Social Preferences or Sacred Values? Theory and Evidence of Deontological Motivations

Daniel L. Chen and Martin Schonger*

Latest version at: http://nber.org/∼dlchen/papers/Deontological.pdf

August 2013

Abstract

Recent advances in economic theory, largely motivated by experimental findings, have led to the adoption of models of human behavior where a decision-maker not only takes into consideration her own payoff but also others' payoffs and any potential consequences of these payoffs. Investigations of deontological motivations, where a decision-maker makes her choice not only based on the consequences of a decision but also the decision per se have been rare. We propose an experimental method that can reveal individual's deontological motivations by varying the probability of the decision-maker's decision being consequential. It uses two states of the world, one where the decision has consequences, and one where it has none. A decision-maker whose preferences satisfy first-order stochastic dominance will choose the decision that leads to the best consequences regardless of the probability of the consequential state. A mixed consequentialist-deontological deicision-maker's choice changes with the probability. The direction of change indicates how deontological motivations are incorporated into preferences. Our model and experimental findings suggest that when moral decisions are involved, the random lottery incentive method of eliciting preferences results in responses different than when their choice is implemented with 100% certainty. Donation decisions are 50-85% more generous when implemented with low probability. Calibrations suggest deontological motivations constitute 40-58% of moral decision-making.

Keywords: Consequentialism, Deontological Motivations, Normative Commitments, Social Preferences, Revealed Preference, Decision Theory, First Order Stochastic Dominance, Random Lottery Incentive Method

JEL codes: D6, K2

*Daniel L. Chen (dlchen@post.harvard.edu) and Martin Schonger (mschonger@ethz.ch) both Center for Law and Economics ETH Zürich. For comments we thank seminar participants at ALEA, Bar-Ilan, Basel, Bonn, ESA, Haifa, Hebrew, MPSA, Tel Aviv, and Zurich (ETH & University). This project was conducted while Chen received funding from the Institute for Humane Studies, the John M. Olin, the Ewing Marion Kauffman, and the Templeton Foundations. This paper is preliminary, please ask before citing or circulating.

1 Introduction

Economics has traditionally taken a consequentialist view of both individual and social behavior, at least if one defines, as we do, consequentialism broadly. In the conventional homo oeconomicus model, a decision-maker (DM) only cares about her own payoff. Behavioral economists have extended this model to allow for a DM to take into account others' payoffs (Andreoni 1990; Rabin 1993; Fehr and Schmidt 1999). Even more recently, models and experimental evidence have shown that people may also care about others' feelings and attitudes as a result of one's decision. Under a broad conception of consequences, like ours, even these non-monetary outcomes are consequences, but note that some of the literature (Gneezy 2005) has a narrower view of consequences and therefore use the term "non-consequential" to refer even to what we would view as consequences such as others' feelings. The question of this paper is whether and in which situations the range of motivations considered should be extended further, to allow for non-consequentialist, specifically deontological (internal duty-oriented) motivations. Deontological motivations are present when people care about their decisions per se. Any direct or indirect consequence arising from payoffs or others' observation of or inference about the DM's behavior, intentions or type, we still consider to be consequentialist. In this delineation, we try to adapt major concepts of moral philosophy to economics, and to bring the precision of economic methodology, in particular revealed preference, to moral philosophy. Philosophers and legal theorists commonly assume that people have deontological motivations (Greene and Cohen 2001; Mikhail 2007). While moral philosophers have long debated whether people should have deontological preferences (Kant 1785; Nagel 1978), we investigate the positive question of whether people have deontological motivations and what are people's deontological motivations.

Deontological motivations do not coincide with what is sometimes called duty in the political science or behavioral economics literature. For example people may participate in elections even when their vote is not pivotal due to a "duty" to vote (Riker and Ordeshook, 1968). But participation in elections is observable to family members and neighbors, and thus the word duty may just refer to fulfilling one's social obligations, i.e. be motivated out of concern for the consequences of one's actions rather than deontologically (for evidence that people vote out of these consequentialist social image concerns see DellaVigna et al., 2013). Moreover, people might also worry how large the mandate is or send a message by voting for a particular candidate. Such expressive voting falls under our broad conception of consequences and is thus not deontological.

A reason for the lack of attention to deontological motivations is likely due to the difficulty of designing studies that can detect and distinguish the presence of non-consequentialist motivations. This paper makes three contributions: It formalizes the notion of consequentialist as well as deontological motivations as properties of preference relations; it suggests a method to use revealed preference to detect deontological motivations; and, using that method, provides experimental evidence for the existence and content of deontological motivations, and thereby hopes to contribute to the understanding of human behavior, specifically behavior that might be viewed as morally motivated or constrained.

Before previewing our formal interpretation of the philosophical concepts of consequentialist and deontological moral philosophy, let us review what they are. Sinnott-Armstrong (2012) define consequentialism as, "the view that normative properties depend only on consequences" and explains that "[c]onsequentialists hold that choices — acts and/or intentions — are to be morally assessed solely by the states of affairs they bring about." Utilitarianism is one example of a consequentialist moral philosophy (Bentham 1791), in fact any welfarist view is consequentialist (Arrow 1970). By contrast, deontological ethics holds that "some choices cannot be justified by their effects — that no matter how morally good their consequences, some choices are morally forbidden."(Alexander and Moore 2012). Immanuel Kant, one of the most famous proponents of deontological ethics for instance claims that lying is absolutely prohibited, even if the lie brought about great good. Famously Kant thinks that it is morally impermissible to even lie to an ax-murderer about the whereabouts of a friend whom the former is in pursuit of (Kant 1797).

We propose the following formalization of these moral philosophies: Consider a decision d, that may cause (a vector of) consequences x. Assume that preference can be represented by a utility function. Then the preferences of a (pure) consequentialist can be represented by a utility function of the form $u = u(x)$. If in addition to consequentialist motivations the DMalso has deontological ones, then the utility function takes the form $u = u(x, d)$. At the other extreme, a DM who is purely deontological would have a utility function of the form $u = u(d)$. Or more precisely, since deontological ethics unlike consequentialism is supererogatory and thus may not lead to a unique morally permissible decision, they are lexicographic with the first component being based only on d, and the second component being based only on x. But as a shortcut to grasp the intuition, it is helpful to think of a purely deontological DM to have preferences represented by $u = u(d)$. It may seem odd to model deontological motivations by utility functions since one may view "utility" as a consequence, but since ours is a revealed preference approach, we follow the usual economics approach (Friedman and Savage, 1948) of modeling decision makers behaving as if they maximized that objective function and refrain from interpreting the function as standing for utility or happiness.

Since decisions and consequences are closely tied together, it is non-obvious how to cleanly distinguish between these motivations. We provide a method to identify deontologicalism. A key aspect of our thought experiment is to vary the probability that one's moral decision is consequential (i.e. carried out). For a consequentialist, the optimal decision is independent of the probability that the action will be enacted, because the marginal cost (e.g. lost money or time) and marginal benefit (e.g. recipient's well-being) are both affected equally by the probability. For a deontologist, the optimal decision is also independent of the probability, since the duty to make a decision is unaffected by the probability. Only mixtures of both consequentialist and deontological motivations predict changes in behavior as the probability changes.

Our thought experiment can be viewed as testing the joint hypothesis of consequentialism and first-order stochastic dominance (FOSD): if your decision has only a consequence with a certain probability and something outside of your control happens, the probability cannot affect your choice of best action. Behavioral changes with the reduction in probability of an act being executed indicate that people care about their decisions even when they are inconsequential. For example, as the likelihood of actually having to pay for the moral decision decreases, the marginal cost to making the moral decision decreases while the feeling of duty (internal non-consequentialist price of the moral decision) to make the moral choice remains the same, so the decision should typically become more moral. Only those who have both motivations and trade them off change their behavior.

The direction of change indicates how consequentialist and deontological motivations are incorporated into preferences. Under additive utility and preferences that are globally convex, a decrease in the probability of being consequential leads to a decrease in marginal cost (external consequentialist price of the moral decision), so to equate marginal costs and marginal benefits, the decision becomes more moral. The decision can become less moral if consequentialist and deontological motivations are non-additive or if deontological motivations are such that utility from the decision per se decreases if the DM decides to do more than what duty calls for (for example, if utility functions are not globally convex).

We report the results of two experiments. Our shredding experiment involves undergraduates making a decision to donate to Doctors Without Borders. Since the DM may care about what the experimenter thinks of the decision and the expected benefit of the decision per se remains even in the non-consequential state, we shred the envelopes containing the decisions that are not implemented, thus, no one ever knows the decision in the non-consequential state. Our MTurk experiment involves data entry workers splitting a bonus with the Red Cross. We find that when the probability of being consequential decreases, altruism increases. We use demographic information to explore who has mixed consequentialist-deontological motives. We then structurally estimate the relative weights individuals place on consequentialist and deontological motivations and the location of one's sacred values.

Two aspects of the experimental design address possible confounds, where individuals are consequentialist but violate FOSD. First, the DM may wish to target the ex ante utility of individuals. Such a DM, would, as the probability of her decision being implemented declines, donate more. We design a treatment arm where the non-consequential state of the world involves the entire sum being donated. Our results are not consistent with targeting of ex ante utility. Second, the decision may be cognitively costly, and so the cost of the decision per se remains even in the non-consequential state. If cognitive costs change with the probability, subjects make be more likely to report their heuristic decision, which may be a more moral one (Rand et al. 2012). We measure time spent making the decision.

The remainder of the paper is organized as follows. Using a thought experiment, section 2 defines consequentialism, deontologicalism, and mixed motivations as properties of a preference relation and proves that under first-order stochastic dominance, behavior is invariant to the probability. Section 3 describes how we implemented the thought experiment, and section 4 reports the reduced form results and structural estimates of deontological motivations. Section 5 discusses the relation of this paper to important strands of related literature, such as expressive voting, intentions-based preferences, self-signaling and repugnance. Section 6 discusses policy implications and concludes.

2 Formal investigation

2.1 Thought Experiment

The idea to identify non-consequentialist motivations by varying the probability of the DM's decision being consequential guides this paper. The DM has a real-valued choice variable d which influences both her own monetary payoff x_1 as well as the payoff x_2 of a recipient R. There are two states of the world, state C and state N. In state C, the DM's decision d fully determines both x_1 and x_2. In state N, both x_1 and x_2 take exogenously given values, and the decision d has no impact at all. Thus, in state C, the decision is consequential, while in state N, it is not. After DM chooses d, nature randomly decides which state is realized. State C occurs with probability $\pi > 0$, state N with probability $1 - \pi$. The structure of the game

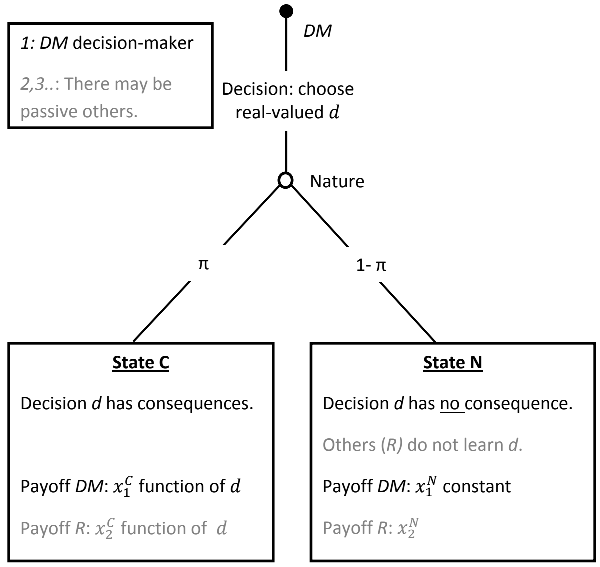


Figure 1:

is public, but the decision d is only known to DM. In state N, therefore, R has no way of knowing d, but, in state C, R knows d, indeed he can infer it from x_2. Superscripts indicate the realized state, so that the payoffs are (x_1^C, x_2^C) in state C, and (x_1^N, x_2^N) in state N. Figure 1 illustrates this.

This general experimental design could be used for many morally relevant decisions; here we apply our identification method to the dictator game and thus to the moral decision to share. As shown in figure 2, the DM gets an endowment of ω, and her decision is how much to give to R. She may choose any d such that $0 \leqslant d \leqslant \omega$ and the resulting payoffs are $x_1^C = \omega - d$ and $x_2^C = d$. For $\pi = 1$, the game thus reduces to the standard dictator game. In state N, a pre-determined, exogenous κ will be implemented, where $0 \leqslant \kappa \leqslant \omega$, and $x_1^N = \omega - \kappa$ and $x_2^N = \kappa$ are the resulting payoffs.

2.2 A testable implication of the standard framework

In the following we sketch the standard, consequentialist approach to choice under uncertainty where the central assumption for choice behavior regarding uncertainty is first-order stochastic dominance (FOSD). A wide variety of models of choice under uncertainty satisfies FOSD and thus fall within this framework, among them most prominently, expected utility theory, its generalization by Machina (1982), but also cumulative prospect theory (Tversky and Kahneman, 1992) or rank-dependent utility theory (Quiggin, 1982).

In the following paragraph and the axioms up to FOSD, we closely follow the canonical framework as laid out in Kreps (1988). Let there be outcomes x. x can be a real valued vector.

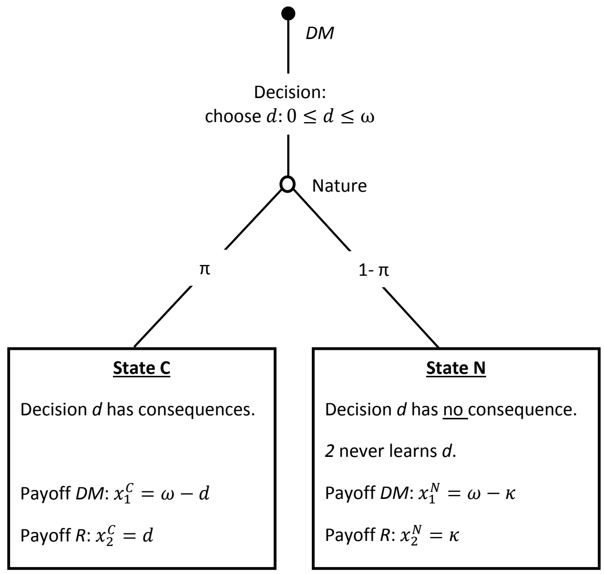


Figure 2:

In the thought experiment, it would be $x = (x_1, x_2)$. Let the set of all x be finite and denote it by X. A probability measure on X is a function $p : X \rightarrow [0, 1]$ such that $\sum_{x \in X} p(x) = 1$. Let P be the set of all probability measures on X, and therefore it, in the thought experiment, a subset of it, is the choice set of the decision-maker. Axiom 1 is the standard one saying that the preference relation is a complete ordering. It implicitly includes consequentialism since the preference relation is on P, that is the over lotteries that are over consequences x.

Axiom 1. *(preference relation) Let $\succsim$ be a complete and transitive preference on P.*

Next we define first-order stochastic dominance (FOSD). Often definitions of FOSD are suitable only for preference relations that are montonic in the real numbers, for example see Levhari et al. (1975). These definition define FOSD with respect to the ordering induced by the real numbers, assuming that prices are vectors. Such an approach is inapproporiate in the context of social preferences which are often not monotonic due to envy or fairness concerns. For example, Fehr and Schmidt (1999) preferences, which merely ordinally rank allocations would violate such definitions of FOSD since they do not satisfy monotonicity. This is a problematic since Fehr Schmidt preferences are merely ordinal rankings of certain prospects and do not convey any attitude of the DM about risk.

Definition. (FOSD) p first-order stochastically dominates q with respect to the ordering induced by $\succsim$, if for all x':

$$\sum_{x:x' \succsim x} p(x) \leq \sum_{x:x' \succsim x} q(x).$$

Axiom. *(FOSD) If p FOSD q with respect to the ordering induced by $\succsim$, then $p \succsim q$.*

Definition. (Strict FOSD) p strictly first-order stochastically dominates q with respect to the ordering induced by $\succsim$ if p FOSD q with respect to that ordering, and and there exists an x' such that:

$$\sum_{x:x' \succsim x} p(x) < \sum_{x:x' \succsim x} q(x).$$

Axiom. *(Strict FOSD) If p strictly FOSD q with respect to the ordering induced by $\succsim$, , then $p \succ q$.*

The following theorem implies that in our thought experiment changing the probability of being consequential π does not change the decision. It is this prediction of the theory that we will test and interpret a rejection of the prediction as evidence that people are not purely consequentialist.

