' choices about a risky gamble and the paternalistic restrictions on such choices. Hence, phase 1 is preparatory for our main research objective (namely to study whether people's reactions to paternalistic behavior are contingent upon its consequences or independent of these), which is tackled in the unannounced second phase. The first sentence of the instructions explained the dangers of gambling and how heavily gambling is regulated in many countries.

At the beginning of phase 1, one participant in each session was randomly assigned the role of "patron" and the remaining participants were assigned the role of "protégés".⁸ The patron received an endowment of 50 ECU (Experimental Currency Unit) and each protégé received an endowment of 100 ECU (one ECU was worth 0.10 Euros). The protégés had to decide which amount (x) of their endowment they wanted to bet in a lottery, where x could be any integer value between 0 and 100. The lottery returned two times the bet with one-half probability and nothing with one-half probability, so that its expected return is $-x + \frac{2x+0}{2} = 0$.

While the protégés made their decisions about x , the patron was given the opportunity to impose a limit to their bet – thus restricting their gambling choices – with no monetary consequences to himself.⁹ Subjects were explained that the patron would receive a fixed payment of 50 ECU irrespective of his choice. We discretized the choice set of the patron to six restriction possibilities: $R \in \{0, 20, 40, 60, 80, 100\}$, where $R = 0$ means restricting access to any positive bet (i.e., allowing no gambling), $R = 20$ means restricting access to bets higher than 20, and so forth, until $R = 100$ that implies no restriction at all. The restriction chosen by the patron ap-

⁸Apart from the first sentence, we formulated the instructions as neutrally as possible, avoiding suggestive terms like patron and protégé. A translated version of the instructions can be found in the appendix.

⁹An "impartial" patron, with no stake in the game, should be perceived as someone who is exclusively interested in what is objectively desirable. This should reduce suspicions by the protégés that the patron made his choice out of self-interest.

plied to all protégés in the session, and it was binding only when $x > R$. Protégés did not receive any feedback about the restriction R set by the patron until the end of the session. Depending on the experimental treatment (explained later in this section), the outcome of the lottery was either disclosed or not disclosed to the protégés.

In (unannounced) phase 2, the protégés were allowed to punish or reward the patron for the imposed restriction. In this phase, the patron received an additional endowment of 50 ECU and had no decision to make. The protégés could alter the patron's total endowment (of 100 ECU) by any integer amount $\Delta \in [-50, +50]$, without any cost to themselves. By choosing a negative (positive) Δ , the protégé was punishing (rewarding) the patron. Since protégés were not informed about the restriction R chosen by the patron, we used the strategy method: each protégé had to decide on the amount Δ by which to vary the patron's endowment for every possible level of restriction (hence, the protégés had to make six choices). At the end of the session, one protégé was randomly selected and his conditional choice was applied to the patron. The patron's final payoff was therefore $100 + \Delta_i$, with i being the randomly selected protégé. Denoting the amount actually bet by the protégé by γ , where γ equals x if $x \leq R$ and R otherwise, his payoff was $100 + \gamma$ or $100 - \gamma$ depending on the realization of the lottery.

We implemented two treatments, which differed with respect to the information supplied to the protégés at the end of phase 1. In the *hindsight* treatment, the protégés made their Δ -choices *after* learning the lottery outcome so that they knew with certainty the consequences (beneficial or harmful) of the restriction. A comparison of Δ -choices between the group of subjects that observed a positive outcome and the group of subjects that observed a negative outcome allows us to assess whether the protégés' reaction to the patron's restriction depends on the realization of the lottery. We will distinguish between these two groups and refer to them as *positive outcome-hindsight* and *negative outcome-hindsight* treatment groups.

In the *foresight treatment*, the protégés made their Δ -choices *before* knowing the lottery outcome so that they could not base their decisions on the actual consequences of the restriction. Comparing Δ -choices in the foresight treatment with Δ -choices in the two hindsight treatment groups, we can determine the extent to which the protégés' reaction to the restriction changes according to whether consequences are known or unknown. A

difference in Δ -choices between the foresight treatment and the negative outcome-hindsight treatment group would justify the tenet “you will not understand it now” (because you cannot see the consequences of the action), “but you will thank me later” (when the consequences become clear). From the results of the foresight treatment it will also be possible to gain an insight into the way in which protégés value freedom of choice, though this is not the primary purpose of the study.

It has been shown that eliciting decisions via the strategy method may affect results compared to the direct-response method (see, e.g., Brandts and Charness, 2011), yet we deliberately decided to use it. The reason is that the strategy method allows us to obtain observations at the harshest restriction (namely, $R = 0$) that may otherwise not have occurred. The protégés’ reaction to $R = 0$ provides a clear-cut assessment of their genuine attitudes toward paternalistic restrictions because punishing and rewarding behavior in response to $R = 0$ cannot be explained by inequality aversion. To see this, note that when $R = 0$ the protégé earns 100 ECU for sure. If he then chooses not to alter the patron’s endowment (by setting $\Delta = 0$), the patron’s earnings are 100 ECU as well. Conversely, if the protégé sets $\Delta \neq 0$, this reaction comes at the cost of introducing payoff inequalities. Thus, evidence of punishment or reward in response to $R = 0$, creation of inequalities notwithstanding, provides an unconfounded test of our hypotheses.

