Behavioral Sources of the Demand for Carbon Offsets: An Experimental Study

Kai-Uwe Kuhn
School of Economics
University of East Anglia

Neslihan Uler
Department of Agricultural and Resource Economics
University of Maryland

Behavioral Sources of the Demand for Carbon Offsets: An Experimental Study[*]

Kai-Uwe Kuhn[†]   Neslihan Uler[‡]

May 28, 2018

[†]School of Economics, University of East Anglia; e-mail: k.kuhn@uea.ac.uk.

[‡]Agricultural and Resource Economics, University of Maryland; e-mail: neslihan@umd.edu.

1

Abstract

Voluntary carbon offsets–that is, payments toward the reduction of damages to the environment–present firms and individuals with the opportunity to neutralize (or lessen) their carbon footprints. In this paper we report on a controlled laboratory experiment to understand the behavioral motivations driving the purchase of carbon offsets, in addition to investigating the effect of the introduction of voluntary offset markets on emission-causing activities. We find that, when the price is sufficiently low, demand for offsets is stable. Behavior is, however, heterogeneous. Individuals with a high (low) personal-responsibility index increase their offset purchases as their own damage (total damages) increases, but do not condition their offsetting behavior on the total damages (own damages) generated. We also show that, when individuals trade in competitive markets, the availability of offsets does not affect the total damages generated. Instead, introduction of carbon offsets increases individuals' earnings by eliminating some of the damages ex-post.

1 Introduction

In recent years, many organizations have been created to increase individuals and firms awareness of their carbon footprints (emissions resulting from everyday activities like driving, heating homes/workplaces and production activities) and present them with opportunities to offset/lesson resultant environmental damages. In particular, organizations like myclimate.org, carbonfootprint.com, carbonfund.org, and terrapass.com offer carbon offsets to finance emissions-reducing activities.[1] Alongside these developments, we have witnessed the establishment of centralized trading institutions, like the Chicago Climate Exchange, and of private companies serving as aggregators for small scale carbon offset supplies (typical in agriculture).

The purchase and trading of voluntary carbon offsets differs from more traditional emissions-targeting schemes (like those commonly observed in the electricity industry), which place ex-ante caps on aggregate emissions, and allow pollution permits to be traded on formal exchanges. In these schemes emissions reductions are mandated and trade occurs as a consequence of profit-maximizing incentives.

Voluntary carbon offsets, on the other hand, do not mandate emissions reductions per se, instead, they aim at changing the preferences of individuals by making them aware of the externalities they cause, and rely on their willingness to reduce these externalities by either avoiding externality causing activities or financing carbon emissions reducing investments elsewhere (e.g., supporting energy efficiency projects and tree planting). Carbon offset schemes usually depend on the belief that carbon reduction should start with the personal responsibility of individuals for the emissions they cause themselves.[2]

In this paper we design and conduct a laboratory experiment to assess the potential for such behavioral motivations to generate private demand for carbon offsets, in addition to investigating the effect of the introduction of carbon markets on emission-causing activities.[3]

---

[1]See also Capoor and Ambrosi (2009), Peters-Stanley and Yin (2013) and Hamrick and Gallant (2017) for overviews of the state of carbon markets.

[2]In fact, there is a growing literature that links feelings of personal responsibility to voluntary carbon offset purchases (e.g., Brouwer et al., 2008; Blasch and Farsi, 2014; Blasch and Ohndorf, 2015; Lange et al., 2017; Lange and Ziegler, 2017) which we discuss in Section 2 in detail.

[3]Controlled laboratory experiments have been used to evaluate economic behavior and environmental policy for several decades. Early examples of testbedding include Plott and Hong (1982) and Grether, Isaac

Our experiment considers a two-stage environment. In the first stage, subjects are given a chance to trade in a competitive market environment. Each trade among a buyer and a seller generates positive individual returns to the two subjects, but also generates carbon emissions (negative externalities) that harm everyone in the experiment equally. After the trading stage has concluded, agents are asked to decide whether and how many carbon offsets to purchase to ameliorate all (or part of) these damages.

Existing experimental research on behavior in anonymous competitive markets appears to confirm the traditional economic view that, in order for individuals to internalize negative externalities, explicit monetary incentives (*i.e.*, taxes) are necessary. A classic experiment by Plott (1983) showed that, while in a competitive market (modeled as a double auction), individuals ignore the negative externalities arising from their transactions, and trading achieves the inefficient competitive equilibrium, an optimal tax to internalize the externality led reliably to efficient solutions. Harrison et al (1987) showed (in a similar setting to Plott's) that allowing side contracting in face-to-face Coasian bargaining between all subjects generated efficient solutions, as well. The transaction costs of Coasian bargaining over issues like carbon emissions are, however, likely to be prohibitive, limiting the applicability of Harrison et al.'s results in the present context.