Theorem 1. *If the DM satisfies the axioms Preference Relation, FOSD, and Strict FOSD, and there exist $x, x', x'' \in X'$ and $\pi \epsilon (0;1]$ such that $\pi x + (1-\pi)x'' \not\prec \pi x' + (1-\pi)x''$, then for all $\pi' \epsilon (0;1]$: $\pi' x + (1-\pi')x'' \not\prec \pi' x' + (1-\pi')x''$.*

Proof. (i) $x \succsim x'$: Suppose not, then $x' \succ x$, and therefore $\pi x' + (1-\pi)x''$ strongly first-order stochastically dominates $\pi x + (1-\pi)x''$. Then by axiom Strong FOSD $\pi x' + (1-\pi)x'' \succ \pi x + (1-\pi)x''$, a contradiction.

(ii) Since $x \succsim x'$, $\pi' x + (1-\pi')x''$ first-order stochastically dominates $\pi' x' + (1-\pi')x''$. Thus by axiom FOSD $\pi' x + (1-\pi')x'' \succ \pi' x' + (1-\pi')x''$. □

The theorem has a corollary for the case of expected utility:

Corollary. *If the decision-maker satisfies axiom Preference Relation and maximizes expected utility and there exist $x, x', x'' \in X'$ and $\pi \epsilon (0;1]$ such that $\pi x + (1-\pi)x'' \not\prec \pi x' + (1-\pi)x''$, then for all $\pi' \epsilon (0;1]$: $\pi' x + (1-\pi')x'' \not\prec \pi' x' + (1-\pi')x''$.*

The corollary holds since expected utility's indpendence axiom implies the axioms of FOSD and Strong FOSD. Note that in the thought experiment and experimental setup, the only way the recipient can learn about the decision is if the decision is implemented. d affects the recipient only via the payoff x_2^C. Thus, the theorem applies even to situations where the DM cares about not only the recipient's outcome but also about the recipient's opinion or feelings about the DM or her decision d. Thus, for consequentialist preferences, even allowing such consequences as others' opinion or the impact that the opinion has on one's self-identity, the DM's optimal split does not depend on the probability of the DM's split being implemented.

2.3 Defining consequentialism and deontic motivations

While the previous subsection was very general in order to demonstrate an impossibility, namely to explain variance in the probability in a consequentialist framework, we can now become less general and more concrete when we want to show what a model could for example explain the behavior. We now assume that the DM has a state-independent utility function u that ranks certain outcomes and that it is twice continuously differentiable with strictly positive first derivatives with respect to the consequences (we will relax that assumption again for purely deontological preferences which will be lexicographic). Under expected utility that u can then be chosen such that is the DM's Bernoulli utility function. We allow the utility u of the DM to be a function of her own monetary payoff x_1, as well as the monetary payoff of the recipient x_2 to capture consequentialist other-regarding motives, and d to capture deontological motives. So the main difference to the previous subsection is that we extend the domain of the preference beyond consequences to decisions.

In the general case with all motivations present, the Bernoulli utility function satisfies $u = u(x_1, x_2, d)$. Here we can see what identifies non-consequentialist motivations. In state N, the decision d has no consequence for payoffs or for what others think or know about the DM, yet the decision does enter the utility function equally in all states of the world. This general framework now allows us to formalize the notion of decision-makers that are purely consequentialist, purely deontological, and consequentialist-deontological. Consequentialist preferences are preferences that depend on monetary payoffs and other consequences such as others' opinion of the DM. Conventional preferences in economics, even other-regarding ones, are all consequentialist.

Definition 1. CONSEQUENTIALIST PREFERENCES: A preference is *consequentialist* if there exists a utility representation u such that $u = u(x)$.

We call a preference consequentialist-deontological if it incorporates concerns beyond the consequences, and considers actions or decisions good or bad per se:

Definition 2. CONSEQUENTIALIST-DEONTOLOGICAL PREFERENCES: A preference is *consequentialist-deontological* if there exists a utility representation u such that $u = u(x, d)$.

Now let us turn to purely deontological preferences. At first one might think that they are just the mirror other extreme of consequentialist preferences and could thus be represented by $u = u(d)$. But since duty is like an internal moral constraint, even fully satisfying one's duty may leave the DM with many morally permissible options, rather than a unique one. As the Stanford Encyclopedia of Philosophy (Alexander and Moore, 2012) puts it, "Deontological moralities, unlike most views of consequentialism, leave space for the supererogatory. A deontologist can do more that is morally praiseworthy than morality demands. A consequentialist cannot. For the consequentialist, if one's act is not morally demanded, it is morally wrong and forbidden. For the deontologist, there are acts that are neither morally wrong nor demanded." We could model duty as a moral, and thus internal constraint on the DM set of feasible decisions. But we decide not to follow down that path and rather model these internal constraints as the first component of a lexicographic preference. The reason we model duty not like a budget constraint but as part of preferences, and thus lexicographic is twofold: first, unlike budget constraints internal moral constraints are not directly observable, and second for consequentialist-deontological preferences which feature a tradeoff, not a lexicographic ordering of these motivations, one could not model duty as an inviolable constraint. This can be formalized as a lexicographic preference, with deontological before consequentialist motivations. Note that while it is attractive to think of our method as detecting which of competing duties a DM feels more duty towards, there is no possibility of a genuine conflict of duties in deontological ethical theory, which can distinguish between a duty all-other-things-being-equal (prima facie duty) and a duty-all-things-considered (categorical duty) (Alexander and Moore, 2012).

Definition 3. Deontological Preferences: A preference is called *deontological* if there exist u, f such that $u = u(d)$, and $f = f(x)$, and f.a. $(x, d), (x', d')$: $(x, d) \succsim (x', d')$ if and only if $u(d) > u(d')$ or $[u(d) = u(d')$ and $f(x) \geqq f(x')]$.

Observable choice behavior then allows us to experimentally identify whether subjects have preferences where both motivations are present, i.e. whether their preferences belong into the category of consequentialist-deontological preferences. We will in particular ask how exogenous variation in the probability π of the decision being consequential impacts the optimal decision. Note that the DM has one choice variable only, d, but by varying the probability of her decision being consequential we can identify whether she cares only about the consequences or also about the decision per se. Since she has only one choice variable it is often useful to consider her indirect objective function $V(d)$.

2.4 Consequentialists who maximize Expected Utility

Given expected utility the DM maximizes

$$E[u(x, d)] = \pi u(x_1^C, x_2^C, d) + (1\text{-}\pi) u(x_1^N, x_2^N, d)$$

and her indirect objective function in case of the dictator game can be written as:
$V(d) = \pi u(\omega - d, d, d) + (1-\pi)u(\omega - \kappa, \kappa, d)$.
Limiting attention to pure consequentialists the problem simplifies to:

$$E[u(x)] = \pi u(x_1^C, x_2^C) + (1-\pi)u(x_1^N, x_2^N)$$

and the indirect objective function to:

$$V(d) = \pi u(\omega - d, d) + (1-\pi)u(\omega - \kappa, \kappa).$$

Note that now the d does not enter in the second term, which corresponds to state N at all.

Let us first consider the simplest example of a consequentialist preference, homo oeconomicus:

Example 1. (Homo oeconomicus) Homo oeconomicus is a consequentialist whose preferences depend only on her own outcome. Her preference can be represented by a Bernoulli utility function with $u = u(x_1)$. Her constrained maximization problem is thus $max_d V(d) = \pi u(\omega - d) + (1-\pi)u(\omega - \kappa)$ subject to $0 \leq d \leq \omega$. As the objective function is proportional to $u(\omega - d)$ the unique maximizer is $d^* = 0$. Observe that the optimal decision d^* of homo oeconomicus does not depend on the probability π of the decision being consequential.

Is this independence of the optimal decision d^* true for consequentialist preferences more generally? Intuitively, as the probability of the sharing decision being implemented varies, both its benefits and costs vary in the same way. Let us investigate this:

Next, let us consider another example of a consequentialist preference, Fehr-Schmidt:

Example 2. (Fehr-Schmidt preferences) Fehr and Schmidt (1999) propose preferences such that the DM is a consequentialist who cares about her own and others' monetary payoffs. The idea is that decision-makers dislike inequality, but dislike inequality in their disfavor even more. Recall that the Fehr-Schmidt utility function is $u(x) = x_1 - \alpha max\{x_2 - x_1, 0\} - \beta max\{x_1 - x_2, 0\}$, where $\beta < \alpha$ and $0 \leq \beta < 1$. We can write the decsion-maker's expected utility as

$$\begin{aligned} E[u(x)] =& \pi \left(x_1^C - \alpha max\{x_2^C - x_1^C, 0\} - \beta max\{x_1^C - x_2^C, 0\} \right) \\ &+ (1-\pi) \left(x_1^N - \alpha max\{x_2^N - x_1^N, 0\} - \beta max\{x_1^N - x_2^N, 0\} \right) \end{aligned}$$

The indirect objective function is then:

$$V(d) = \pi\left(\omega - d - \alpha max\{2d-\omega,0\} - \beta max\{\omega - 2d, 0\}\right) \\ + (1-\pi)\left(\omega - \kappa - \alpha max\{2\kappa - \omega, 0\} - \beta max\{\omega - 2\kappa, 0\}\right)$$

V obtains a maximum wherever the first summand does. Thus as usual when an FS-decisionmaker is confronted with a DG she will choose $d = 0$ if $\beta < \frac{1}{2}$, and $d = \frac{\omega}{2}$ if $\beta > \frac{1}{2}$, and for $\beta = \frac{1}{2}$ she is indifferent between all donations that are no more than half the endowment. The optimal donation does not depend on the probability.

Another famous example of social preferences are Andreoni preferences:

Example 3 (**Andreoni preferences**). In a seminal paper Andreoni (1990) points out that DMs in a public goods contribution framework empirically seems to derive utility (i.e behave as if) not only from the total amount of the public good G provided, but also from her contribution g. First note that in the Andreoni-framework there is no nonconsequential state so even from a theoretical point of view, one cannot tell the consequential and the deontological apart. Both interpretations are consistent with the formal model, but the verbal description of "impure altruism" suggests a consequentialist understanding, that is G is a consequence of the decision as are social reactions to the generosity of the individual decision. Thus a DM with warm-glow preferences is a consequentialist whose preferences depend on her own outcome, the charitable recipient's outcomes, and the decision that is implemented. In the Andreoni-framework it is not possible that a DM decides to contribute g but then her decision is not carried out. Assume this can happen and that in this case she contributes some constant κ (think of it as zero). All others contribute G_{-DM} to the public good in every state of the world. Then we can write the decision-maker's expected utility as

$$E[u(x_1, g, G)] = \pi u(x_1^C, g^C, G^C) + (1-\pi)u(x_1^N, g^N, G^N)$$

The indirect objective function is then:

$$V(d) = \pi u(\omega - d, d, G_{-DM} + d) + (1-\pi)u(\omega - \kappa, \kappa, G_{-DM} + \kappa)$$

Note that the objective function is affine in $u(\omega - d, d, G_{-DM} + d)$. Thus d^* does not depend on π.

Fact. ***(consequentialist EU maximizers)*** *For a consequentialist DM who satisfies the assumptions of expected utility theory the optimal d^* does not depend on π.*

Proof. A consequentialist who satisfies the axioms of expected utility theory faces the constrained optimization problem $V(d) = \pi u(\omega - d, d) + (1-\pi)u(\omega - \kappa, \kappa)$ subject to $0 \leq d \leq \omega$. Note that the indirect objective function V function is affine in $u(\omega - d, d)$. Thus the optimal d does not depend on π.[1] □

Another way to see this is to look at the first-order condition, which can be written as $\frac{u_1(\omega-d,d)}{u_2(\omega-d,d)} = 1$, so the marginal rate of substition equals the marginal cost of donation of 1. Therefore under expected utility, for any consequentialist DM the amount shared does not vary in the probability.

2.5 Purely deontological preferences

We say that the DM has deontological motivations if her Bernoulli utility does not only depend on the consequences but also on the decision itself. The DM cares about her decision even if is without any monetary or non-monetary consequence for herself and others. Thus even if the decision is never implemented, and which decision she took is never learned by anyone, and thus no one's opinion about the DM as a consequence of her decision changes, she still cares about the decision.

Theorem 2. ***(Deontological preferences)*** *For purely deontological preferences the optimal decision d^* is constant in the probability π.*

It is a natural question to ask if deontological moral philosophy even applies to situations under (objective) uncertainty. A first response is that the natural world is always uncertain, so any moral philosophy that aims to provide guidance to people in the natural world presumably must apply to decision-making under uncertainty. Let us illustrate this point with what Kant thinks about uncertainty in the famous ax-murderer example: "Es ist doch möglich, daß, nachdem du dem Mörder auf die Frage, ob der von ihm Angefeindete zu Hause sei, ehrlicherweise mit ja geantwortet hast, dieser doch unbemerkt ausgegangen ist und so dem Mörder nicht in den Wurf gekommen, die That also nicht geschehen wäre" (Kant 1797). He states that there is always some uncertainty about the consequences of saying the truth to the ax-murderer, so therefore, one should do one's duty to say the truth regardless of what happens to the ax-murderer or the victim (i.e. the victim happens to have left the house unnoticed by you).

This is because in these lexicographic preferences, a person is either pure deontological or pure consequentialist in comparing possible decisions. Formally, there is no trade-off.

[1] Note that there could be more than one optimal d, but then the solution set does not depend on π.

A lexicographic deontologist maximizes $u(d)$ first, then there is a compact set where she maximizes $v(x)$ next. Our theorem applies to either the pure consequentialist portion $v(x)$ or the deontological portion $u(d)$.

2.6 Consequentialist-deontological preferences

Theorems 1 and the fact show that neither consequentialist nor purely deontological preferences predict behavioral changes as the probability of being consequential changes. Now we give a simple example of consequentialist-deontological preferences where the optimal decision changes as the probability of being consequentialist changes. To that end, we consider an additive utility functions that depends on the decision d and only on one consequence, the payoff for the DM herself:

Example 4. $u = u(x_1, d) = x_1 + b(d)$, where $b_1 > 0$ and $b_{11} < 0$.

Then $V(d) = \pi(\omega - d) + (1 - \pi)(\omega - \kappa) + b(d)$ is strictly concave in d. The first-order condition is $b_1(d) = \pi$ and thus for an interior solution $\frac{\partial d^*}{\partial \pi} = \frac{1}{b_{11}(d)} < 0$.

Thus in the above example d^* is decreasing in the probability of being consequential. This result means that the lower the probability that the DM's decision is implemented, the more she donates. At first glance, this result may seem somewhat counter-intuitive, but it is consistent with the intuition that wealthy people vote for generous redistribution when their probability of being pivotal is low, while in economic life, where their decision gets implemented for sure, they may donate only a little. *In our setup, the benefit of altruism is always there, but the costs are only incurred with probability* $1-\pi$*.* So, the lower the probability the decision is executed, the lower the cost of making the decision, and thus we should expect to see *more* altruism.[2]

For a slightly more general model: let $u(x_1, d) = f(x_1) + b(d)$. Then, $U(x_1, d) = \pi(f(x_1^C) + b(d)) + (1 - \pi)(f(x_1^N) + b(d))$ and $V(d) = \pi f(\omega - d) + (1 - \pi) f(\omega - \kappa) + b(d)$. The first order condition is: $\frac{\partial V(d)}{\partial d} = -\pi f_1(\omega - d) + b_1(d) = 0$. For d^* to be a maximum, then $\frac{\partial^2 V(d)}{\partial d^2} = \pi f_{11}(\omega - d) + b_{11}(d) < 0$. By the implicit function theorem, $\frac{\partial d^*}{\partial \pi} = \frac{f_1(\omega - d^*)}{\pi f_{11}(\omega - d^*) + b_{11}(d^*)} < 0$, since utility is increasing in own outcomes and the denominator, which is the second derivative of the indirect objective function, is negative. Note that the recipient's payoff is a function of the DM's payoffs, but so long as other-regarding concerns are concave then the sum of utility from own payoffs and utility from others' payoffs is still concave and the above result holds. Decisions do not have to be continuous for this result to obtain. If decisions are

[2] Utility in money does not have to be linear for this result to obtain.

discrete, then the behavior of mixed consequentialist-deontological person is jumpy, i.e. it weakly increases as their decision becomes less consequential. Note that if the consequentialist and deontological choice is the same, then the choice is still invariant to the implementation probability: $f_1(\omega - d) = b_1(d) = 0$, then $\frac{\partial d^*}{\partial \pi} = 0$.

Example 5. (Impure altruism and deontological motivations) In example 3 we showed that Andreoni-preferences for warm-glow are purely consequential. Now let us extend the Andreoni-preferences to allow for deontological motivations and assume a utility function of the form $u = u(x_1, g, G, d)$. Thus the consequences x are now the triple $x = (x_1, g, G)$. The DM decides how much to contribute, that is chooses a d (where d is affordable $0 \leq d \leq \omega$). In the consequential state the decision gets implemented and thus $g^C = d$, wheras in the nonconsequential state $g^N = \kappa$.

$$E[u(x_1, g, G, d)] = \pi u(x_1^C, g^C, G^C, d) + (1 - \pi) u(x_1^N, g^N, G^N, d)$$

The indirect objective function is then:

$$V(d) = \pi u(\omega - d, d, G_{-DM} + d, d) + (1 - \pi) u(\omega - \kappa, \kappa, G_{-DM} + \kappa, d)$$

Now d^* can vary in π.