The reaction to $R > 0$ could be explained by aversion to inequality (see, e.g., Fehr and Schmidt 1999; Bolton and Ockenfels 2000). Suppose, for instance, that $R > 0$ and a protégé in the hindsight treatment observes that the amount bet is lost. Then, the patron’s payoff would be higher than the protégé’s payoff. An inequality averse protégé may decide for $\Delta < 0$ (seemingly punishing the patron) in order to restore equality. Similar arguments hold for the foresight treatment where protégés who are extremely averse to disadvantageous inequality (i.e., inequality which is to their material disadvantage) may choose Δ such that they ensure to be never worse off than the patron.¹⁰ For example, a protégé who bets 20 ECU should opt for $\Delta = -20$, and a protégé who bets 60 ECU should choose $\Delta = -60$. Obviously protégés’ aversion to disadvantageous inequality may be much less

¹⁰While we acknowledge that preferences for inequality may interact with risk attitudes (for experimental research on this topic, see Amiel et al. 2001; Kroll and Davidovitz 2003; Carlsson et al. 2005; Bolton and Ockenfels 2010), our aim here is to provide a simple scenario that clearly illustrates why we state and test our hypotheses conditional on $R = 0$.

severe (see Fehr and Schmidt 1999). Yet, this aversion would always cause protégés with different positive bets to choose different Δ -values. Since any $R > 0$ entails different actual bets by the protégés, inequality aversion may account for their Δ -choices. Only $R = 0$ guarantees that all protégés make the same actual bet of zero.

2.2 Procedures

The experiment was programmed in Fischbacher's (2007) z-Tree and conducted in the experimental laboratory of the Max-Planck Institute of Economics in Jena (Germany). The participants, undergraduate students from the Friedrich-Schiller University of Jena, were recruited using Greiner's (2004) ORSEE software. Upon entering the laboratory, they were randomly assigned to visually isolated computer terminals.

The full sequence of events unfolded as follows. First, the instructions for the first phase were distributed and read aloud to establish public knowledge. Then, the participants were randomly assigned the role of either patron or protégé. Once the protégés made their bets and the patron chose the restriction, a computerized random draw determined, for each protégé, whether the amount bet was doubled or lost. Protégés in the positive outcome-hindsight treatment group learned that they had doubled the amount bet; protégés in the negative outcome-hindsight treatment group learned that they had lost the amount bet; protégés in the foresight treatment were left in the dark about the lottery outcome. Next, the instructions for the unannounced second phase were displayed on the subjects' screens. The protégés decided whether to increase, decrease, or leave unchanged the patron's total endowment for each and every possible restriction. Afterward, one of the protégés was randomly selected and his choices became payoff relevant for the patron. The restriction set by the patron was revealed to the protégés, the experimental earnings were computed, and the subjects were paid accordingly. Protégés in the foresight treatment were also informed about the realization of the lottery.

Finally, protégés were administered a post-experimental questionnaire asking for background information (such as age and gender), eliciting their beliefs about the intentions motivating the patron's choice, and assessing their emotions toward the patron. The protégés were told to suppose that the patron had allowed no gambling ($R = 0$). Then they had to report

on a seven-point scale with what intensity they experienced both anger and gratitude toward the patron in the case the outcome of the lottery was positive and in the case it was negative. The scale ranged from “no intensity at all” (1) to “very intensely” (7).

We used a between-subjects design where each subject was randomly assigned to one of the three treatment groups. We ran twelve sessions with 30 to 32 subjects in each session. In total, 379 students participated in the experiment and 367 acted as protégés. To balance the number of protégés in each treatment group, we had eight sessions for the hindsight treatment and four sessions for the foresight treatment. This yields data for 118 protégés that knew they had doubled the amount bet (positive outcome-hindsight treatment group), 126 protégés that knew they had lost the amount bet (negative outcome-hindsight treatment group), and 123 protégés that were not informed about the lottery outcome (foresight treatment).

3 Main research hypotheses

Our experiment is expressly designed to test two main hypotheses. In this section, we state the hypotheses in their alternative form (H_1) and rigorously distinguish each one of them from the null (H_0). To this aim, let us denote the mean of Δ conditional on a certain set of events Ω by $E(\Delta \mid \Omega)$.

The first hypothesis tests whether the protégés’ behavior toward the patron is consistent with consequentialist reasoning. It is based on the hindsight treatment data and compares the protégés’ Δ -choices in response to the harshest restriction when the restriction makes them worse off with their Δ -choices when the restriction makes them better off.