This paper examines a setting in which subjects first trade in a double-auction market similar to the one in Plott (1983) and then are given the opportunity to buy carbon offsets to mitigate the negative externalities created in the trading stage. Our baseline treatment is based on Plott (1983) and it does not have a carbon offset market (*i.e.*, there is no second stage). In our two other main treatments, we introduce a carbon offset market after the trading stage. The price of offsets varies between these two treatments in order to capture the effect of prices on offset purchases and trading behavior.

While Plott's results suggest that negative externalities cannot be mitigated in a double-auction environment without monetary incentives, this does not necessarily imply that carbon offsets would necessarily be ineffective in diminishing externalities. In a double-auction with external effects (without the possibility of carbon offsets as in Plott's set-up), social

and Plott (1981). For surveys see, for example, Mestelman (2000), Sturm and Weimann (2006), Ehmke and Shogren (2008), Cason (2010), Friesen and Gangadharan (2013), and Noussair and van Soest (2014).

preferences may not be observed because a given individual's effect on others is not directly observable and, in addition, involves a complex tradeoff by requiring that fewer privately advantageous trades be completed. Cognitively, it may therefore be difficult to compare costs and benefits of other-regarding behavior in a double auction setting (Sandel, 2012; Kube, Marechal and Puppe, 2012; Falk and Szech, 2013).[4] On the other hand, our environment introduces a carbon offset market after the double-auction takes place. This could alleviate negative externalities through two different channels. First, the mere presence of a mechanism to make offset payments may activate the type of other-regarding preferences that we see in a public goods game, and might lead to carbon offset purchases to mitigate negative externalities, ex-post. Second, making the external effect of trading more salient through an offset market may also trigger norm activation for norms requiring the avoidance of causing damage to others (and may lead to ex-ante externality mitigation via a lower number of trades in the double-auction).[5] For example, Bartling, Weber and Yao (2015) find evidence for socially responsible behavior in a competitive market set-up, where sellers decide on a price and on which type of product, dirty or clean, they want to offer for sale, and buyers decide which product to buy.[6] Similarly, in our experiment, the availability of offset payments could modify behavior in the initial double-auction in comparison to the results of Plott (1983).

The second stage resembles a classic public goods game. Several research studies have documented that individuals may have other-regarding preferences and contribute to public goods at much higher levels than predicted by standard theory (*i.e.*, Andreoni, 1989, 1990; Ledyard, 1995; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Andreoni and Miller, 2002; Charness and Rabin, 2002). Another major finding in this literature is that contributions decline and free-riding increases with repetition (Ledyard, 1995; Chaudhuri,

---

[4]Alternatively, Andreoni (1995) argues that the reason negative externalities are ignored is because of negative framing (cold-prickle of doing something bad), instead of positive framing (the warm-glow of doing something good), as in public goods experiments.

[5]For a general overview of how social norms may shape economic behavior, see Elster (1989). In addition, Henrich et al. (2001, 2005) demonstrate how market integration may affect social norms of cooperation in a cross-cultural study with 15 small-scale societies. Vandenbergh and Steinemann (2007) have argued that activation for an emerging norm for carbon neutrality could have a major impact on reducing carbon emissions.

[6]Bartling, Weber and Yao (2015) show that while standard theory predicts only the cheaper good will be produced and traded, in an experiment conducted in Switzerland, a significant proportion of products traded are socially responsible.

2011). Nevertheless, contributions usually stay above the complete free-riding equilibrium prediction.

The setting in this paper differs from the previous literature by studying the impact of behavior in the double-auction (first) stage on behavior in the carbon offset (second) stage, as well as the potential feedback effect from the opportunity to purchase carbon offsets on trading behavior. There are several potential reasons why behavior in our set-up might differ from that observed in a typical public goods game. First, the degree to which a public goods problem exists depends on the degree of trading in the first stage. Second, behavioral motivations in the carbon offset stage may be quite different than in a public goods game. For example, subjects may treat damages caused by themselves and those caused by others very differently. Since personal contribution to the negative externalities is observable by subjects, this may affect their decisions regarding their purchases of carbon offsets.