Non-Additive Utility Now for more complicated utility functions, non-additive or non-globally convex ones, it is possible to generate examples where $\frac{\partial d^*}{\partial \pi} = \frac{1}{b_{11}(d)} > 0$. Suppose the DM has preferences represented by $u = u(x_1, d)$. Assume that the first derivatives are positive (monotonicity), and that $u_{11} < 0$ and $u_{22} < 0$ (risk-aversion). Then the DM maximizes $V(d) = \pi u(\omega - d, d) + (1 - \pi) u(\omega - \kappa, d)$. The first order condition is $-\pi u_1(\omega - d, d) + \pi u_2(\omega - d, d) + (1 - \pi) u_2(\omega - \kappa, d) = 0$. By the implicit function theorem, and simplifying using the first order condition gives

$$\frac{\partial d^*}{\partial \pi} = \frac{1}{\pi^2}[-2u_{12}(\omega - d, d) + u_{11}(\omega - d, d) + u_{22}(\omega - d, d) + \frac{1 - \pi}{\pi} u_{22}(\omega - \kappa, d)]^{-1} u_2(\omega - \kappa, d)$$

So for sufficiently negative $u_{12}(\omega - d, d)$ we can get $\frac{\partial d^*}{\partial \pi} > 0$. Utility functions that are not globally convex can lead to local maxima that, when the decision is less consequential, can lead to jumps to maxima involving lower d.

Example 6. (Bliss Point) $u(x_1, x_2, d) = (1 - \mu) \left(- (1 - \lambda) (\omega - x_1)^2 - \lambda (\omega - x_2)^2 \right)$

$-\mu(\delta-d)^2$, where $0 \le \delta \le \omega$ and $0 \le \mu, \lambda \le 1$. In our thought experiment $V(d) = \pi(1-\mu)\left(-(1-\lambda)d^2 - \lambda(\omega-d)^2\right) + (1-\pi)(1-\mu)\left(-(1-\lambda)\kappa^2 - \lambda(\omega-\kappa)^2\right) - \mu(\delta-d)^2$. For a DM who is pure consequentialist ($\mu = 0$), the function obtain its global maxima at a blisspoint: $d^* = \lambda\omega \equiv d_c^*$. For a DM who is pure deontological ($\mu = 1$) there exist a blisspoint $d^* = \delta \equiv d_d^*$. We now look at a person with mixed preferences. There is a unique critical point where the function obtain its global maxima: $d^* = \frac{\pi(1-\mu)}{\pi(1-\mu)+\mu}\lambda\omega + \frac{\mu}{\pi(1-\mu)+\mu}\delta = \frac{\pi(1-\mu)}{\pi(1-\mu)+\mu}d_c^* + \frac{\mu}{\pi(1-\mu)+\mu}d_d^*$. As you can see, d^* is a weighted mean of the two bliss points and if $d_c^* \neq d_d^*$ it depends on π: $\frac{\partial d^*}{\partial \pi} = \frac{(1-\mu)\mu\left(d_c^*-d_d^*\right)}{(\pi(1-\mu)+\mu)^2}$. The relation between d_c^* and d_d^* determine the sign of this expression. If $d_c^* > d_d^*$, so the bliss point for a consequentialist is to the right of the bliss point for a deontologist, then as the probability of being consequential increases, the d^* increases as well. Such a situation can arise, for example, if social audience concerns are strong and the duty to donate to others is weak perhaps because the duty to one's own family is strong. We might expect this situation to occur in developing countries or in coworker experiments, where the duty to donate to colleagues is weaker than the duty to provide money for one's children. Bell-shaped utility functions commonly used in estimates of policy choices by politicians lead to such a scenario.

This formulation also addresses the possibility of competing duties, such as the duty to others versus the duty to oneself and one's family. Reducing the probability a decision is implemented should lead to decisions that align more with the direction where one feels the greatest duty. The direction of the decision changes gives insight into these competing duties, and the location of the optimand for one's greatest duty.

2.7 Other explanations

2.7.1 Ex-ante fairness

A potential confound to our explanation of deontological motivations is that people could have preferences over the lotteries themselves if they view them as procedures, rather than if their preferences are fundamentally driven by the prizes (consequences or the decision). Formally, this is a violation of first-order stochastic dominance, and as such might be viewed implausible, but a famous example articulated by Machina shows how this might not be as implausible as one might first think. Machina gives the example of a mother:

> "Mom has a single indivisible item -a "treat"- which she can give to either daughter Abigail or son Benjamin. Assume that she is indifferent between Abigail getting the treat and Benjamin getting the treat, and strongly prefers either of

these outcomes to the case where neither child gets it. However, in a violation of the precepts of expected utility theory, Mom strictly prefers a coin flip over either of these sure outcomes, and in particular, strictly prefers $\frac{1}{2}, \frac{1}{2}$ to any other pair of probabilities." (Machina 1989)

In our experimental setup for example a subject might target the expected value of the recipient and thus vary the decision in the probability:

Example 7. Targeting the recipient's expected value. Consider the following preferences $U(x_1, x_2) = E[x_1] + a(E[x_2]) = \pi x_1^C + (1-\pi)x_1^N + a\left(\pi x_2^C + (1-\pi)x_2^N\right)$. Let a be a function that captures altruism and let it be strictly increasing and strictly concave. Note that this objective function is not linear in the probabilities. The indirect objective function is $V(d) = \pi(\omega - d) + (1-\pi)(\omega - \kappa) + a(\pi d + (1-\pi)\kappa)$. The first-order condition is $a_1(\pi d + (1-\pi)\kappa) = 1$. By the implicit function theorem, $\frac{\partial d^*}{\partial \pi} = \frac{\kappa - d^*}{\pi}$. Thus the optimal decision changes in the probability. In two special cases, it is easy to determine the sign of the derivative, even if d^* itself is not (yet) known: if $\kappa = 0$, then $\frac{\partial d^*}{\partial \pi} \leqq 0$, and if $\kappa = \omega$, then $\frac{\partial d^*}{\partial \pi} \geqq 0$.

Let us look at a more general case: $U = f(E[u(x_1)], E[\widetilde{u}(x_2)])$, where f is $f_1, f_2 > 0$ (strictly increasing), $f_{12}f_1f_2 - f_{11}f_2^2 - f_{22}f_1^2 > 0$ (strictly quasi-concave), $f_{12}f_2 - f_{22}f_1 > 0$, $f_{12}f_1 - f_{11}f_2 > 0$ (normal in both arguments) and $u, \widetilde{u}$ is $u_1, \widetilde{u}_1 > 0$ (strictly increasing) and $u_{11}, \widetilde{u}_{11} < 0$ (strictly concave). Then, the indirect objective function is

$$V(d) = f\left(\pi u(\omega - d) + (1-\pi)u(\omega-\kappa), \pi\widetilde{u}(d) + (1-\pi)\widetilde{u}(\kappa)\right)$$

Note that $V(d)$ is globally strongly concave:

$$\frac{1}{\pi}\frac{\partial^2 V(d)}{(\partial d)^2} = -\left(2f_{12}f_1f_2 - f_{11}f_2^2 - f_{22}f_1^2\right)\frac{1}{f_2^2}\pi u_1^2(\omega - d) + f_1 u_{11}(\omega - d) + f_2\widetilde{u}_{11}(d) < 0$$

So, there exist a unique solution. The First-order condition for this problem is $\frac{\widetilde{u}_1(d)}{u_1(\omega-d)} - \frac{f_1}{f_2} = 0 \equiv F$. The FOC defines d^* implicitly as a function of π. By the implicit function theorem $\frac{\partial d^*}{\partial \pi} = -\frac{\frac{\partial F(d^*,\pi)}{\partial \pi}}{\frac{\partial F(d^*,\pi)}{\partial d^*}}$. As $\frac{\partial F(d^*,\pi)}{\partial d^*}$ has sign of $\frac{\partial^2 V(d)}{(\partial d)^2} < 0$: $sgn\left(\frac{\partial d^*}{\partial \pi}\right) = sgn\left(\frac{\partial F(d^*,\pi)}{\partial \pi}\right)$. It can be shown that:

$$\begin{aligned}\frac{\partial F(d^*,\pi)}{\partial \pi} &= \frac{\widetilde{u}_1(d^*)}{f_1}(f_{12}f_1 - f_{11}f_2)\left[u(\omega - d^*) - u(\omega - \kappa)\right] \\ &+ \frac{u_1(\omega - d^*)}{f_2}(f_{12}f_2 - f_{22}f_1)\left[\widetilde{u}(\kappa) - \widetilde{u}(d^*)\right]\end{aligned}$$

So the sign of $\frac{\partial d^*}{\partial \pi}(\pi)$ depends on the difference between $d^*(\pi)$ and κ:

For $d^*(\pi) = \kappa$: $\frac{\partial F(d^*,\pi)}{\partial \pi} = 0$ thus $\frac{\partial d^*}{\partial \pi}(\pi) = 0$

For $d^*(\pi) < \kappa$: $\frac{\partial F(d^*,\pi)}{\partial \pi} > 0$ thus $\frac{\partial d^*}{\partial \pi}(\pi) > 0$

For $d^*(\pi) > \kappa$: $\frac{\partial F(d^*,\pi)}{\partial \pi} < 0$ thus $\frac{\partial d^*}{\partial \pi}(\pi) < 0$

Now if $\kappa = 0$, then $\frac{\partial d^*}{\partial \pi} \leqslant 0$, while for $\kappa = \omega$ $\frac{\partial d^*}{\partial \pi} \geqslant 0$. Thus experimentally, by varying κ we can test whether people have such ex-ante considerations.

2.7.2 Cognition costs

Another possible explanation for invariance in the probability might be cognition costs. Cognition costs are a consequence, but unlike the other consequences they are not captured in our consequentialist framework, since they are incurred during the decision and are a consequence that even arises if the nonconsequential state realizes. To fix ideas, consider the following model: $u = u(x_1, x_2, \gamma)$, where $u_1, u_2 > 0$, $u_\gamma < 0$ and $\gamma \geq 0$. In addition, let's assume that utility is continious. The DM can compute the optimal decision, but to do so, she incurs a cognition cost $\gamma > 0$, or otherwise she can make a heuristic (fixed) choice $\bar{d}$ for which (normalized) costs are 0. We have no model of what the heuristic choice is, and in principle it could be anything, but recent experimental work argues that the heuristic choice tends to be a cooperative or fair one (Rand et al. 2012) so the reader might wish to think for example of $\bar{d} = \frac{\omega}{2}$. In any case expected utility from the heuristic choice is $V(\bar{d}) = \pi u(\omega - \bar{d}, \bar{d}, 0) + (1-\pi)u(\omega - \kappa, \kappa, 0)$. By constrast for a non-heuristic choice, $V(d) = \pi u(\omega - d, d, \gamma) + (1-\pi)u(\omega - \kappa, \kappa, \gamma)$. Define $\check{d} \equiv argmax V(d)$. Obviously, $\check{d}$ does not vary in π. The DM will choose to act heuristically iff $V(\check{d}) < V(\bar{d})$ or

$$F(\pi) \equiv V(\check{d}) - V(\bar{d}) = \pi \left(u(\omega - \check{d}, \check{d}, \gamma) - u(\omega - \bar{d}, \bar{d}, 0) \right) \\ + (1-\pi) \left(u(\omega - \kappa, \kappa, \gamma) - u(\omega - \kappa, \kappa, 0) \right) < 0$$

We can differ two cases:

i) If $u(\omega - \check{d}, \check{d}, \gamma) - u(\omega - \bar{d}, \bar{d}, 0) < 0$, $F(\pi)$ is always negative, so the person uses the heurisic choice, independent of π.

ii) In the other case, $u(\omega - \check{d}, \check{d}, \gamma) - u(\omega - \bar{d}, \bar{d}, 0) > 0$, there exist a unique $\widetilde{\pi}$ with $0 < \widetilde{\pi} < 1$ such that $F(\widetilde{\pi}) = 0$, the person switch from heuristic to non heuristic. This comes from the fact that in this case$F(\pi)$ is strictly monotone in π, $F(0) < 0$ and $F(1) > 0$, so for probabilities of being consequential close to 1 computing is better, and for probabilities close to zero, the heuristic is better. Since $\check{d} \neq \bar{d}$, this means that such cognition costs predict that even a consequentialist DM will not be invariant to the probability. For the rest of this section, we will focus on this case.

Now suppose we vary the cognition cost, that is, we do an exercise in comparative statics and investigate how $\widetilde{\pi}$ varies in γ, and note that

$$\frac{\partial \widetilde{\pi}}{\partial \gamma} = \frac{-\widetilde{\pi} u_3(\omega - \check{d}, \check{d}, \gamma) - (1 - \widetilde{\pi})\, u_3(\omega - \kappa, \kappa, \gamma)}{u(\omega - \check{d}, \check{d}, \gamma) - u(\omega - \bar{d}, \bar{d}, 0) + u(\omega - \kappa, \kappa, 0) - u(\omega - \kappa, \kappa, \gamma)} > 0,$$

that is, the higher the cognition costs, the higher the thresold for probability being consequential such that computation is the better choice. Obviously there are some very low γ and some very high γ such that locally, $\widetilde{\pi}$ is a constant function of γ, but there, the above assumptions are violated. The figure shows when, as a function of a probability, someone would incur a given cognition cost. So if we could experimentally vary not only probability but also cognition costs, and then observe it, the cognition cost story predicts the pattern shown in the figure.

In summary, variation in the decision d with respect to π is consistent with decision-makers switching to a heuristic $\bar{d}$, which may be higher or lower than the preferred choice $\check{d}$, leading to the inability to infer consequentialist-deontological preferences. If decision-makers have different γ or different $\bar{d}$, then we might observe a smooth $\frac{\delta d}{\delta \pi}$. A cognition costs model, however, would predict that time spent on the survey also changes as d changes with π. Our research design provides a second test of the cognition costs model. By increasing the cognition cost of making a decision, we should observe $\frac{\delta d}{\delta \pi}$ to be larger for larger values of π, but smaller for smaller values of π. Thus, we should observe an S-shape curve in decisions with respect to π, the higher are cognition costs. This formal modeling and experimental test of cognition costs seems to be rare in the literature, for a previous example, albeit one that does not have the decision-maker solve the metaproblem optimally, see Wilcox (1993).

3 Experimental Methodology

We implement two experimental designs. Our shredding experiment involves undergraduates making a decision to donate to Doctors Without Borders. Since the DM may care about what the experimenter thinks of the decision and the expected benefit of the decision per se remains even in the non-consequential state, we shred decisions that are not implemented, thus, no

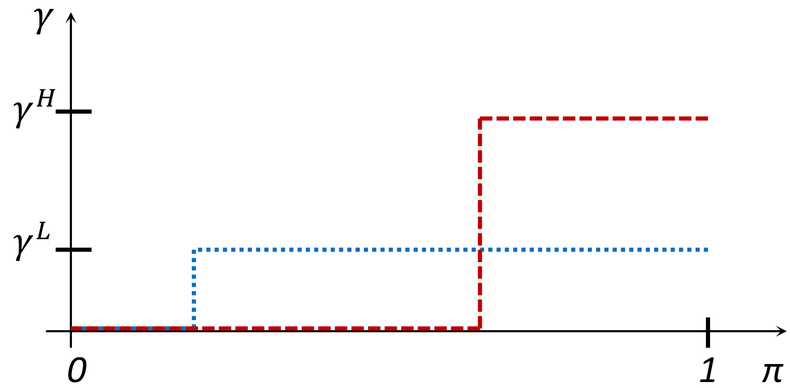


Figure 3:

one ever knows the decision in the non-consequential state. Our Amazon Mechanical Turk (MTurk or AMT) experiment involves anonymous data entry workers splitting a bonus with the Red Cross.

Participants in the shredding experiment first see a demonstration of a public randomization device (Wheel of Fortune) and paper shredder; the shredding bin is opened to publicly verify that materials will truly be destroyed. Participants then solve three IQ tasks. If at least one answer is correct, they proceed to the donation decision. Subjects are randomly assigned to low (3/16) or high probability (15/16) π of implementation and to 0 or maximum κ donation in the non-consequential state. The decision is written on a piece of paper and placed in an envelope. The donation is made out of 20Chf. After the Wheel of Fortune is spun, envelopes that are to be destroyed are collected and shredded. The remainder are opened and participants are paid. Undergraduates were recruited in the lab and classroom.

Participants in the MTurk experiment were recruited as workers in Amazon Mechanical Turk, a labor market intermediary (LMI). The LMI can be used to implement anything from a natural field experiment to a laboratory experiment (Harrison and List 2004). Workers come to the marketplace naturally and are unaware they are in an experiment at the time of arrival, and this lack of awareness alleviates Hawthorne effects (Orne 1962; Rosenthal 1966). Through an interface provided by the LMI, registered users perform tasks posted by buyers for money. The tasks are generally simple for humans to do yet difficult for computers. Common tasks include captioning photographs, extracting data from scanned documents, and transcribing audio clips.