Hypothesis 1. *Protégés in the positive outcome-hindsight treatment group will punish more (or reward less) in reaction to $R = 0$ than protégés in the negative outcome-hindsight treatment group:*

$$H_1 : E(\Delta \mid R = 0, \text{positive outcome-hindsight}) < E(\Delta \mid R = 0, \text{negative outcome-hindsight}).$$

The corresponding null hypothesis is:

$$H_0 : E(\Delta \mid R = 0, \text{positive outcome-hindsight}) = E(\Delta \mid R = 0, \text{negative outcome-hindsight}).$$

Rejecting the null hypothesis of no difference in protégés’ behavior when

the restriction proves to be harmful and when it proves to be beneficial provides support for the view that people take a consequentialist position when appraising paternalism. Yet, evidence in favor of Hypothesis 1 does not say anything about whether and to what extent knowledge (as opposed to ignorance) of the consequences influences behavior. To draw inference about this issue, we compare the protégés' Δ -choices in the foresight treatment with their Δ -choices in each of the two hindsight treatment groups. When the consequences of the restriction are unknown (as in the foresight treatment), consequentialist protégés should consider two Δ -values, one if the outcome of the lottery is positive and one if it is negative, and then take some (weighted) average between the two. Conversely, when the consequences of the restriction are known (as in the hindsight treatment), the protégés must decide on just one Δ , without any need to compute averages. This implies that Δ -choices in the foresight treatment should lay in-between those in the two hindsight treatment groups. Specifically, we expect that, compared to Δ -choices conditional on full restriction in the foresight treatment, Δ -choices in the positive outcome-hindsight treatment group will be smaller while Δ -choices in the negative outcome-hindsight treatment group will be larger. Thus, we test:

Hypothesis 2. *Protégés who are prevented from gambling will punish (reward) the patron more (less) in the positive outcome-hindsight treatment group than in the foresight treatment; they will punish (reward) the patron less (more) in the negative outcome-hindsight treatment group than in the foresight treatment:*

$$H_1 : E(\Delta \mid R = 0, \text{positive outcome-hindsight}) < E(\Delta \mid R = 0, \text{foresight});$$

$$E(\Delta \mid R = 0, \text{negative outcome-hindsight}) > E(\Delta \mid R = 0, \text{foresight}).$$

The corresponding null hypothesis is:

$$H_0 : E(\Delta \mid R = 0, \text{positive outcome-hindsight}) = E(\Delta \mid R = 0, \text{foresight});$$

$$E(\Delta \mid R = 0, \text{negative outcome-hindsight}) = E(\Delta \mid R = 0, \text{foresight}).$$

4 Experimental results

4.1 Data description and preliminary tests

On average, over both treatments, the amount that the protégés intend to bet in the lottery is 34 ECU, and the restriction set by the patrons is 77

Table 1: Summary statistics of Δ -choices at each restriction, separately for each treatment group.

R	Positive outcome-hindsight		Negative outcome-hindsight		Foresight	
	Mean (1)	Std. dev. (2)	Mean (3)	Std. dev. (4)	Mean (5)	Std. dev. (6)
0	-13.97	30.37	2.41	36.86	-7.20	33.20
20	-6.07	29.46	1.59	32.30	-4.11	30.07
40	1.08	29.74	2.33	31.13	1.37	28.73
60	8.74	28.02	2.21	32.39	4.65	27.99
80	12.93	27.05	0.61	32.82	6.41	27.68
100	15.72	29.50	3.27	34.79	9.47	30.49

ECU. Table 1 presents descriptive statistics of the protégés' Δ -choices for each possible restriction, separately for each treatment group.¹¹ A graphical representation of the average Δ -choices in the treatments is provided in Figure 1.

Table 1 and Figure 1 show two things. First, average Δ -choices in the positive outcome-hindsight treatment group and the foresight treatment follow a similar pattern (compare columns (1) and (5) in Table 1): they are negative for $R = \{0, 20\}$ and increase monotonically with decreasing restriction. A series of sign tests reveals that Δ -choices for two adjacent restrictions are always significantly different from each other at the 1% level for the positive outcome-hindsight treatment group, and at the 5% level for the foresight treatment. Second, average Δ -choices in the negative outcome-hindsight treatment group (see column (3) in Table 1) are all positive and do not vary significantly across restriction levels. Sign tests confirm that Δ -choices for two adjacent restrictions are never significantly different from each other (not even the extreme restrictions, $R = 0$ and $R = 100$, turn out

¹¹The median values equal zero in all except two cases, both of which refer to protégés in the positive outcome-hindsight treatment group (namely $\text{median}(\Delta) = 0.50$ and 1.50 when $R = 80$ and 100 , respectively).

Figure 1: Bar-diagrams of average Δ -choices by restriction and treatment group. The abbreviations NO-H, PO-H, and F stand for negative outcome-hindsight treatment group, positive outcome-hindsight treatment group, and foresight treatment, respectively. The bar height indicates, for each treatment group, the average Δ in response to the corresponding restriction.

to differ significantly).

To check whether the harshest restriction causes a significant reaction by the protégés, we performed, for each of the three treatment groups, sign tests of the null hypothesis that Δ -choices conditional on $R = 0$ are equal to zero.¹² We are able to reject the null for the foresight and the positive outcome-hindsight treatment group (p -values = 0.004 and 0.000, respectively), but not for the negative outcome-hindsight treatment group (p -value = 0.223). Hence, forbidding gambling triggers significant punishment if the protégés do not know about the lottery outcome and if they know that the restriction makes them worse off. We can interpret the latter finding as a first piece of evidence that people strongly dislike paternalism when it precludes the achievement of better states.¹³

¹²We focus on the harshest restriction because, as pointed out in Section 2.1, this case controls for inequality aversion.