We show that, when the price is low, there is a moderate demand for carbon offsets and no decay over periods (in contrast to typical public goods experiments). More importantly, we find that many subjects act based on their personal contribution to the damage: they increase their offset purchases as their own damages increase–a behavior that we will call "personal responsibility" driven contributions. In particular, we construct a "personal responsibility index" derived from a series of survey questions eliciting information on subjects' concern for their contribution to externalities–such as engaging in non-monetary, time consuming projects like volunteering and/or recycling. We show that personal responsibility driven contributions occur among the set of subjects who have a high personal-responsibility index.

Interestingly, trading in the double auction market does not seem to vary with the presence of the opportunity to purchase offsets, nor with the cost at which carbon offsets can be acquired. In other words, the presence of the second stage carbon offset market does not affect the level of negative externalities, ex-ante. There is a demand for carbon offsets ex-post, however, which leads to a decrease in the level of negative externalities, ex-post. This seems to indicate a strong tendency for mental accounting (where externalities are ignored in trading, but make a difference when considering carbon offsetting later on).

We stress that our results are not only relevant for carbon offset markets, but have much broader implications for collective action problems with negative externalities, in general.

6

Our experiment refrains from using carbon offset terminology. We use terms like "damages" and "damage offsets", instead of "environmental damages" and "carbon offsets". This neutral terminology mitigates framing and experimenter demand effects, in addition to making our results generalizable to other settings. Our results show that individuals' reactions to collective action problems will partially depend on how they contributed to the problem arising in the first place. Specifically, our results shed light on how the personal perception of a contribution to a negative externality problem will affect the willingness to contribute to its solution.[7]

The remainder of this paper is organized as follows: In Section 2, we briefly review the related literature. In section 3, we introduce our novel two-stage trading environment. The experimental design and procedures are presented in detail in Section 4. We provide our experimental findings for the second stage of damage offset purchases in Section 5, while Section 6 explores the impact of anticipated damage offset purchases on trading in stage 1. Section 7 provides further discussion. Section 8 concludes.

2   Related Literature

According to the State of Voluntary Carbon Markets 2017 report from Forest Trends' Ecosystem Marketplace, as of 2016, offsets equivalent to 1.1 billion metric tonnes of carbon dioxide emissions have been transacted voluntarily by governments, companies, individuals and intermediary brokers. Prices for voluntary carbon offsets in 2016 were affordable overall, but there was also a high variance in prices. Prices are influenced by factors such as project type and location.[8] For example, offsets from wind projects in Asia were sold at an average of

---

[7]One timely example is renewable power purchase agreements. Presently, many large companies in the Unites States are either making or committing to make renewable energy purchases to reduce their carbon footprint. For example, in January 2018, T-Mobile US announced its intention to transition to 100% renewable electricity by 2021 and joined the 100% renewable electricity initiative RE100 (https://cleantechnica.com/2018/01/30/t-mobile-announces-commitment-to-100-renewable-electricity-joins-re100/). In February 2018, AT&T announced that they will purchase 520 megawatts of wind energy from wind farms in Oklahoma and Texas through NextEra Energy Resources (https://cleantechnica.com/2018/02/12/att-announces-mammoth-520-mw-corporate-ppa-oklahoma-texas-wind-farms/). While these are not carbon offset purchases per se, our set-up has implications for these power purchase agreements, as well.

[8]Conte and Kotchen (2010) systematically investigate the factors that explain large variability in prices for carbon offsetting. While location of providers, location of projects, types of projects, third-party certification

$0.7/tCO_2e$, while the price for offsets from afforestation/reforestation projects in Africa was on average $6.7/tCO_2e$. Approximately 46% of all offsets transacted worldwide originated in Asia, followed by North America, Latin America and Africa.

The first theoretical analysis of voluntary carbon offsets was provided by Kotchen (2009b). Kotchen showed that mean donations do not converge to zero as the economy grows large. The vast majority of other papers pertaining to carbon offsets are empirical studies based on survey data. These papers find substantial demand for voluntary carbon offset purchases in different contexts and investigate the underlying motivations behind these purchases. For example, in the context of air travel, Brouwer, Barnder and Van Beukerign (2008) find that a significant share of air travel passengers are willing to pay a carbon tax to offset their contribution to the emission of green house gases based on the 'polluter pays principle'. Travelers' perception of their own responsibility for climate change is shown to have a significant effect on stated willingness to pay (WTP) for offsetting their emissions.