To make the experiment appear more natural to workers and to lock them in to prevent selective attrition, we first asked them to transcribe paragraphs from a Dutch translation of Adam Smith's *The Wealth of Nations*. This task is sufficiently tedious that no one is likely to do it "for fun", and it is sufficiently simple that all participants can do the task. The source text was machine-translated to prevent subjects from finding the text elsewhere on the Internet. In all treatment conditions, workers face an identical "lock-in" task in order to minimize differential attrition when the treatment is revealed (Chen 2012; Chen and Horton 2009). The lock-in task was the data entry of three paragraphs. A paragraph takes about 100 seconds to enter so a payment of 10 cents per paragraph is equivalent to $28.80 per 8-hour day. The current federal minimum wage in the Unites States is $58/day.

After the lock-in task, subjects have an opportunity to split their bonus with the charitable recipient, the Red Cross. Workers then provided their gender, age, country of residence, religion, and how often they attend religious services. After work was completed, according to the original expiration date listed on the LMI, bonuses were calculated and workers were notified of their earnings.

The empirical specification examines the effect of treatment on donation:

$$Donation_i = \beta_0 + \beta_1 Treatment_i + \beta_2 X_i + \varepsilon_i \tag{1}$$

$Treatment_i$ represents the treatment group for individual i (sometimes represented as π, the probability a decision is consequential) and X_i represents individual demographic characteristics. We display the raw data means, distributions, results from the Wilcoxon-Mann-Whitney test for differences in distributions of donations, and ordinary least squares regressions. When we included covariates, country of origin was coded as United States and India with the omitted category as other, religion was coded as Christian, Hindu, and Atheist with the omitted category as other, religious services attendance was coded as never, once a year, once a month, once a week, or multiple times a week.[3]

We next estimate for each individual how sensitive their decision d is to π as predicted from their demographic characteristics. In essence, we construct synthetic cohorts to emulate a within-subject design. Formally, we estimate:

$$Donation_i = \beta_0 ProbabilityConsequential_i + \beta_1 \mathbf{X}_i ProbabilityConsequential_i + \alpha \mathbf{X}_i + \varepsilon_i$$

We interpret the change in d to π as measuring the mixed consequentialist-deontological motives. Intuitively, if $\mathbf{X}_i$ were country fixed effects, this would be like computing country-

[3]Some regressions also code for levels of respect for their parents, police, and their boss, respectively: not at all, not much, some, a little, and a lot.

level averages of $\frac{\delta d}{\delta \pi}$. Each demographic variable contributes to the effect of probability of being consequential on the donation.

We then compute for each individual:

$$MixedConsequentialistDeontological_i = |\hat{\beta}_0 + \widehat{\beta}_1 \mathbf{X}_i|$$

We use all the demographic characteristics to construct a mixed consequentialist-deontological score. Each demographic variable contributes to the effect of probability of being consequential on the donation. Each subject's demographic variables are then used to calculate a predicted mixed consequentialist-deontological score by taking the absolute value of the sum of the contributions of their demographic characteristics along with the constant term. We interpret the change in d to π as measuring mixed consequentialist-deontological motives. Intuitively, if $\mathbf{X}_i$ were a dummy indicator for being male, this would be like computing $\frac{\delta d}{\delta \pi}$ for the average male. Males may be less generous than females, but generosity of both males and females may decrease with π. Whether $\frac{\delta d}{\delta \pi} < 0$ in different sub-populations allows investigating the possibility that people's duties differ.

3.1 Power Calculation to Determine Optimal Treatment Ratio

In the shredding study, we conduct a power calculation to determine the optimal ratio of treatment to control subjects. Our two probabilities, 15/16 and 3/16, mean that data collection for low π is five times more expensive. Our estimand is: $\hat{\text{ß}} = E(T) - E(C) = \frac{\sum T}{n_T} - \frac{\sum C}{n_C}$. We seek to minimize $Var(\hat{\text{ß}}) = \frac{\hat{\sigma_T^2}}{n_T} - \frac{\hat{\sigma_C^2}}{n_C}$ subject to the budget constraint that $n_T c_T + n_C c_C \leq I$. The first-order conditions of the Lagrangian are $-\frac{1}{n_T^2} = -\lambda c_T$ and $-\frac{\gamma}{n_C^2} = -\lambda c_C$ where $\gamma = \frac{\hat{\sigma_C^2}}{\hat{\sigma_T^2}}$. This determines the optimal ratio of data collection to be: $\frac{n_T^2}{n_C^2} = \gamma \frac{c_C}{c_T}$. Intuitively, as the cost of data collection for treatment increases, we collect more control. As the variance of the treatment sample increases, we collect more treatment. Sample variance among low π subjects was higher in the pilot, which required a roughly 1:1 ratio of opened envelopes.

3.2 Amazon Mechanical Turk Experiment

We run two MTurk experiments, where κ, the amount given to the Red Cross when the decision is not consequential, varies within and across experiments. In both experiments, participants get randomly assigned to one of five groups that differ in π, the probability of the decision being consequential: 100%, 66%, 33%, 5%, and 1%. All subjects are in the role of dictator. In one experiment, when workers' decisions are not consequential, half the time κ is unknown to workers (they are told the computer is making a determination) and we draw κ from a uniform distribution, while in the other half, κ is 10 cents and they are told the computer will split the 50 cent bonus into 10 cents for the Red Cross and 40 cents for the worker. In the second experiment, we set κ be 50 cents and 0 cents in the different treatment arms. We present both the raw data as well as regression specifications that indicator variables for κ. Varying κ addresses the possible confounding explanation that subjects maximize ex ante utility. If workers target an expected donation, then the effect of the probability of being non-consequential should have opposite signs when the computer assigns 0 cents as donation or 50 cents as donation. When κ was unknown, we asked workers what they believed would be the amount donated if the computer made the decision.

3.3 Demographics and vignette questions

Our experiments end with a battery of demographic questions. In addition, in the shredding experiment, we ask questions in order to investigate how our revealed preference method of detecting deontological motivations relates to more traditional, vignette-based studies. The goal here is to find out to what extent both yield the same results, which would then constitute a validation of certain vignette-based approaches.

First, we ask a set of questions related to protected values. This concept stems from psychology and are "values that a community treats as absolute, not tradeable and exchangeable for otehr values," which we take from (Tanner et al., 2009). In using these protected values in experimental economics, we follow Gibson et al. (2012), whose English translation we use. Specifically, we ask about truthfulness as a protected value and have 4 statements and ask participants how much they agree with each on a Likert-type scale. We also elicit responses to two moral dilemmas, a variant of the moral trolley problem, except the number of lives saved differs between the dilemmas.

4 Experimental Results

4.1 Shredding Experiment

On average, participants donated 20% when π was high and 36% when π was low (Figure 1). The roughly 80% increase in donation is observed in both $\kappa = 0$ and $\kappa = Max$ treatments

(Figure 1). The results are significant at the 10% level with κ fixed effects (Table 1). As π changes, expected donations are not fixed; they increase when κ is high and decrease when κ is low (Figure 2). Expected giving is also not fixed. For both κ, expected giving drops by half as π goes from high to low. The statistical significance of these results are displayed in Table 1.

4.2 MTurk Experiment

Our main finding shows that the lower the probability that the decision is consequential, the more generous is the worker (Figure 3). Donations increased from 18% when $\pi = 1$ to 27% when $\pi = 0.01$. Effect of π is significant at the 5% level (Table 2). Expected donations and expected giving are not fixed, suggesting that participants are neither targeting expected receipt nor expected giving.

When we examine each κ treatment arm separately, we find a quantitatively similar 5.3% to 7.8% decrease as π goes from 0 to 1 (Table 3). The effects are not significantly different across treatment arms. Other significant predictors of donations are being Indian, who donate 8.4% less, and people under 25, who donate 5.6% less than others.

We next examine whether the distribution of donation decisions is significantly affected by π. Along most thresholds for π, the distribution of donations as π increases is significantly different (Table 4). For example, 0.05 in Column 1 means that we reject with 95% confidence the hypothesis that the distribution of decisions for people treated with $\pi = 1, 0.67, 0.33$ is the same as the distribution of decisions for people treated with $\pi = 0.05, 0.01$.

The raw data suggests that there may be many people who do not respond to treatment, always donating 0%, 50%, or 100%. One interpretation of our results could be that there are sizeable fractions of people who are pure consequentialist or pure deontological and a large fraction of people who have hybrid motivations.

The distribution of donations do not significantly vary by κ. Qualitatively similar results are found in the shredding experiment; differences by π are more significant than differences by κ (Table 5). Means are also not significantly different by κ in either experiment.

4.3 Who is Deontological?

Along all demographic groups, $\frac{\delta d}{\delta \pi} < 0$. Americans, Christians, Atheists, and those who are less likely to attend religious services are particularly likely to have steeper $\frac{\delta d}{\delta \pi}$ (Table 6). Column 3 displays a significant coefficient on the interaction with being Indian that is positive. Summing this interaction term with the level effect indicates that Indians (40% of the sample) are more pure consequentialist or deontological than others. Even when all covariates' interactions are included, Atheists appear to be the most mixed consequentialist-deontological in their motivations.

4.4 Rejection of Ex-Ante Consequentialism

Calculating expected donations, we observe strong rejection of ex ante consequentialism: the visual plot and regressions show that the expected donation changes unambiguously with π for κ separate or pooled in the shredding experiment. In the MTurk experiment, when κ is unknown, we calculate expected donations using data on perceived donation when κ is unknown. Similar variance of $E(x_2)$ with respect to π are found. Importantly in neither experiment do we see $\frac{\delta d}{\delta \pi} > 0$ when $\kappa = Max$ (Figure 1 and Table 3).

The results also do not support the hypothesis that individuals target their expected act of giving, πd. When $\kappa = 0$, expected giving is the same as expected donation, which varies with π. Regression results indicate that expected giving increases as π increases.

4.5 Rejection of Cognition Cost Explaination

Under the cognitive cost model in Section 2.6.3, individuals use heuristics and spend less time thinking when their decision is less likely to be implemented. We do find that individuals spend fewer seconds when π is low, but time spent does not predict donations (Figure 4), which is inconsistent with the cognition cost explanation explaining increased generosity and decreased time spent when π is low.

In MTurk, time spent is only affected (and reduced) by $\pi = 1$ (Figure 5); this result would appear inconsistent with cognition costs, however, $\pi = 1$ might be qualitatively different from other π. In analyses available on request, we show that females were faster, Americans were faster, Indians were slower, and people with fewer errors in data entry were slower. Donations were not associated with time spent.

Time spent is less for those with high mixed consequentialist-deontological score, $\frac{\delta d}{\delta \pi}$ (Figure 6). This is somewhat of a surprising result. One interpretation may be that people who are purely consequentialist or purely deontological take time to suppress temptation to be mixed in motives.

Importantly, those with high $\frac{\delta d}{\delta \pi}$ do not vary time spent as π changes (Figure 7). In fact, at low π, those with low $\frac{\delta d}{\delta \pi}$ spend less time than those with high $\frac{\delta d}{\delta \pi}$. In other words, those that a cognition cost explanation would resort to heuristics when π is low and think when π is high actually do the reverse, spending less time thinking about their decision as π increases, relative to those who are more pure consequentialist or deontological (Table 7).

4.6 Rejection of Diffusion of Responsibility, Loss of Control, and Disappointment Aversion

If someone feels less responsible for the outcome, they may choose to be more selfish, under the argument that the lottery chose the final outcome for the recipient. Decision-makers should become more generous with a high probability of implementation, but we find the opposite.

If individuals value authority (Fehr and Wilkening 2012), they may compensate themselves for the loss of control, which would also predict decisions to become more generous with a high probability of implementation.

4.7 Trading Off Consequentialist and Deontological Motivations

If we make functional form assumptions about consequentialist and deontological motivations, we can obtain estimates about how individuals trade off between consequentialist and deontological motivations.

4.7.1 Homo Oeconomicus and Deontological Motivations

We might write the consequentialist portion of the utility function using homo oeconomicus and the deontological portion using bliss point preferences. We would like to estimate the bliss point, δ, and the weight individuals place on the consequentialist motivations, λ.

$u(x_{DM,}, x_2, d) = \lambda(x_1) + (1-\lambda)(-(\delta-d)^2) = \lambda(\omega-d) + (1-\lambda)(-(\delta-d)^2)$

Our goal is to write the first-order condition for individuals' utility, treat the data as if they are the outcome of utility maximization, and then estimate the parameters that achieve the maximum likelihood for the observed data. In particular, the first-order conditions provide moment conditions that we try to fit. Since we are interested in the first-order condition with respect to individuals' decisions, we can focus on the decision-dependent portion of expected utility. The first-order condition is: $0 = \pi\lambda(-1) + 2(1-\lambda)(\delta-d)$. This results in a linear regression, $\pi\frac{-\lambda}{2(1-\lambda)} + \delta = d^*$.

Note that we can interpret the constant term of the linear regression as the bliss point. This is intuitive since the constant term represents the decision when $\pi = 0$. We can precisely estimate this term as 25%. Our estimate of -0.073 from Table 2 implies that $\lambda = 0.13$. This small weight is intuitive since the data reveals that many people donate more than the bliss point of 25%.

4.7.2 Fehr-Schmidt and Deontological Motivations

Next, we might write the consequentialist portion using Fehr-Schmidt preferences and the deontological portion using bliss point preferences. In principle, we would like to separately estimate the bliss point, δ, the weight individuals place on the consequentialist motivations, λ, and the inequality parameters, α and β.

$u(x_{DM,}, x_2, d) = \lambda(x_1 - \alpha max\{x_2 - x_1, 0\} - \beta max\{x_1 - x_2, 0\}) + (1 - \lambda)(-(\delta - d)^2)$.

We again seek the first-order condition for individuals' utility, treating the data as if they are the outcome of utility maximization. GMM estimation will reveal the parameters that achieve the maximum likelihood for the observed data, where we use the moment conditions derived from the first-order conditions. We can focus on the decision-dependent portion of expected utility and plug in d for x_2, x_1: $\pi\lambda(\omega - d - \alpha max\{d - (\omega - d), 0\} - \beta max\{(\omega - d) - d, 0\}) + (1 - \lambda)(-(\delta - d)^2)$.

We can rewrite this as: $\pi\lambda(\omega - d - \alpha max\{2d - \omega, 0\} - \beta max\{\omega - 2d, 0\}) + (1 - \lambda)(-(\delta - d)^2)$. This expression is quadratic in d, so the first-order condition, and hence moment conditions, will be linear in d. Thus, we will be estimating a linear regression to back out our parameters of interest. To see this, first observe that the decision-dependent portion of expected utility if $\frac{\omega}{2} > d$, is: $\pi\lambda(\omega - d - \beta(\omega - 2d)) + (1 - \lambda)(-(\delta - d)^2)$, else $\pi\lambda(\omega - d - \alpha(2d - \omega)) + (1 - \lambda)(-(\delta - d)^2)$.

The individual's first-order condition over their choice d is then given by the following expression. If $\frac{\omega}{2} > d$, then: $0 = \pi\lambda(2\beta - 1) + 2(1 - \lambda)(\delta - d)$, else $0 = \pi\lambda(-2\alpha - 1) + 2(1 - \lambda)(\delta - d)$.

Thus, our linear regression is: If $\frac{\omega}{2} > d$, then $\pi\frac{\lambda(2\beta-1)}{2(1-\lambda)} + \delta = d^*$, else $\pi\frac{\lambda(-2\alpha-1)}{2(1-\lambda)} + \delta = d^*$. This expression motivates our GMM condition:

$$E\left[\pi\left(1[\tfrac{\omega}{2} > d]\left[d - \pi\frac{\lambda(2\beta-1)}{2(1-\lambda)} - \delta\right] + 1[\tfrac{\omega}{2} \leq d]\left[d - \pi\frac{\lambda(-2\alpha-1)}{2(1-\lambda)} - \delta\right]\right)\right] = 0.$$

Equivalently, we can run a linear regression of d on $1[\frac{\omega}{2} > d]\pi$ and $1[\frac{\omega}{2} \leq d]\pi$. However, the ordinary least squares version of this regression is somewhat problematic because the decision appears on both the left-side of the equation as outcome and the right-side in the indicator function, which would drive a spurious correlation on β_2 were we to estimate $d_i = \beta_0 + \beta_1\pi_i + \beta_2 1[\frac{\omega}{2} \leq d_i]\pi_i + \varepsilon_i$. We thus need to instrument for $1[\frac{\omega}{2} \leq d_i]$ that is not directly correlated with d_i.