¹³We will provide an interpretation of Δ -choices in the foresight treatment in Section 4.3.

4.2 Testing the main research hypotheses

Here we test the two main hypotheses formulated in Section 3. A comparison of the values in columns (1) and (3) of Table 1 shows that the protégés' reaction to the harshest restriction hinges upon the outcome of the lottery (see also Figure 1). Specifically, average Δ -choices conditional on $R = 0$ are negative (−13.97 ECU) when the protégés know that they would have benefited from gambling, whereas they are positive (2.41 ECU) when the protégés know that they would have been harmed by gambling. Statistical corroboration of the differences in Δ -choices at the harshest restriction between the two hindsight treatment groups is provided by a Mann-Whitney test, which rejects the null of Hypothesis 1 in favor of its alternative (p -value = 0.000). This evidence supports our first research hypothesis that protégés' behavior toward a patron restricting freedom of choice is in line with consequentialist reasoning.

Turning to Hypothesis 2, which compares Δ -choices in foresight and in hindsight, from Table 1 and Figure 1 we can see that protégés who are kept in the dark about the lottery outcome make average Δ -choices that lie between the average Δ -choices made in the positive outcome- and the negative outcome-hindsight treatment groups. Mann-Whitney tests allow the rejection of the null of Hypothesis 2 at the 10% level for the comparison between the foresight and the positive outcome-hindsight treatment group (p -value = 0.096), and at the 1% level for the comparison between the foresight and the negative outcome-hindsight treatment group (p -value = 0.011).¹⁴

This gives support to our second research hypothesis and to the tenet “you will not understand it now, but you will thank me later”. In comparison to behavior in foresight, with the wisdom of hindsight, protégés make larger Δ -choices if they see that the patron's action has prevented them from harm, even though they are not willing to reward the patron to a great extent.

¹⁴Note that we find no difference when comparing Δ -choices made in the foresight treatment to average Δ -choices over the two hindsight treatment groups (p -value = 0.527; two-sided Mann-Whitney test). This suggests that protégés are consistent in the sense that their reaction after observing the realized outcome does not, on average, contradict their reaction in foresight.

4.2.1 Econometric analysis using Tobit regressions

To gain further insight into the relationship between observed lottery outcome and Δ -choices in the hindsight treatment, Table 2 reports the results of three Tobit regressions with individual Δ -choices in response to $R = 0$ in the hindsight treatment as dependent variable. Model 1 is a basic specification, which includes only a treatment dummy (namely the “outcome” of the lottery that equals 1 for the positive outcome-hindsight treatment group) and the protégés’ intended bet x . It confirms the impression from Table 1 and the non-parametric analysis that Δ -values are significantly lower in the positive outcome- than in the negative outcome-hindsight treatment group. Additionally, the model shows that “intended bet” is significantly negative: Δ decreases by close to 0.5 ECU per unit increase in x .

Model 2 adds additional controls for the intentions ascribed by the protégés to the patron, and the protégés’ gender. The variables “outcome” and “benevolence” are included separately as main effects and in combination as interaction effects. Beliefs in the benevolent (rather than malevolent) intentions of the patron lead protégés that are prevented from harm to increase their Δ -choices by approximately 40 ECU, and the effect is significant. However, the large and significant interaction term indicates that benevolent intentions lose importance if the protégés would have doubled the amount bet (the marginal effect reduces to about 7 ECU). The coefficient of “outcome” is still significantly negative: protégés in the negative outcome-hindsight treatment group reduce their Δ -choice by 21.8 ECU if they attribute malevolent intentions to the patron. On the other hand, the coefficient of “benevolence $\times$ outcome” shows that protégés in the positive outcome-hindsight treatment group punish the patron more severely (by close to 55 ECU) although they think of him as benevolent. The results of Model 2 are important as they indicate that the protégés’ reaction toward the patron remains consistent with consequentialist thinking even after controlling for the intentions attributed to the patron. The coefficient of “male” is positive and significant at the 10% level, meaning that male, compared to female, protégés tend to punish (reward) the patron less (more). The present gender effect is in line with the results of Johansson-Stenman (2012), who observes that women are more likely than men to classify behavior as unethical if it violates someone else’s rights.

Finally, in Model 3, we control for the protégés’ feelings toward the

Table 2: Tobit regressions controlling for lottery outcome and intended bet (Model 1), intentions (Model 2), and emotions (Model 3)

Model	Dependent variable: protégés' Δ -choices when $R = 0$ in the hindsight treatment		
	1 (Basic)	2 (Intentions)	3 (Emotions)
outcome (=1 if positive)	-28.868*** (7.376)	-21.809*** (8.177)	-21.040*** (8.138)
intended bet x	-0.463*** (0.139)	-0.498*** (0.137)	-0.440*** (0.140)
benevolence		39.794*** (12.696)	37.729*** (12.538)
male		13.716* (7.259)	13.010* (7.208)
angry _{won}			-1.199 (2.927)
grateful _{lost}			3.854 (2.925)
angry _{won} $\times$ outcome			-0.451 (3.854)
grateful _{lost} $\times$ outcome			-8.652** (4.125)
benevolence $\times$ outcome		-33.411** (16.996)	-28.948* (16.978)
Constant	20.812*** (7.181)	7.948 (7.826)	6.375 (7.862)
Observations	244	244	244
R^2	0.01	0.02	0.03

Note: Standard errors in parentheses.

*** Significant at the 1% level. ** Significant at the 5% level.

* Significant at the 10% level.

patron that sets $R = 0$. The variable “angry_{won}” (“grateful_{lost}”) is a self-reported measure of the intensity of anger (gratitude) felt by the protégé

had he been prevented from gambling and experienced a positive (negative) lottery outcome. These variables are included separately and in combination with “outcome”.¹⁵ The interaction terms allow assessing whether the self-reported emotions vary with the treatment. The main effects of both “angry_{won}” and “grateful_{lost}” are insignificant, meaning that neither anger in case of assumed winnings nor gratitude in case of assumed losses can explain the behavior of the protégés in the negative outcome-hindsight treatment group. The coefficient of “grateful_{lost} × outcome” is significantly negative, i.e., the greater the gratitude of the protégés when they had to imagine that the patron had set $R = 0$ and they had lost, the lower their Δ -choices when the outcome of the lottery turned to be positive. Although emotions seem to correlate with behavior, the coefficient of “angry_{won} × outcome” is not significant, implying that the self-reported intensity of anger in the case of assumed winnings does not explain Δ -choices when the realization of the lottery was indeed positive. Note, finally, that the remaining coefficients in Model 3 are similar in value to those in Model 2.

4.3 How do protégés value freedom of choice?

We devote the last part of the analysis to the question of how protégés may value freedom of choice, distinguishing between two alternative concepts of freedom.¹⁶ The first concept is the *instrumental* one that freedom of choice is “a means to other ends” (Sen 1988) and, thus, has a value insofar as it helps to maximize utility. Under this view, a reduction in the opportunity set does not matter if the utility maximizing solution is included in the resulting subset. The second concept emphasizes the *intrinsic* importance of freedom of choice. From this perspective, freedom has a value in itself and, as such, it increases with the size of the opportunity set. An example taken from Sen (1991, p. 25) illustrates not only the difference between the concepts, but also how we can detect them in our context. Sen considers a person who (1) in a first scenario, decides to read a particular book, Cymbeline;

¹⁵To ease interpretation of the coefficients, the emotion variables are centered. This means that the average self-reported emotion is subtracted from each observation of these variables (the interaction terms are also computed using the centered variables). The effect of “outcome” on the dependent variable is therefore given for the “average” person on the emotion scale.

¹⁶See Verme (2009) for a more comprehensive list of possible views of how people value freedom.

(2) in an alternative scenario, is forced to read another book which he would not have chosen to read; and (3) in a third scenario, is given no choice and simply ordered to read *Cymbeline*. As Sen notes, the person's freedom is certainly reduced in the last two scenarios, but he is not equally unfree in the two cases: in the second scenario he is forced to read a book he would not have chosen, whereas in the third scenario he has to read the book he would have chosen to read anyway. While the instrumental view of freedom would consider the first and third scenarios as equivalent, the intrinsic view would insist that only the first scenario respects one's freedom of choice.

Sen's example makes it evident that a simple way to assess whether protégés in our experiment assign an intrinsic value to freedom is to consider how those who had chosen $x = 0$ in phase 1 react to $R = 0$ in phase 2.¹⁷ It seems reasonable to expect that protégés who value freedom as such should punish a patron forbidding gambling even when the restriction does not apply to them. Thus, using the data from the foresight treatment, we test the null hypothesis that average Δ -choices conditional on $R = 0$ and $x = 0$ are equal to zero, against the alternative that such Δ -choices are negative.¹⁸ A sign test does not allow us to reject the above null hypothesis (p -value = 0.363), indicating that protégés are not willing to punish the patron when they are not affected by the restriction. This result is reminiscent of Ahlert and Crüger's (2004) finding that players, having to choose between an ultimatum game and a dictator game, tend to trade their freedom to veto for monetary rewards.

Protégés for whom freedom has an instrumental role should punish the patron when he sets a restriction that forces them to bet an amount lower than the one they would have bet. Clearly, a binding restriction should trigger a negative reaction also by those who value freedom in itself. Hence, to investigate whether freedom is valued both as an instrument and because of its intrinsic value, we study the reaction to $R = 0$ of the protégés that had chosen $x > 0$. Based on a sign test, we can reject the null hypothesis that average Δ -choices conditional on $R = 0$ and $x > 0$ are equal to zero, in favor of the alternative that these Δ -choices are negative (p -value = 0.001).

¹⁷Recall that any positive restriction would result in different actual bets, so that inequality aversion might explain Δ -choices.