Estimates using two different instruments, being Indian or being under 25, result in similar point estimates (Table 8). The bliss point is to donate 25% of endowment. The first coefficient indicates that while $d < 50\%$, donation increases as π decreases. However, once $d > 50\%$, donation decreases as π decreases. This is intuitive. Since the bliss point for duty is below 50%, then for people to meet their duty as π falls, they should be moving towards 25% donation, which is less than 50%.

Our results suggest that $\frac{\lambda(2\beta-1)}{2(1-\lambda)} = -0.36$ and $\frac{\lambda(-2\alpha-1)}{2(1-\lambda)} = 1.16$. Of course, with 2 equations and 3 unknowns, we cannot identify our parameters. However, we can choose values for β and α in the range of values in Fehr and Schmidt (1999), e.g. $\beta = 0.8$ and $\alpha = 0.5$, implies $\lambda = 7.24$. Note, however, $\beta > \alpha$. These numerical examples provide a starting point for similar calculations of how individuals may trade-off consquentialist and deontological motivations.

5 Discussion and relation to previous literature

5.1 Expressive voting

Voting is a situtation that is similar but not identical to our thought experiments and actual experiments. The voting literature distinguishes between instrumental voting and expressive voting. A vote can only be instrumental insofar as the voter expects to be pivotal. All other reasons to vote and to vote for a certain option are defined as expressive in that literature. Note that this does not coincide with our definitions of consequentialist and deontological motivations. Nor does its conception of duty, which is an external one of moral obligation rather than an internal, moral one like ours. A few examples might clarify concepts: Going to the polls in order to cast a likely pivotal vote for one's preferred option would be instrumental. Voting for another option when one is not pivotal would be expressive. Going to vote even though one has no chance to affect the outcome would sometimes be called "duty" in that literature, but note that this likely refers to the social expectation and social sanctions not an interior conception of duty. Despite these conceptual differences the situation of voting can help inform our paper and is related to it. Our experiment creates a state of the world where a decision has no consequence whatsoever, and no one except the subject herself, not even the experimenter, knows the decision. In voting the decision to go vote obviously has the consequence that family, neighbors and friends may observe that. What has only a tiny consequence and thus comes closest to our clean experimental design is casting a particular non-pivotal vote for one candidate rather than another. There the only consequence is that in the published tally of votes even non-pivotal votes are of impact (for example it matters not only which candidates wins a U.S. presidential election but also by which margin for reasons of perceived legitimacy.).

5.2 Repugnance and refusal to engage in trade-offs

There is a growing empirical interest in sacred values and non-consequentialist motivations. Motivations such as repugnance are invoked to explain the lack of certain market transactions (Roth 2007) and the reluctance to tax immutable characteristics (Mankiw and Weinzierl 2010). Experimentalists have begun to document behaviors resembling sacred values. An experiment asking subjects to imagine they are a CEO finds preferences for procedural fairness (Gibson et al. 2012; Brock et al. 2011). Another set of work suggests that social audience effects underpin aversion to lying, unfair decisions, breaking of promises, and not being directly responsible for doing some misdeed (Battigalli and Dufwenberg 2007; Dana et al. 2007; Lundquist et al. 2009; Andreoni and Bernheim 2009; Coffman 2011). We argue, however, some of these behaviors previously attributed to (social) consequences, may instead be due to deontological motivations.

5.3 Intentions-based social preferences

Second, theorists have broadened what can be modeled under a consequentialist framework. For example, when agents have more control, she may identify more with the mission of the principal (Ellingsen and Johannesson 2008); a decision's consequences may signal the actor's intentions because of actions not chosen (McCabe et al. 2003) and those intentions may be rewarded or punished by others; and, actors may care about inequality, reciprocity, or simply the well-being of others stemming from the implementation of their decision (Andreoni 1990; Rabin 1993; Fehr and Schmidt 1999). We distinguish our model in that individuals might care about others or care about fulfilling a duty to help completely apart from *any* consequences of this decision.

5.4 Self-image

Third, economists are formally incorporating psychological concepts such as self-image and expressive voting. Self-image models to date, however, focus on individuals changing their self-image upon learning the consequences of their decision (Benabou and Tirole 2011). Expressive voting has a social element, because people want others to know a norm is expressed, even if they are not pivotal (Feddersen et al. 2009; Shayo and Harel 2011). Models of duty-orientation to date focus on individuals caring about their decision's consequences on public goods provision (Nyborg 2011).[4]

5.5 Organ donation

One recent study—an observational study that mimics our experimental design with changes in the probability of being consequential (Bergstrom et al. 2009)—is suggestive that deontological motivations are present outside the lab. Conditional on registration for bone marrow donation, African-Americans in the U.S. are more likely to be asked to donate than Caucasian-Americans. African-Americans are also less likely to register for bone marrow donation while Caucasian-Americans are more likely to register for bone marrow donations. This finding is similar to ours in that those whose actions are less likely to be carried out are also more likely to be generous. Our model suggests that ethnicities that have a low probability of being called to donate bone marrow are going to be more altruistic in signing up for bone marrow donation. Thus, our results may help explain existing puzzles in charitable giving. Even though our results indicated that reducing the probability of being consequential also reduced expected donations, research building on ours may devise better means of increasing donations.

5.6 Biological and internal consequences, neuroeconomics

There may be some biological mechanism that makes people behave deontologically. So in some very trivial sense, people always behave biologically not deontologically.

Further research might argue if satisfying duty activates the same pleasure centers in the brain as, say, eating ice cream does, then the best interpretation might not be duty. This paper does not take a stand. We use revealed preferences, but we are open to the possibility that neuroeconomics could make a contribution.

6 Conclusion

In recent decades behavioral economics has shown that individuals make decisions not solely based on self-interest, that is considering only consequences for oneself, but that decisions are also based on the consequences for others. This paper provides clean experimental evidence

[4] An important distinction between our paper and a large literature in psychology examining whether the probability that one's help will have an impact affects the decision to help (Batson et al. 1991; Smith et al. 1989) is that these psychological studies examine whether one's help *actually helps*, rather than whether one's help *will be carried out*, an important distinction, since in those previous experiments, the cost of the decision is experienced by subjects whether or not their decision to help actually helps.

that a focus solely on consequences is too narrow, but rather that individuals seem to also care about decisions in-and-of-themselves, independently, of consequences.

The traditional approach to measuring consequentialist vs. deontological preferences is through vignette studies (Greene and Cohen 2001; Chen 2012). We derive predictions from a simple, but general, model of social preferences to infer the presence of consequentialist or deontological preferences (or both) in an actual experiment involving costly actions. In several experiments, we investigate whether individuals care about actions per se rather than about the consequences of actions. We begin with a formal investigation of whether individual pro-sociality varies with the likelihood that their pro-social decision will actually be implemented. We show that any change is inconsistent with standard behavioral preferences. The intuition is that preferences that depend only on outcomes, whether directly *or indirectly*, would predict that decisions (altruism, truth-telling, promise-keeping) are constant in the probability because varying the probability affects equally the benefits and costs of some action.

We find that individuals share more with charitable organizations, the lower the likelihood that their sharing decision will actually be implemented. We show that this result is inconsistent with standard behavioral preferences that depend directly or indirectly on outcomes only. Our results suggest that people care about decisions in-and-of-themselves. Decisions are remembered and relevant even when they are inconsequential—in the strong sense that they not only do not affect payoffs, but moreover other agents never learn about them.

Future research should examine to what extent deontological motivations exist, how they come about, whether they differ between individuals, how to screen or select for the presence of deontological motivations, and whether deontological motivations exist outside the lab.

Our results contribute to several methodological debates in experimental economics, law and public policy, and courts. First, experimental designs that indicate to subjects only one of their decisions is payoff-relevant may cause behavioral changes because the probability their act has consequences is reduced. A seminal paper showed that utility must be approximately linear (Rabin 2003), so if expected utility also holds, when only one decision is payoff-relevant, risk aversion should not affect decision-making in these games. Our results show that even if utility is approximately linear and decision-makers employ narrow bracketing, focusing only on one decision or game at a time, changes in the probability of decisions being payoff-relevant should affect decision-making when the decision has a moral element. Our model and results contribute a novel critique of the randomized-lottery-incentive method frequently used in experimental economics. Holt (1986) is the most well-known theoretical critique of the random-lottery incentive method, which shows that if subjects understand the whole experiment as a single game and violate the independence axiom that then considering each experiment by itself does not give subjects true preferences. This critique already applies for decision-problems where there is no other player or even passive recipient and no moral dimension whatsoever, as in choices among lotteries. Starmer and Sugden (1991) experimentally test whether this potential problem identified by theory is a problem in practice, and conclude that it is not. They conclude that resesarchers using the method need not be worried about this particular problem.

Second, contingent valuation studies that query individuals' hypothetical preferences have a formal reason to differ from studies that use actual decisions with consequences. While others have pointed out puzzles in contingent valuation studies (Diamond and Hausman 1994; List and Gallet 2001; List 2001), our model provides a structure for differences between vignette studies and revealed preference.

Third, in legal settings, we are sometimes interested in the motivations of the plaintiff or the defendant (mental state), for example in copyright disputes when we care about a creator's moral rights or in equity law when we care about opportunistic behavior. A distinction between *mens rea* (intention) and *actus reus* (act) is often made. For this, a method to detect deontological motivations in other contexts seems urgent. More generally, to design optimal policies, we need to understand what people believe as moral.

The role of markets in moral behavior appears poorly understood. Market intermediation (non-consequential nature of some decisions) has been suggested to play a role in the financial crisis, though it's exact role has not been clearly delineated. Future research can also examine to what extent deontological motivations are innate/unalterable, culturally determined, or affected by particular policies. To what extent might war be about "sacred values" (Bowles and Polania-Reyes 2012)? What kinds of consequentialist policy responses or types of welfare economics might incorporate sacred values? Little also is known about the relationship between deontological motivations and economic growth, democracy, and rights (women's rights, sexual rights, religious rights, environmental rights, rights of future generations, etc.).

On a fundamental level, our results suggest that deontological motivations can explain some of the patterns previously attributed to consequences. Observations linking social preferences with outcomes, in some cases, may be due to individuals simply being hardwired to display those preferences even with no possibility of punishment or reward. Empathic concern, duty-driven, deontological decision-making may occur, regardless of the consequences for the potential beneficiaries. Behaviorally, such effects may be more prevalent than previously thought.

By showing the presence of deontological preferences, we also provide empirical foundations for broader frameworks linking identity, morals, and beliefs (Benabou and Tirole 2011), for philosophers who argue that human dignity derives from the possibility of deontological decision-making, and for theoretical frameworks investigating whether evolution selects for altruistic behavior. Alger and Weibull (2012) investigate what preferences will be selected for when preferences rather than strategies are the unit of selection and find that preferences that are a convex combination of homo oeconomicus and homo kantiensis, which is similar to our definition of purely deontological preferences, will be evolutionarily stable. Linking these theoretical predictions about the positive prevalence of deontological motivations to results about the actual prevalence of deontological motivations from studies using our revealed preference method for detecting them seems a promising research program.

References

Alexander, Larry and Michael Moore, *Stanford Encylopedia of Philosophy* Winter 2012. http://plato.stanford.edu/archives/win2012/entries/ethics-deontological/. 3, 10

Alger, Ingela and Jörgen W Weibull, "Homo Moralis," 2012. 33

Andreoni, J. and B.D. Bernheim, "Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects," *Econometrica*, 2009, *77* (5), 1607–1636. 29

Andreoni, James, "Impure altruism and donations to public goods: a theory of warm-glow giving," *Economic Journal*, 1990, *100* (401), 464–477. 2, 12, 29

Arrow, K.J., *Social choice and individual values*, Vol. 12, Yale university press, 1970. 3

Batson, C.D., J.G. Batson, J.K. Slingsby, K.L. Harrell, H.M. Peekna, and R.M. Todd, "Empathic joy and the empathy-altruism hypothesis.," *Journal of Personality and Social Psychology*, 1991, *61* (3), 413. 30

Battigalli, P. and M. Dufwenberg, "Guilt in games," *The American economic review*, 2007, *97* (2), 170–176. 29

Benabou, Roland and Jean Tirole, "Identity, morals, and taboos: Beliefs as assets," *Quarterly Journal of Economics*, 2011, *126* (2), 805–855. 29, 33

Bentham, J., *Panopticon*, T. Payne London, 1791. 3

Bergstrom, T.C., R. Garratt, and D. Sheehan-Connor, "One chance in a million: Altruism and the bone marrow registry," *American Economic Review*, 2009, *99* (4), 1309–1334. 30

Bowles, S. and S. Polania-Reyes, "Economic incentives and social preferences: substitutes or complements?," *Journal of Economic Literature*, 2012, *50* (2), 368–425. 32

Brock, J.M., A. Lange, and E.Y. Ozbay, "Dictating the risk–experimental evidence on giving in risky environments," Technical Report 2011. 29

Chen, Daniel L., "Markets and Morality: How Does Competition Affect Moral Judgment," Technical Report, Working Paper, Duke University School of Law 2012. 21, 31

_ **and John J. Horton**, "The Wages of Paycuts: Evidence from a Field Experiment," Technical Report, Working Paper 2009. 21

Coffman, L.C., "Intermediation Reduces Punishment (and Reward)," *American Economic Journal: Microeconomics*, 2011, *3* (4), 77–106. 29

D., R. Brian Sommerville Leigh E. Nystrom John M. Darley Greene Joshua and Jonathan D. Cohen, "An fMRI Investigation of Emotional Engagement in Moral Judgment," *Science*, 2001, *293* (5537), 2105–2108. 2, 31

Dana, Jason, Roberto A. Weber, and Jason X. Kuang, "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness," *Economic Theory*, 2007, *33* (1), 67–80. 29

DellaVigna, Stefano, John A. List, and Ulrike Malmendier, "Voting to Tell Others," Technical Report, Working Paper February 2013. 2

Diamond, P.A. and J.A. Hausman, "Contingent valuation: Is some number better than no number?," *The Journal of economic perspectives*, 1994, *8* (4), 45–64. 32

Ellingsen, T. and M. Johannesson, "Pride and prejudice: The human side of incentive theory," *The American Economic Review*, 2008, *98* (3), 990–1008. 29

Feddersen, Timothy, Sean Gailmard, and Alvaro Sandroni, "Moral bias in large elections: theory and experimental evidence," *American Political Science Review*, 2009, *103* (2), 175–192. 29

Fehr, E. and K.M. Schmidt, "A theory of fairness, competition, and cooperation," *Quarterly Journal of Economics*, 1999, *114* (3), 817–868. 2, 7, 11, 28, 29

Fehr, Ernst and Tom Wilkening, "The lure of authority: Motivation and incentive effects of power," 2012. 26

Friedman, Milton and L. J. Savage, "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, 1948, *56* (4), pp. 279–304. 3

Gibson, R., C. Tanner, and A. Wagner, "Preferences for truthfulness: heterogeneity among and within individuals," *American Economic Review*, 2012. 23, 29

Gneezy, Uri, "Deception: The Role of Consequences," *The American Economic Review*, 2005, *95* (1), pp. 384–394. 2

Harrison, Glenn W. and John A. List, "Field Experiments," *Journal of Economic Literature*, 2004, *42* (4), 1009–1055. 20

Holt, Charles A., "Preference Reversals and the Independence Axiom," *The American Economic Review*, 1986, *76* (3), pp. 508–515. 32

Kant, I., "The categorical imperative," 1785. 2

_ , "Über ein vermeintes Recht aus Menschenliebe zu lügen," *Berlinische Blätter*, 1797, *1* (1797), 301–314. 3, 13

Kreps, David M, *Notes on the Theory of Choice*, Westview Press Boulder, 1988. 6

Levhari, David, Jacob Paroush, and Bezalel Peleg, "Efficiency Analysis for Multivariate Distributions," *The Review of Economic Studies*, 1975, *42* (1), pp. 87–91. 7

List, John, "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards," *American Economic Review*, 2001, *91* (5), 1498–1507. 32

List, John A and Craig A Gallet, "What experimental protocol influence disparities between actual and hypothetical stated values?," *Environmental and Resource Economics*, 2001, *20* (3), 241–254. 32

Lundquist, T., T. Ellingsen, E. Gribbe, and M. Johannesson, "The aversion to lying," *Journal of Economic Behavior & Organization*, 2009, *70* (1), 81–92. 29

Machina, Mark J, ""Expected Utility" Analysis without the Independence Axiom," *Econometrica*, 1982, pp. 277–323. 6

Machina, Mark J., "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *Journal of Economic Literature*, 1989, *27* (4), pp. 1622–1668. 17

Mankiw, N.G. and M. Weinzierl, "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution," *American Economic Journal: Economic Policy*, 2010, *2* (1), 155–176. 29

McCabe, K.A., M.L. Rigdon, and V.L. Smith, "Positive reciprocity and intentions in trust games," *Journal of Economic Behavior & Organization*, 2003, *52* (2), 267–275. 29

Mikhail, J., "Universal Moral Grammar: Theory, Evidence and the Future," *Trends in Cognitive Sciences*, 2007, *11* (4), 143–152. 2