¹⁸Unawareness of the lottery outcome appears to be the best scenario for looking at this issue, since the analysis in Section 4.2 has shown that knowledge of the outcome affects average Δ -values.

This means that the protégés significantly punish the patron when both dimensions of freedom of choice are violated.

Finally, if the instrumental value of freedom is causing the reaction of the protégés, one can expect that, in response to the harshest restriction, the protégés who had chosen $x > 0$ punish more than the protégés who had chosen $x = 0$. A two-sided Mann-Whitney test of the null hypothesis $E(\Delta \mid R = 0, x > 0) = E(\Delta \mid R = 0, x = 0)$ confirms our expectation: the null can be rejected at the 5% level.¹⁹ To sum, protégés refrain from punishing the patron unless they themselves are restricted (i.e., unless both the instrumental and the intrinsic values of freedom are violated). Punishment seems to be triggered mainly by instrumental considerations.

As we considered only Δ -choices in response to $R = 0$ (in order to control for inequality aversion), for the sake of completeness, Figure 2 presents Δ -choices for each possible restriction level, separately for the case in which the restriction was not binding ($x \leq R$) and for the case in which it was binding ($x > R$). The figure shows an interesting pattern. While protégés tend on average to punish the patron if the restriction was binding, they do not choose negative Δ -values if the restriction was not binding. In the latter case, they are even willing, on average, to reward the patron.

5 Summary and concluding remarks

Our goal in this study was to understand how people perceive paternalism. The question we asked was not merely whether people like or dislike paternalistic interventions, but whether people's reaction to paternalism is contingent upon the consequences of the intervention or independent of these. We feel this issue merits attention in view of the increasing and widespread

¹⁹Note that subjects are not randomly assigned to the two groups that are compared. Rather, they are assigned to the groups depending on whether their intended bet was greater than or equal to zero. Consequently, the Δ -choices of the protégés in the two groups may be different for reasons other than whether or not the restriction is binding. For example, protégés who intended to bet zero are likely to be more risk averse than the others. Our conclusions then hold only under the assumption that people with different levels of risk aversion do not value freedom of choice differently. Although we are not aware of any evidence to the contrary, we acknowledge that the results are suggestive rather than conclusive. Not distinguishing between the two concepts of freedom provides a way to make controlled inference. Based on the sign test, we can reject the null hypothesis that average Δ -choices are equal to zero, in favor of the alternative that they are negative (p -value = 0.004). That people value freedom of choice is thereby confirmed.

Figure 2: Protégés' Δ -choices for each restriction and separately for the case where $x \leq R$ and the case where $x > R$. Beginning from the top left panel, sample sizes are as follows: $R = 100$ ($n_{NB} = 123$, $n_B = 0$), $R = 80$ ($n_{NB} = 116$, $n_B = 7$), $R = 60$ ($n_{NB} = 109$, $n_B = 14$), $R = 40$ ($n_{NB} = 84$, $n_B = 39$), $R = 20$ ($n_{NB} = 47$, $n_B = 77$), $R = 0$ ($n_{NB} = 19$, $n_B = 104$).

use of paternalistic practices in all aspects of life, from transport security to retirement savings to consumption behavior.

Although paternalism has been traditionally regarded as wrong because it represents a violation of freedom of choice, advances in behavioral economics have led economists and policy makers to become increasingly prone to supporting paternalistic interventions. After all, it is often argued, if individuals are boundedly rational, paternalistic interventions may help them to avoid mistakes and to make rational decisions. However, what has remained largely unexplored is how the protégés (i.e., those affected by the paternalistic intervention) perceive paternalism. Do the protégés realize that the intervention is in their own interest? Do they instead consider the intervention as an usurpation of their freedom? Or, rather, does their behavior toward paternalism depend on whether the paternalistic intervention proved to be beneficial or harmful to them? To remain in the setting envisioned in our experiment, consider people who, against their expressed wish, are not allowed to gamble. They may perceive this prohibition as positive because it is likely to keep them away from poverty. Alternatively, they may regard the prohibition as a violation of the liberty to decide for themselves whether or not to risk their money. Or the reaction to the prohibition may depend on whether its consequences have been experienced and proved to enhance well-being. According to a consequentialist perspective, after losing money gambling, the protégé may appreciate the prohibition.

Our results suggest that with the wisdom of hindsight (i.e., being aware of the consequences of the paternalistic act to their own payoff), protégés take predominantly a consequentialist stand on paternalism. Protégés punish an impartial and randomly appointed patron who forbids gambling when this action makes them worse off, and, notably, they reward the patron when the action makes them better off. Additionally, in comparison to protégés who are kept in the dark about the consequences of the restriction, protégés who know that they would have won, had they gambled, punish the patron more heavily, whereas protégés who know that they would have lost are willing to reward the patron. The regression results reveal that the protégés' reaction to paternalism remains consistent with consequentialist reasoning even after controlling for the intentions (malevolent vs. benevolent) that the protégés attribute to the patron.