Nagel, Thomas, *The possibility of altruism*, Princeton University Press, 1978. 2

Nyborg, Karine, "I don't want to hear about it: Rational ignorance among duty-oriented consumers," *Journal of Economic Behavior & Organization*, 2011, *79* (3), 263–274. 30

Orne, M.T., "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications," *American Psychologist*, 1962, *17* (11), 776. 20

Quiggin, John, "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 1982, *3* (4), 323–343. 6

Rabin, M., "Risk aversion and expected-utility theory: A calibration theorem," *Econometrica*, 2003, *68* (5), 1281–1292. 31

Rabin, Matthew, "Incorporating fairness into game theory and economics," *American Economic Review*, 1993, pp. 1281–1302. 2, 29

Rand, David G, Joshua D Greene, and Martin A Nowak, "Spontaneous giving and calculated greed," *Nature*, 2012, *489* (7416), 427–430. 5, 18

Riker, William H. and Peter C. Ordeshook, "A Theory of the Calculus of Voting," *American Political Science Review*, 2 1968, *62*, 25–42. 2

Rosenthal, R., *Experimenter Effects in Behavioral Research*, Appleton-Century-Crofts, 1966. 20

Roth, Alvin E., "Repugnance as a Constraint on Markets," *Journal of Economic Perspectives*, 2007, *21(3)*, 37–58. 29

Shayo, Moses and Alon Harel, "Non-consequentialist voting," *Journal of Economic Behavior & Organization*, 2011. 29

Sinnott-Armstrong, Walter, *Stanford Encyclopedia of Philosophy* Winter 2012. http://plato.stanford.edu/archives/win2012/entries/consequentialism/. 3

Smith, K.D., J.P. Keating, and E. Stotland, "Altruism reconsidered: The effect of denying feedback on a victim's status to empathic witnesses.," *Journal of Personality and Social Psychology*, 1989, *57* (4), 641. 30

Starmer, Chris and Robert Sugden, "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation," *The American Economic Review*, 1991, *81* (4), pp. 971–978. 32

Tanner, Carmen, Bettina Ryf, and Martin Hanselmann, "Geschützte Werte Skala (GWS)," *Diagnostica*, 2009, *55* (3), 174–183. 23

Tversky, Amos and Daniel Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 1992, *5* (4), 297–323. 6

Wilcox, Nathaniel T, "Lottery choice: Incentives, complexity and decision time," *The Economic Journal*, 1993, *103* (421), 1397–1417. 19

Web Appendix: Experimental Instructions

Placeholder Task at Amazon Mechanical

Turk Transcribe Text

Instructions:

- After you have read the instructions, go to this site to begin work: Please Right Click Here (to open job in a new window)
- Copy text exactly as it appears in the scanned image.

Payment:

- You will receive 10 cent reward for completing the first paragraph. You can earn much more in bonus.
- When you complete the survey at the end, you will receive a completion code in order to receive payment.

You MUST keep this window open in order to enter the completion code. Bonuses will be paid after the HIT expires or after the work has been completed.

Enter completion code here: ____________________

First Page at External Host

Introduction

Task:
You will be presented with three (3) text paragraphs. Please enter the paragraphs wordfor-word in the text box below each paragraph, ignoring hyphenation. For example, if a word is split over two lines, i.e. "cup-cake," type "cupcake." Once you have transcribed as many paragraphs as you would like, hit "next," leaving the text-boxes blank - you will eventually get to the last questions.

Payment:
You must complete at least 3 paragraphs to have your work accepted. A sample paragraph is shown below. Note: Once you click "Next" you will not be able to navigate to previous pages.

De jaarlijkse arbeid van elk volk is het fonds die oorspronkelijk levert hij met alle benodigdheden en conveniencies van het leven die het jaarlijks verbruikt, en die altijd bestaan, hetzij in de onmiddellijke produceren van die arbeid, of in wat wordt gekocht met die van andere landen. Volgens dus, als deze producten, of wat is gekocht met het, draagt een grotere of kleinere verhouding tot het aantal van degenen die zijn om te consumeren, het volk zal beter of slechter geleverd met alle de benodigdheden en conveniencies waarvoor zij gelegenheid. Maar dit deel moet in elk volk worden geregeld door twee verschillende

Red Cross Experiment 1

Sixth Page at External Host – Sharing Bonus (Control)

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross.

Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Sixth Page at External Host – Sharing Bonus (1% Consequential, 99% Unknown[5] - Treatment Group 2)

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross. There is a 1% chance that your action will be carried out. Because our computer program that makes the payment to the charities account is still rough, there is a 99% chance that the program will choose an amount to donate, that may be different than yours.

Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Sixth Page at External Host – Sharing Bonus (5% Consequential, 95% Unknown - Treatment Group 3)

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross. There is a 5% chance that your action will be carried out. Because our computer program that makes the payment to the charities account is still rough, there is a 95% chance that the program will choose an amount to donate, that may be different than yours.

Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Sixth Page at External Host – Sharing Bonus (33% Consequential, 64% Unknown - Treatment Group 4)

[5] The computer's split was a uniform random number in "Unknown" treatments.

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross. There is a 33% chance that your action will be carried out. Because our computer program that makes the payment to the charities account is still rough, there is a 67% chance that the program will choose an amount to donate, that may be different than yours.
Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Sixth Page at External Host – Sharing Bonus (66% Consequential, 34% Unknown - Treatment Group 5)

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross. There is a 66% chance that your action will be carried out. Because our computer program that makes the payment to the charities account is still rough, there is a 34% chance that the program will choose an amount to donate, that may be different than yours.
Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Sixth Page at External Host – Sharing Bonus (1% Consequential, 99% 10 Cents - Treatment Group 6)

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross. There is a 1% chance that your action will be carried out. Because our computer program that makes the payment to the charities account is still rough, there is a 99% chance that the program will modify your chosen amount to 10 cents.
Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Sixth Page at External Host – Sharing Bonus (5% Consequential, 95% 10 Cents - Treatment Group 7)

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross. There is a 5% chance that your action will be carried out. Because our computer program that makes the payment to the charities account is still rough, there is a 95% chance that program will modify your chosen amount to 10 cents.
Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Sixth Page at External Host – Sharing Bonus (33% Consequential, 64% 10 Cents - Treatment Group 8)

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross. There is a 33% chance that your action will be carried out. Because our computer program that makes the payment to the charities account is still rough, there is a 67% chance that program will modify your chosen amount to 10 cents.

Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Sixth Page at External Host – Sharing Bonus (66% Consequential, 34% 10 Cents - Treatment Group 9)

You will receive a 50 cent bonus for completing this survey that you can split with the Red Cross. There is a 66% chance that your action will be carried out. Because our computer program that makes the payment to the charities account is still rough, there is a 34% chance that program will modify your chosen amount to 10 cents.

Please decide an amount between 0 and 50 to donate to the Red Cross (if you decide to keep 50 cents for yourself, type 0).

Red Cross Experiment 2

Same as Red Cross Experiment 1 except the exogenous splits were 0 and 50 cents respectively.

Web Appendix:

Shredding Experiment Instructions Donation Screen for Subject with $\pi = 3/16$ and $\kappa = 0$

Congratulations, you did well!

Therefore you have earned CHF20.00 to be divided between the charity *Doctors Without Borders* and you.

Look at the paper in the envelope on your desk. To indicate your donation decision mark the appropriate box, then seal the envelope.

After everyone has made their decision, we will spin the wheel of fortune. Depending on where it stops your decision will be implemented or not:

If the wheel stops on 1, 10 or 13:
Your envelope will be brought into the experimenter's room and **opened** there. The **donation** will be made **according to your instructions.**

If the wheel stops on 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15, 16:
Your envelope will be **shredded. No one will ever know how much you wanted to donate.** Instead, by default, the charity gets CHF0.00 and you get CHF20.00.

Please make your choice now, mark the appropriate box on the paper, place the paper in the envelope, and seal the envelope.

After you have sealed the envelope, please press OK.

OK

Sheet of paper participants fill out, put in an envelope, and seal.

Donation decision of subject number: 2

If you see the congratulations screen:

Of the CHF20 I want to donate

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

CHF to Doctors Without Borders.

If you have made too many mistakes:

Please check this box:

☐

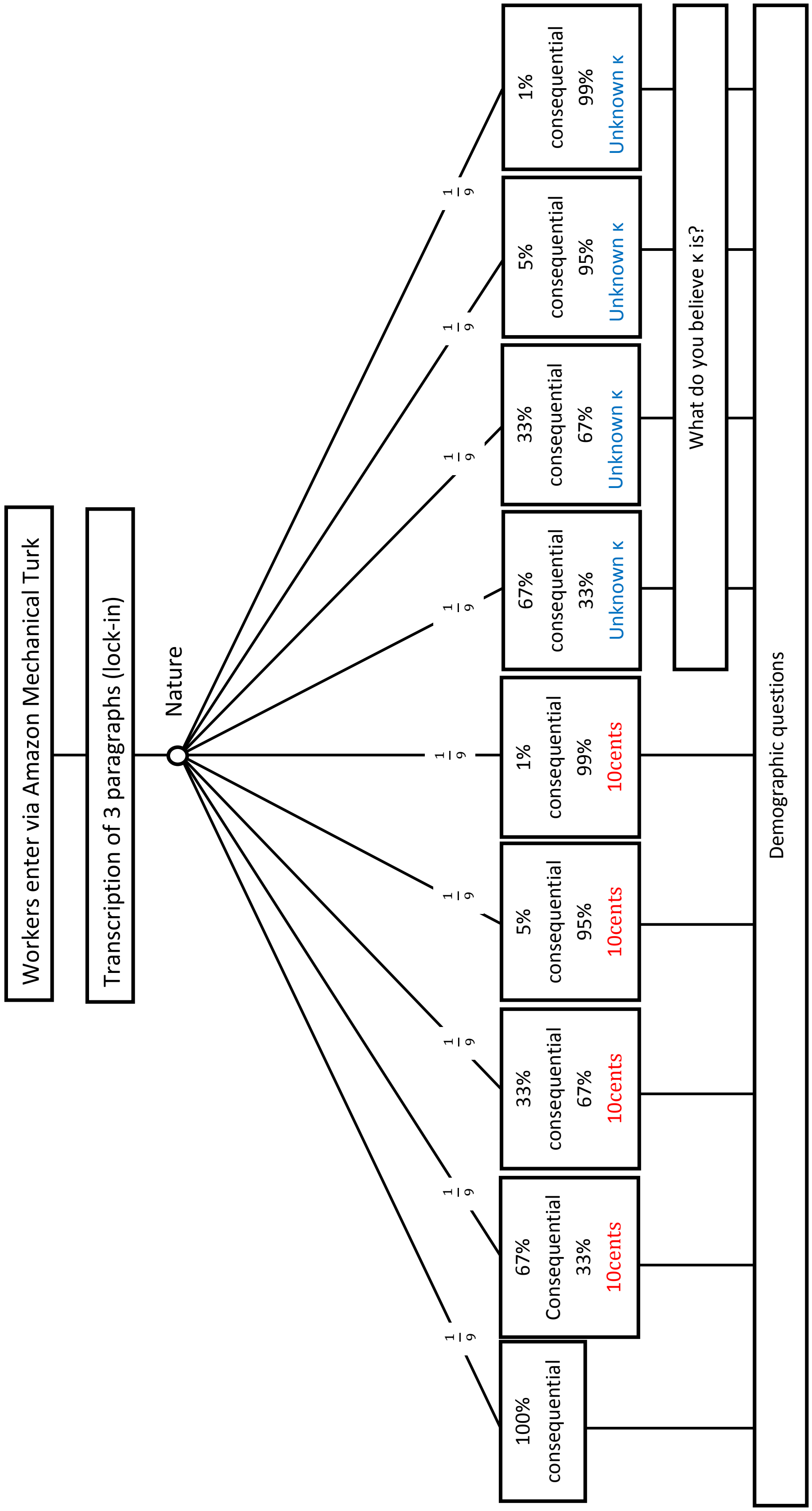
After marking exactly one box, please put this sheet in the envelope and seal it.

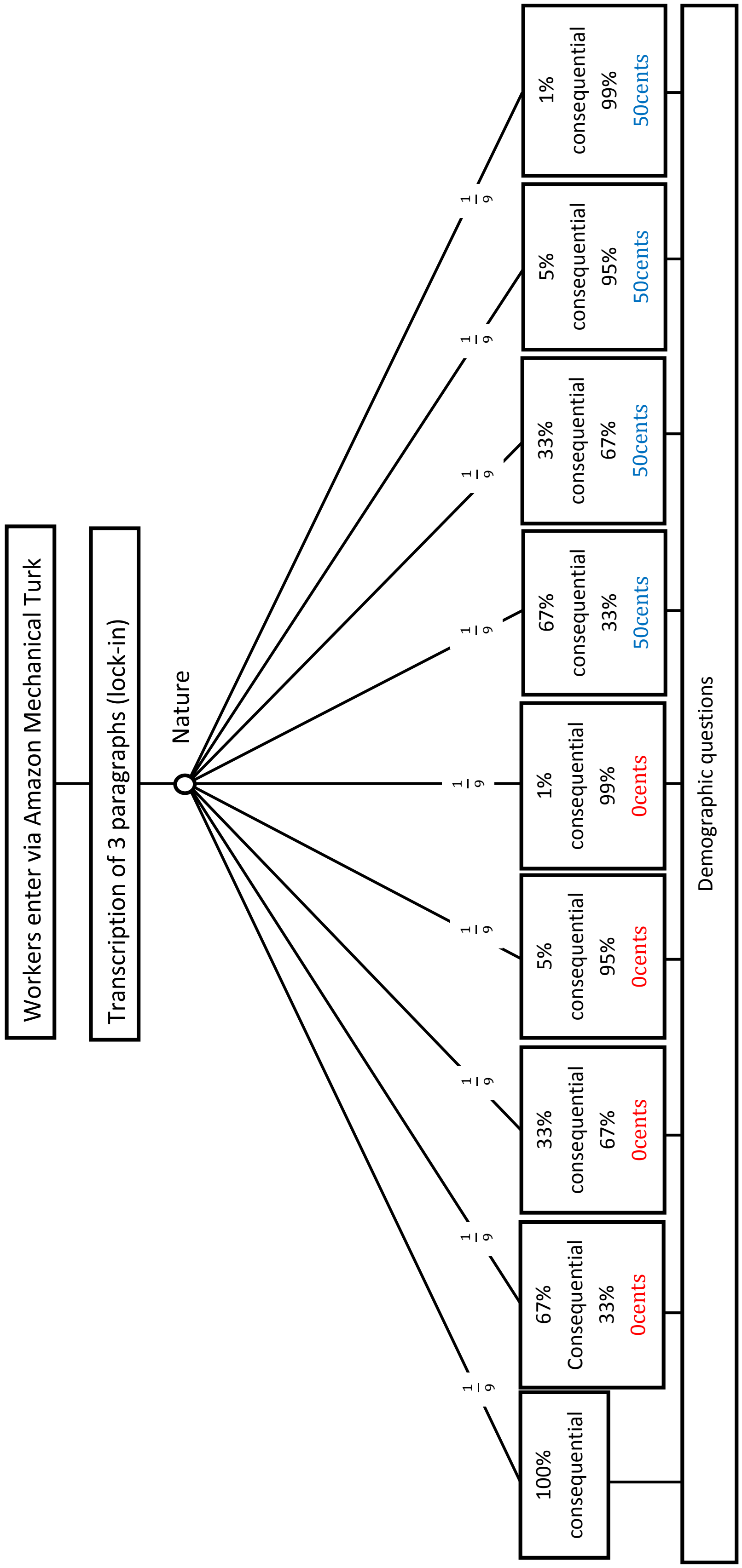
→ Then click OK on the screen so the experiment can proceed!