We interpret our results as evidence that the experienced or divulged consequences of paternalistic interventions play an important role when it comes to the evaluation of the interventions. Folk intuitions, which are es-

essential for the successful implementation and communication of any policy, do not reject paternalism on principled grounds (i.e., on grounds that value freedom of choice in itself). The finding that consequences have a significant impact upon protégés' reaction to paternalism may persuade governments either to undertake paternalistic policies well before the election year so that their presumed positive effect becomes visible or to promote the salience of prevented harm to citizens. Yet, we are hesitant to give sound policy advice based on the experimental results we present. On the one hand, people's actual perception of paternalism is still at an early stage of scrutiny to make any conclusive suggestion. On the other hand, the possibility of self-justifying paternalism (i.e., "paternalistic behavior which effects its own justification", Archard 1993, p. 341) needs to be considered. Some scholars worry that allowing paternalism to be justified by subsequent consent might lead paternalistic governments "to create that very consent" by somewhat distorting people's preferences and desires (e.g., Carter 1977; Van de Veer 1979; Archard 1993). Uncovering evidence of self-justifying paternalism could provide a fertile ground for future experimental work.

References

- Ahlert, Marlies, and Arwed Crüger.** 2004. "Freedom to Veto." *Social Choice and Welfare* 22(1): 7–16.
- Amiel, Yoram, Frank A. Cowell, and Avraham Polovin.** 2001. "Risk Perceptions, Income Transformations and Inequality." *European Economic Review* 45(4): 964–976.
- Appiah, Kwame A.** 2008. *Experiments in Ethics (Flexner Lectures)*. Cambridge: Harvard University Press.
- Archard, David.** 1993. "Self-justifying Paternalism." *The Journal of Value Inquiry* 27: 341–352.
- Bolton, Gary E., and Axel Ockenfels.** 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90(1): 166–193.
- Bolton, Gary E., and Axel Ockenfels.** 2010. "Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States: Comment." *American Economic Review* 100(1): 628–633.
- Brandts, Jordi, and Gary Charness.** 2011. "The Strategy Versus the Direct-Response Method: A First Survey of Experimental Comparisons." *Experimental Economics* 14: 375–398.
- Carter, Rosemary.** 1977. "Justifying Paternalism." *Canadian Journal of Philosophy* 7(1): 133–145.
- Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin.** 2003. "Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism"." *University of Pennsylvania Law Review* 151(3): 1211–1254.
- Carlsson, Fredrik, Dinky Daruvala, and Olof Johansson-Stenman.** 2005. "Are People Inequality-Averse, or Just Risk-Averse?" *Economica* 72(287): 375–396.
- Dworkin, Gerald.** 2010. "Paternalism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. <http://plato.stanford.edu/archives/sum2010/entries/paternalism/>

- Fehr, Ernst, and Klaus M. Schmidt.** 1999. "A Theory of Fairness, Competition and Cooperation." *Quarterly Journal of Economics* 114: 817–868.
- Fischbacher, Urs.** 2007. Zurich Toolbox for Readymade Economic Experiments. *Experimental Economics* 10(2): 171–178.
- Greene, Joshua, Fiery Cushman, Lisa Stewart, Kelly Lowenberg, Leigh Nystrom, and Jonathan Cohen.** 2009. "Pushing Moral Buttons: The Interaction between Personal Force and Intention in Moral Judgment." *Cognition* 111(3): 365–371.
- Greiner, Ben.** 2004. "An online recruitment system for economic experiments." In *Forschung und wissenschaftliches Rechnen 2003*, edited by Kurt Kremer and Volker Macho, 79–93. Göttingen: GWDG Bericht 63, Ges. für Wiss. Datenverarbeitung.
- Gruber, Jonathan, and Botond Köszegi.** 2008. *A Modern Economic View of Tobacco Taxation*. Paris: International Union Against Tuberculosis and Lung Disease.
- Honderich, Ted.** 1995. *The Oxford Companion to Philosophy*. Oxford: Oxford University Press.
- Johansson-Stenman, Olof.** 2012. "Are Most People Consequentialists?" *Economic Letters* 115(2): 225–228.
- Knobe, Joshua.** 2003 "Intentional Action in Folk Psychology: An Experimental Investigation." *Philosophical Psychology* 16: 309–324.
- Kroll, Yoram, and Liema Davidovitz.** 2003 "Inequality Aversion versus Risk Aversion." *Economica* 70(277): 19–29.
- O'Donoghue, Ted, and Matthew Rabin.** 2003. "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes." *American Economic Review* 93(2): 186–191.
- Schelling, Thomas C.** 1984. "Self-Command in Practice, in Policy, and in a Theory of Rational Choice." *American Economic Review Papers and Proceedings* 74(2): 1–11.
- Sen, Amartya.** 1987. *On Ethics and Economics*. New York: Basil Blackwell.