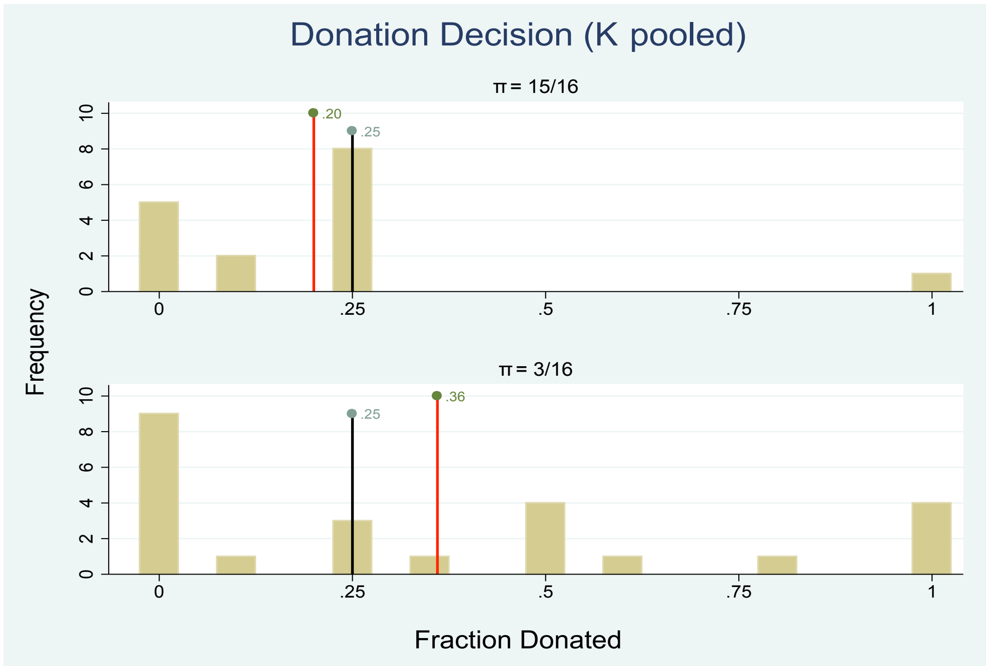
As *DM* becomes less consequential, what happens to the donation? (=What is the sign of $-\frac{\partial d}{\partial \pi}$?)

Experiment	Experimental evidence	Consequentialism	Purely Deontological	Targeting Ex Ante	Consequentialist-deontological	
					additive, each concave	general
Red Cross κ=ω=50	+	0	0	-	+	-+
Red Cross κ=0	+	0	0	+	+	-+
Red Cross κ=10	+	0	0	?	+	-+
Red Cross κ=unknown	+	0	0	?	+	-+
Co-worker κ=unknown	-	0	0	?	+	-+

Note: Self-signalling can be interpreted as consequentialist, deceptive self-signalling as consequentialist-deontological.


Workers enter via Amazon Mechanical Turk
Transcription of 3 paragraphs (lock-in)
Nature
1/9
100% consequential
67% Consequential 33% 10cents
33% consequential 67% 10cents
5% consequential 95% 10cents
1% consequential 99% 10cents
67% consequential 33% Unknown κ
33% consequential 67% Unknown κ
5% consequential 95% Unknown κ
1% consequential 99% Unknown κ
What do you believe κ is?
Demographic questions


Workers enter via Amazon Mechanical Turk
Transcription of 3 paragraphs (lock-in)
Nature
1/9
100% consequential
67% Consequential 33% 0cents
33% consequential 67% 0cents
5% consequential 95% 0cents
1% consequential 99% 0cents
67% consequential 33% 50cents
33% consequential 67% 50cents
5% consequential 95% 50cents
1% consequential 99% 50cents
Demographic questions



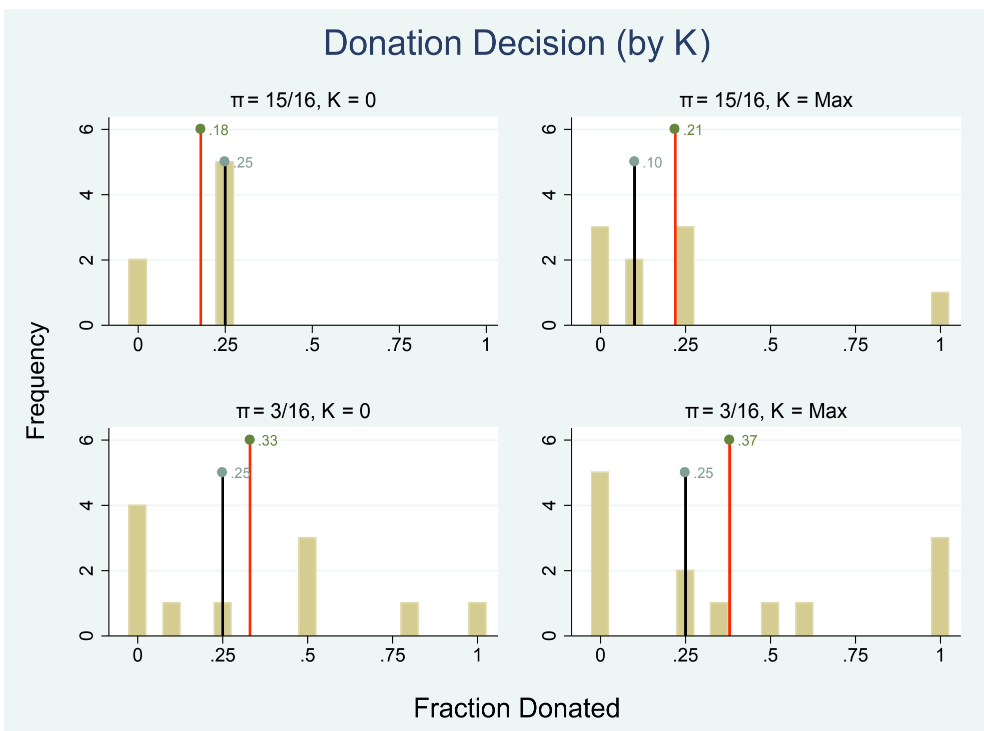
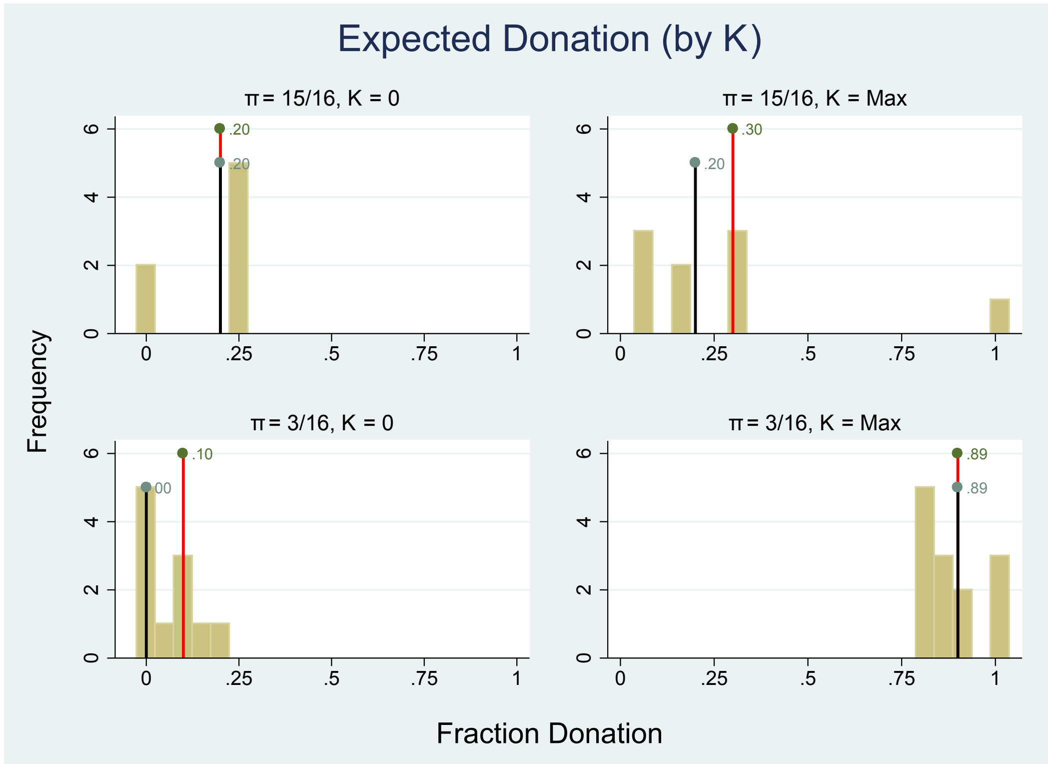


Figure 1:
Red: Mean
Black: Median



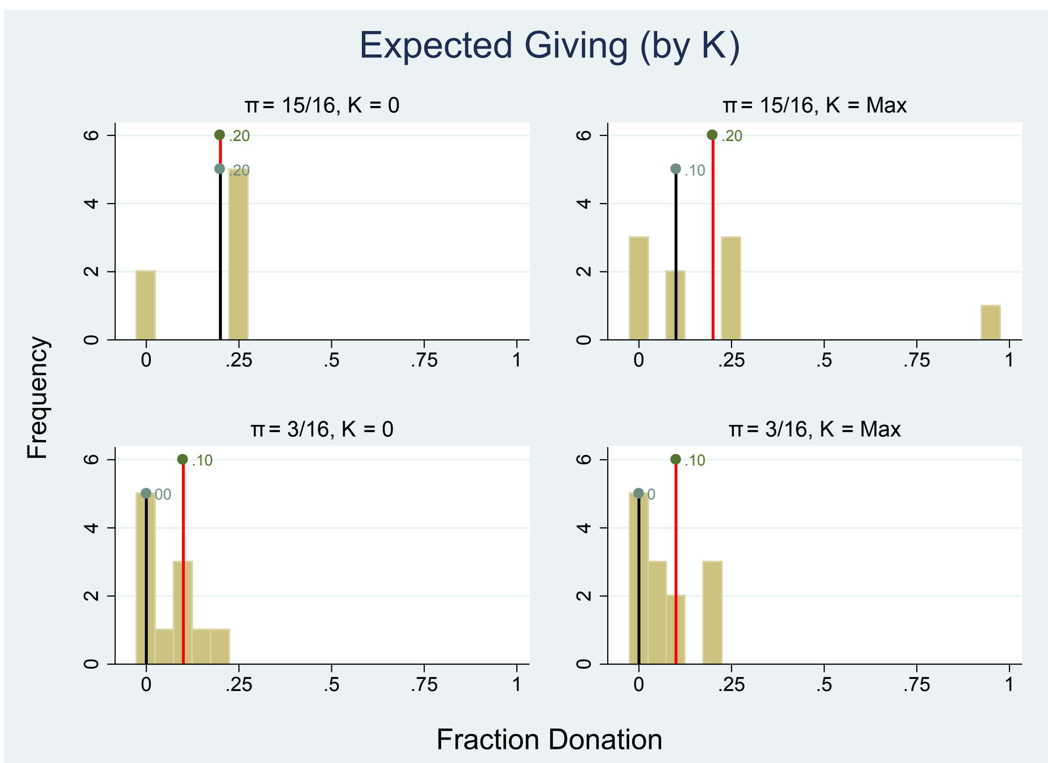


Figure 2:
Red: Mean
Black: Median

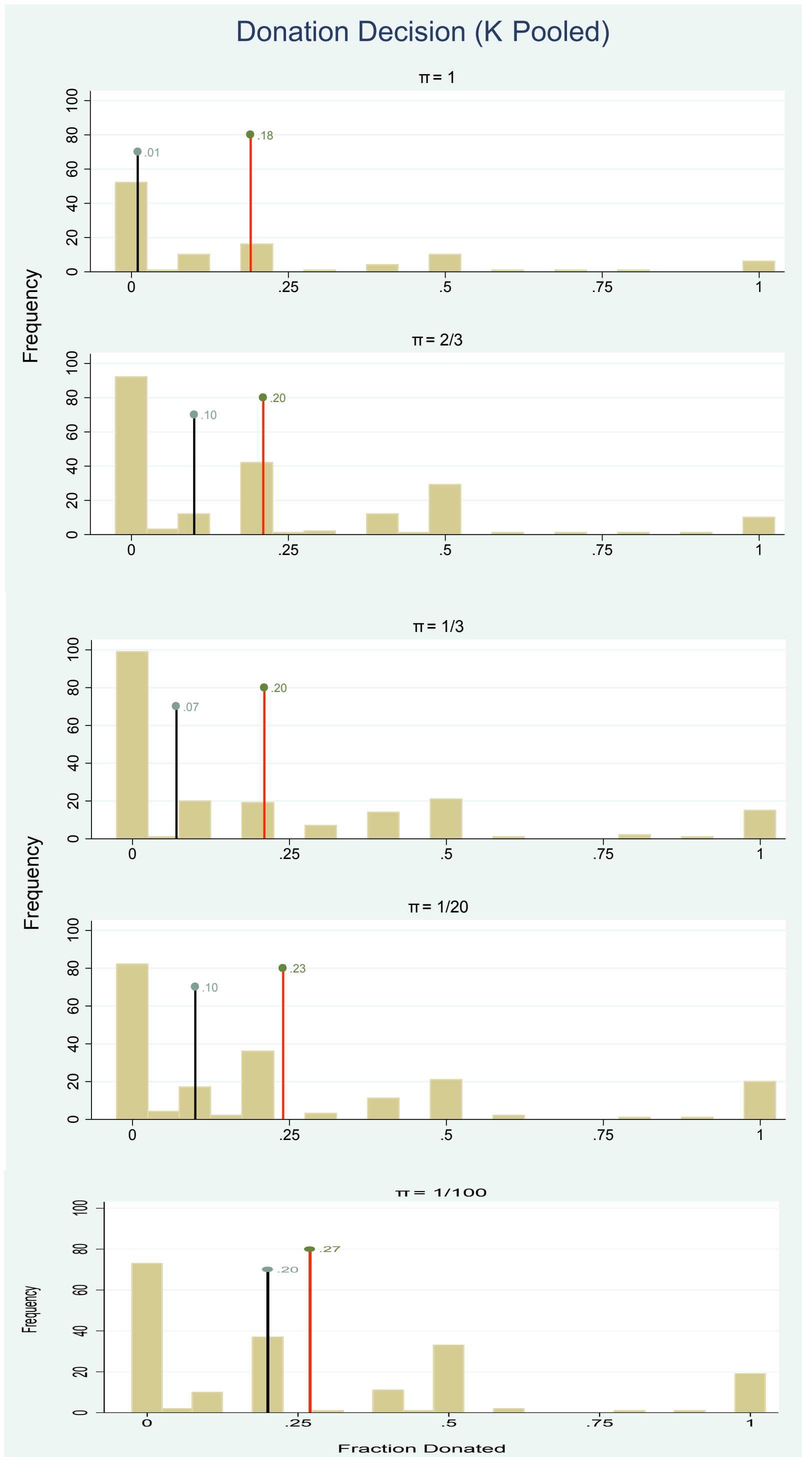
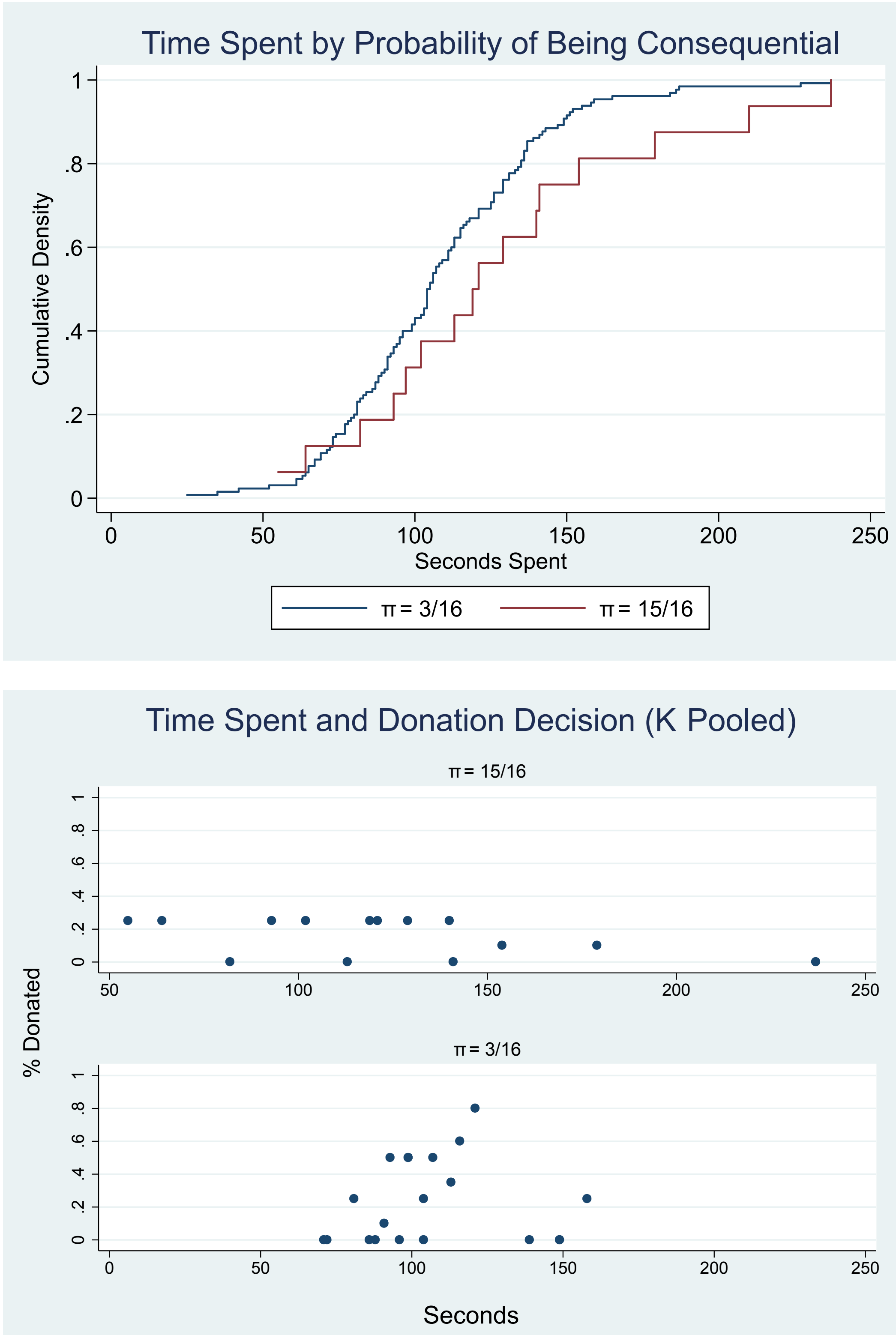


Figure 3:
Red: Mean, Black: Median



Time Spent by Probability of Being Consequential
Cumulative Density
Seconds Spent
π = 3/16
π = 15/16
Time Spent and Donation Decision (K Pooled)
π = 15/16
π = 3/16
% Donated
Seconds

Figure 4:

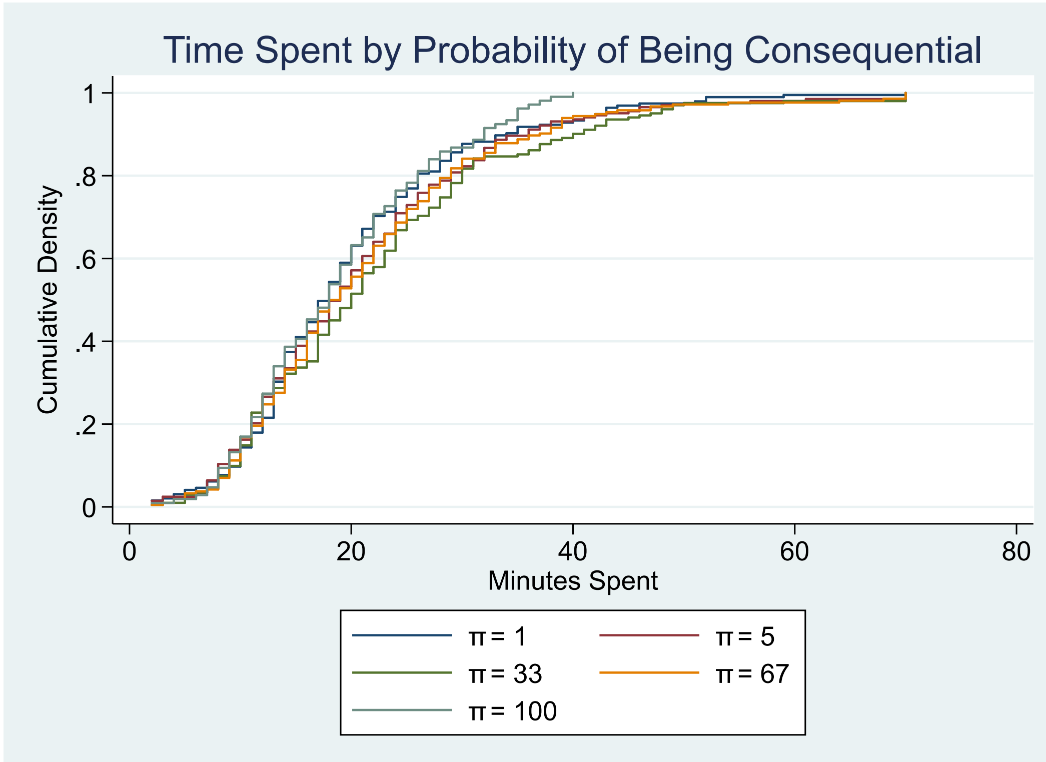


Figure 5:

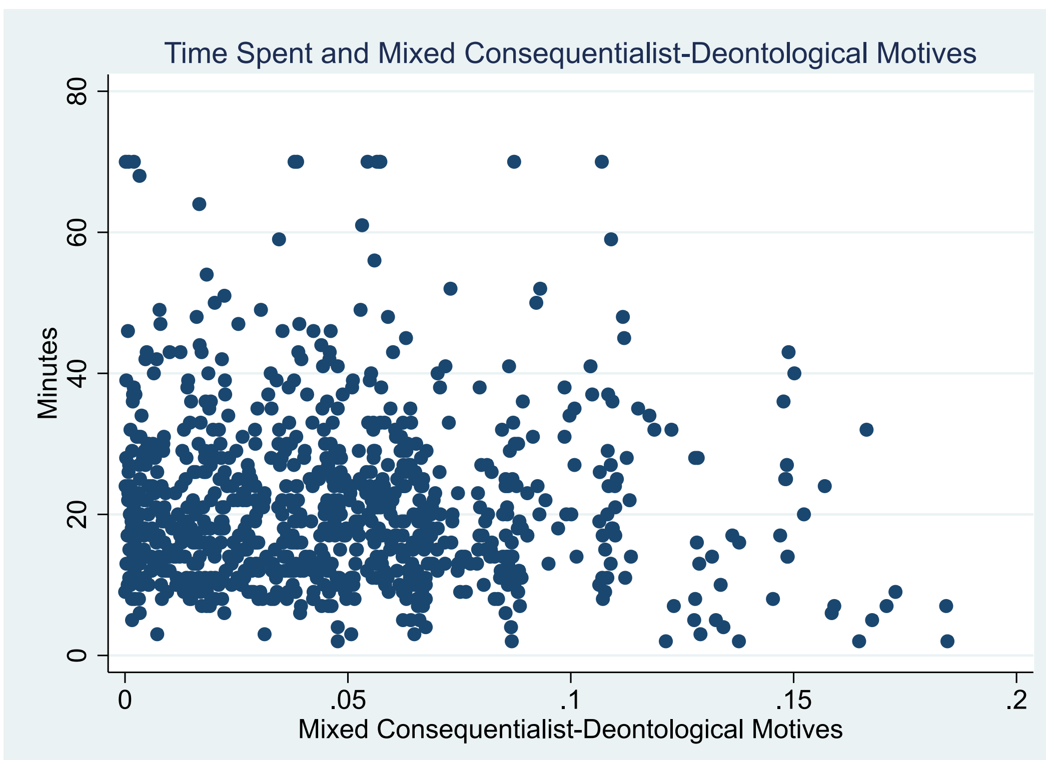


Figure 6:

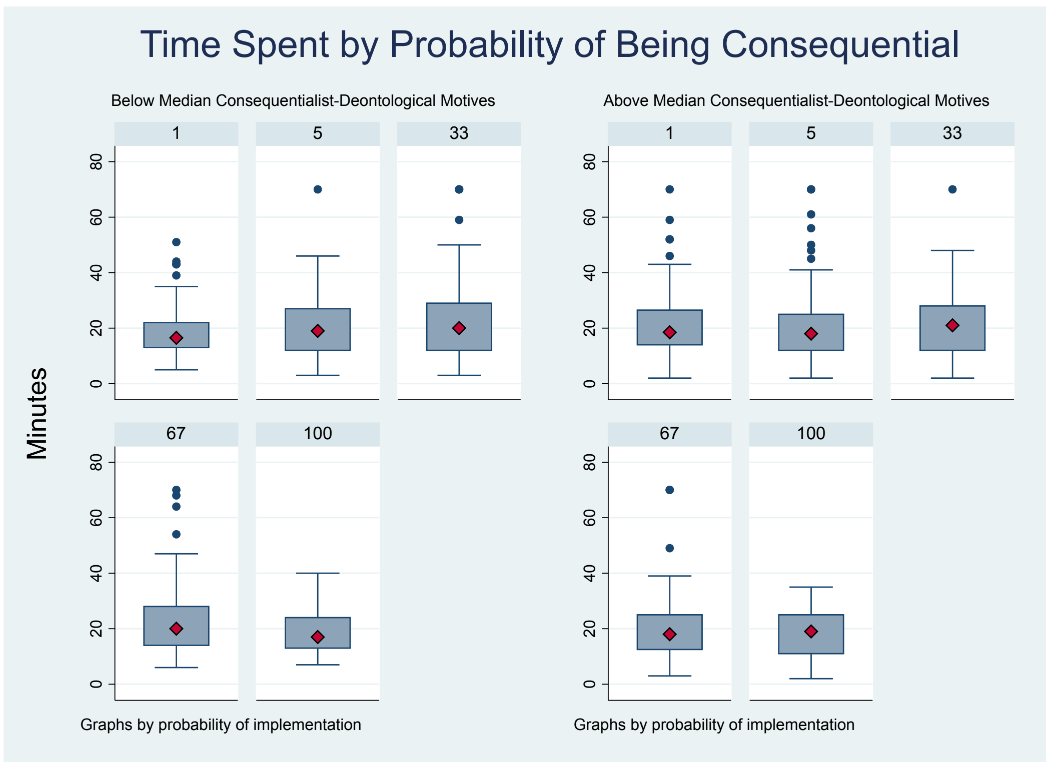


Figure 7:
Red Diamond: Median

Table 1: Shredding Experiment (K Pooled)

	Ordinary Least Squares					
	(1)	(2)	(3)	(4)	(5)	(6)
	Decision (d)		Expected Donation $E(x_R)$		Expected Giving (π* d)	
Mean dep. var.	0.30		0.39		0.12	
% Consequential (π)	-0.211	-0.208*	-0.380**	-0.393***	0.160**	0.161***
	(0.142)	(0.115)	(0.155)	(0.0993)	(0.0658)	(0.0593)
K Fixed Effects	N	Y	N	Y	N	Y
Observations	40	40	40	40	40	40
R-squared	0.055	0.409	0.136	0.665	0.135	0.334

Notes: Standard errors in parentheses. Raw data shown in Figures 1 and 2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: AMT Experiment (K Pooled)

	Ordinary Least Squares					
	(1)	(2)	(3)	(4)	(5)	(6)
	Decision (d)		Expected Donation $E(x_R)$		Expected Giving (π* d)	
Mean dep. var.	0.23		0.34		0.07	
% Consequential (π)	-0.0725**	-0.0684*	-0.224***	-0.219***	0.194***	0.213***
	(0.0288)	(0.0390)	(0.0334)	(0.0299)	(0.0132)	(0.0181)
K Fixed Effects	N	Y	N	Y	N	Y
Controls	N	Y	N	Y	N	Y
Observations	902	900	902	900	902	900
R-squared	0.007	0.059	0.048	0.604	0.194	0.214

Notes: Standard errors in parentheses. Raw data shown in Figure 3. Controls include indicator variables for gender, American, Indian, Christian, Atheist, aged 25 or younger, and aged 26-35 as well as continuous measures for religious attendance and accuracy in the lock-in data entry task. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: AMT Experiment (by K)

	Ordinary Least Squares							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Decision (d)		Decision (d)		Decision (d)		Decision (d)	
	K = Unknown		K = 10¢		K = 0¢		K = 50¢	
Mean dep. var.	0.26		0.22		0.20		0.22	
% Consequential (π)	-0.0778	-0.0654	-0.0525	-0.0321	-0.0711	-0.0708	-0.0644	-0.0675
	(0.0523)	(0.0523)	(0.0526)	(0.0536)	(0.0464)	(0.0466)	(0.0462)	(0.0456)
Male		-0.0909**		-0.0474		0.0108		0.0178
		(0.0399)		(0.0430)		(0.0395)		(0.0362)
American		0.0241		-0.0539		0.0838		0.117*
		(0.0524)		(0.0539)		(0.0664)		(0.0598)
Indian		-0.0672		-0.0785		-0.0673		-0.0626
		(0.0566)		(0.0560)		(0.0630)		(0.0590)
Christian		-0.0295		0.0584		-0.0215		-0.000293
		(0.0483)		(0.0503)		(0.0494)		(0.0479)
Atheist		-0.0188		0.00480		0.0113		-0.0927
		(0.0644)		(0.0649)		(0.0802)		(0.0725)
Religious Services Attendance		-0.00614		0.000508		0.00367		-0.00546
		(0.0145)		(0.0156)		(0.0137)		(0.0137)
Ages 25 or Under		-0.0207		-0.122**		-0.0109		-0.113**
		(0.0518)		(0.0570)		(0.0493)		(0.0474)
Ages 26-35		0.00271		-0.110*		-0.00105		-0.111**
		(0.0548)		(0.0593)		(0.0493)		(0.0480)
Own Errors		-0.000192		-0.000186		0.000220		-0.000148
		(0.000193)		(0.000163)		(0.000194)		(0.000143)
Observations	260	260	218	218	256	255	271	270
R-squared	0.009	0.069	0.005	0.081	0.009	0.052	0.007	0.097

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: AMT Experiment

Thresholds	Wilcoxon-Mann-Whitney 2-sided test (p-values) (1) K Unknown or 10¢	(2) K = 0¢ or 50¢	(3) K Pooled
$\pi = 1$ vs. $\pi \leq 0.67$	0.91	0.05	0.11
$\pi \geq 0.67$ vs. $\pi \leq 0.33$	0.07	1.00	0.20
$\pi \geq 0.33$ vs. $\pi \leq 0.05$	0.05	0.10	0.01
$\pi \geq 0.05$ vs. $\pi = 0.01$	0.15	0.02	0.01

	π Pooled
K ≥ 10¢ vs. K = 0¢	0.40
K = 50¢ vs. K ≤ 10¢	0.11

Table 5: Shredding Experiment

Thresholds	Wilcoxon-Mann-Whitney 2-sided test (p-values) Pooled
$\pi = 3/16$ vs. $\pi = 15/16$	0.30
K = 0 vs. K = Max	0.94

Table 6: Who Is Mixed Consequentialist-Deontological? (AMT Experiment)

	Ordinary Least Squares									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Decision (d)									
Mean dep. var.	0.23									
% Consequential (π)	-0.100**	-0.0493	-0.124**	-0.0500	-0.0522	-0.0774	-0.0618	-0.0548	-0.0839**	-0.0190
	(0.0494)	(0.0429)	(0.0506)	(0.0436)	(0.0403)	(0.0616)	(0.0467)	(0.0443)	(0.0407)	(0.126)
π * Male	0.0612									0.0490
	(0.0577)									(0.0611)
π * American		-0.0675								0.0370
		(0.0627)								(0.0911)
π * Indian			0.0990*							0.0426
			(0.0574)							(0.0963)
π * Christian				-0.0599						-0.0658
				(0.0632)						(0.0783)
π * Atheist					-0.133					-0.145
					(0.0837)					(0.108)
π * Religious Services Attendance						0.00394				-0.00739
						(0.0210)				(0.0224)
π * Ages 25 or Under							-0.0149			-0.0815
							(0.0576)			(0.0787)
π * Ages 26-35								-0.0386		-0.0878
								(0.0597)		(0.0808)
π * Own Errors									0.000402	0.000319
									(0.000299)	(0.000307)
K Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	900	900	900	900	900	900	900	900	900	900
R-squared	0.061	0.061	0.063	0.060	0.062	0.059	0.059	0.060	0.061	0.068

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Time for Completion of Survey (AMT Experiment)

Sample	All Subjects	Above Median Mixed-Consequentialist		Below Median Mixed-Consequentialist	
	(1)	(2)	(3)*	(4)	(5)*
Mean dep. var.			20.8		
% Consequential (π)	0.0123	0.0176	0.0452	0.163***	0.118*
	(0.0162)	(0.0547)	(0.0574)	(0.0548)	(0.0635)
π^2		-0.000482	-0.000452	-0.00167***	-0.00122*
		(0.000573)	(0.000602)	(0.000581)	(0.000674)
Above Median Mixed-Consequentialist	0.755				
	(1.119)				
π * Above Median Mixed-Consequentialist	-0.0386*				
	(0.0227)				
Observations	900	449	449	451	451
R-squared	0.004	0.008		0.019	

Notes: Standard errors in parentheses. Mixed-Consequentialist aggregates for each subject their demographic characteristics' contribution to the effect of π on the Donation decision. Regressions are weighted by the standard deviation of the first regression to account for uncertainty in the calculation of mixed-consequentialist score. Columns 3 and 5 employ median regressions. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 8: Trading Off Consequentialist-Deontological Motivations (AMT Experiment)

	OLS	IV	IV
	(1)	(2)	(3)
		Decision (d)	
Mean dep. var.		0.23	
% Consequential (π)	-0.239***	-0.363***	-0.368***
	(0.0249)	(0.0548)	(0.139)
π * 1(d ≥ w/2)	0.870***	1.516***	1.542**
	(0.0412)	(0.250)	(0.714)
Constant (Duty Bliss Point)	0.251***	0.249***	0.249***
	(0.0116)	(0.0131)	(0.0134)
IV	N	π, Indian	π, Age ≤ 25
Observations	902	902	902
R-squared	0.336	0.155	0.140

Notes: Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01