- Sen, Amartya.** 1988. "Freedom of Choice: Concept and Content." *European Economic Review* 32: 269–294.
- Sen, Amartya.** 1991. "Welfare, Preference and Freedom." *Journal of Econometrics* 50(2): 15–29.
- Thaler, Richard H., and Cass R. Sunstein.** 2003. "Behavioral Economics, Public Policy, and Paternalism: Libertarian Paternalism." *American Economic Review* 93(2): 175–179.
- Thaler, Richard H., and Cass R. Sunstein.** 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Thomas, Matthew, and Luke Buckmaster.** 2010. "Paternalism in Social Policy - When Is It Justifiable?" Parliament of Australia, Department of Parliamentary Services, Research Paper No. 8.
- Van de Veer, Donald.** 1979. "Paternalism and Subsequent Consent." *Canadian Journal of Philosophy* 9(4): 631–642.
- Verme, Paolo.** 2009. "Happiness, Freedom and Control." *Journal of Economic Behavior & Organization* 71(2): 146–161.

Appendix: Experimental instructions (translated from German)

Instructions for phase 1 (distributed and read aloud)

Welcome and thanks for participating in this experiment!

This is an experiment which involves gambling. More than half of all Germans gamble. Studies have found that young people are particularly prone to gambling. Gambling is considered problematic in many societies and often heavily regulated.

In this experiment, we shall speak of ECU (Experimental Currency Units) rather than euros. Your payoff in ECU will be converted to Euros at the end of the experiment with a conversion rate of $1 \text{ ECU} = 0.10 \text{ Euros}$.

One of you will be randomly assigned to role B. The remaining participants will be assigned to role A.

The participants in role A will be endowed with 100 ECU. They can bet any integer amount up to their endowment in a gamble. Depending on the result of a random draw, the amount bet will be either doubled with probability $\frac{1}{2}$ or lost with probability of $\frac{1}{2}$. The participant in role B may, however, impose a gambling limit to the participants in role A.

The participant in role B will be endowed with 50 ECU that she will keep irrespective of her choice. She will be asked to place a gambling limit on the participants in role A. If the bet that a participant in role A wanted to make is higher than the imposed gambling limit, the bet will only be placed up to the imposed limit.

The gambling limit set by the participant in role B is the maximal amount that the participants in role A are allowed to bet in the gamble. The participant in role B must select one of the following gambling limits: 0 ECU (no gambling is allowed), 20 ECU (allowing to bet up to 20 ECU), 40 ECU (allowing to bet up to 40 ECU), 60 ECU (allowing to bet up to 60 ECU), 80 ECU (allowing to bet up to 80 ECU), or 100 ECU (setting no gambling limit). The participants in role A are only able to place a bet up to the imposed gambling limit.

If you have any questions, please raise your hand and an experimenter will approach you.

Instructions for phase 2 (displayed on screens)

The protégés read:

Before informing you about the decision of the participant in role B and your consequent payoff, we kindly ask you to participate in a second experiment.

In this experiment, the participant in role B is endowed with additional 50 ECU, which gives her a total payoff of 100 ECU.

You may, however, increase or decrease B's payoff by any integer amount between 0 and 50 ECU. You may, of course, leave B's payoff unchanged. None of these choices will cost you anything.

Since you do not know the level of the gambling limit imposed by the participant in role B, you will have to make this choice for each possible limit that B may have imposed.

Every participant in role A will make these choices. One of you will be randomly selected, and the choice that she made for the gambling limit actually set by the participant in role B will be implemented, i.e., B's payoff will be altered according to this choice.

The patron read:

We kindly ask you to participate in a second experiment.

In this experiment, you are endowed with additional 50 ECU, which gives you a total payoff of 100 ECU.

The participants in role A will decide whether to increase or decrease your payoff by any integer amount between 0 and 50 ECU. They will make this choice for each gambling limit that you may have imposed.

Every participant in role A will make these choices. One of them will be randomly selected, and the choice that she made for the gambling limit you actually set will be implemented, i.e., your payoff will be altered according to this choice.

Please wait until the participants in role A have made their decisions.

Post-experimental questionnaire: eliciting beliefs about the patron's intentions

Why do you think that the participant in role B set a limit to your choice? Please, place a '1' in the option that you think is the most correct one, a '2' in the option that you think is the second to most correct one, and so on until you have rated all the eight options.

- ☐ She has a moral conviction that gambling is bad.
- ☐ She does not want to see that I lose money.
- ☐ She does not want to see that I win money.
- ☐ She gets joy from being in control.
- ☐ She is spiteful.
- ☐ She is benevolent.
- ☐ She is envious that she did not have the opportunity to gamble.
- ☐ She is envious that she had a lower endowment.

Post-experimental questionnaire: assessing emotions toward the patron

In the experiment you decided to bet ... ECU. Suppose the participant in role B limited your bet to 0 ECU. You now accidentally meet this member. Please indicate your feelings towards this person in the following situations:

Positive lottery outcome

Your bet was lucky and you won 0 ECU. If the participant in role B would have not restricted your bet, you would have won ... ECU.

- Please indicate your anger toward the participant in role B.
- Please indicate your gratitude toward the participant in role B.

Negative lottery outcome

Your bet was unlucky and you lost 0 ECU. If the participant in role B would have not restricted your bet, you would have lost ... ECU.

- Please indicate your anger toward the participant in role B.
- Please indicate your gratitude toward the participant in role B.

[Participants indicated their feelings of anger and gratitude on a 7-point scale].