In the context of vehicle usage, Lange and Ziegler (2017) study the link between offsetting and mitigation activities in reducing environmental pollution using data on driver's license holders in the U.S. and Germany. Lange and Ziegler show, both theoretically and empirically, that individuals are more likely to use carbon offsetting and mitigation channels (*i.e.*, purchase a less-emitting vehicle or a vehicle with alternative drive systems) to reduce the environmental damage, the higher their feeling of responsibility for carbon emissions. Lange et al. (2017) provide both theoretical and empirical evidence on complementarities between offsetting and other pro-environmental activities. Lange et al. (2017) use data from a representative online survey and consider a wider range of seven pro-environmental activities (such as buying an energy-efficient appliance, reducing consumption of meat or dairy products, *etc.*). Their empirical analysis finds that the introduction of carbon offsets has a net positive effect on environmental quality.

Blasch and Farsi (2014), using an online survey in Switzerland, vary both the context and the characteristics of various damage mitigation projects and explore consumers' underlying motivations to buy offsets as well as the relationship between voluntary offsetting and other

---

play a large role in prices, number of projects that a provider manages and a provider's status as for-profit or not-for-profit do not seem to have an effect on prices.

voluntary mitigation behavior. They show that demand for voluntary carbon offsetting is highly context-dependent and strongly varies with the types of mitigation projects offered. Blasch and Ohndorf (2015) test a theoretical framework in which different motivations (pure and impure altruism, internalized norms, and social recognition) for environmental offsetting are considered. They find that motivation for individual offset purchases varies between different types of consumers; WTP for carbon offsetting is mainly driven by internalized norms to avoid pollution (ascribed responsibility), and partly by income, whereas an individual's probability to offset seems to be largely influenced by the expected social (dis)approval for (not) offsetting. Schwirplies and Ziegler (2016) conduct a survey in two countries and investigate the motives for purchasing voluntary carbon offsets as well as green products. They find that environmental awareness, warm glow motives, and the desire to set a good example significantly motivate the choice of both climate protection activities in Germany and the U.S.

Loschel et al. (2013) and Diederich and Goeschl (2014) both conduct framed field experiments to estimate WTP for voluntary climate action with incentivized choices in order to mitigate the potential hypothetical bias of stated preferences methods. Subjects in these experiments decide, under different prices, what quantity of European Union Allowances to buy (which quantities were then retired from the European Emissions Trading Scheme). These papers report much lower WTP compared with the empirical literature based on survey measures.[9] Loschel et al. (2017) also conduct a framed field experiment using European Union Allowances. They study the role of information about the demand revealed by other subjects and the role of collective action on voluntary climate change mitigation. While information on other subjects' demands does not have an effect, collective action increases the WTP by almost 45 percent. Arana and Leon (2013) and Lofgren et al. (2012) both conduct field experiments to study the effect of defaults on voluntary carbon offsetting programs, and provide mixed evidence on the importance of defaults.

Our paper is also related to recent empirical papers that study the impact of green products on energy consumption using field data. The introduction of a green product might

---

[9]Note, however, that these papers focus on a particular climate action policy and results may therefore not be readily generalizable to other types of offsetting. In addition, these papers are susceptible to field price censoring, since prices for these emissions allowances within the experiment cannot be easily distinguished from actual prices in the field.

induce lower levels of guilt in those who use green products and might therefore lead to an overall negative effect, due to a rebound effect (e.g., Kotchen and Moore, 2008; Kotchen, 2009a; Gillingham, Kotchen, Rapson and Wagner, 2013). Kotchen and Moore (2008) study the relationship between voluntary restraint (consuming less electricity) and a voluntary price premium (paying more for green electricity) both theoretically and empirically. They show that non-conservationists reduce their consumption after participating in a voluntary price premium (due to higher prices), but conservationists, who were already exhibiting voluntary restraint, do not reduce their electricity consumption after participation. Jacobsen et al. (2012) find that households participating in a green-electricity program above the minimum threshold level do not change their electricity consumption, but those participating at the minimum threshold increase electricity consumption by 2.5 percent after enrolling in the program. Despite this buy-in mentality, the net effect for the switch to green-electricity for the buy-in households is a reduction in pollution emissions, as the behavioral response is not large enough to offset the environmental benefit of the green-electricity purchase. Harding and Rapson (2014) analyze the introduction of a voluntary carbon offset program on household energy consumption behavior. Using data from Pacific Gas and Electric's ClimateSmart Program, the authors show that many individuals who voluntarily sign up for a carbon offset program increase their electricity consumption following adoption, and argue that the introduction of the carbon offsets is harmful to the environment.[10]

While our paper contributes to the recent and growing literature on carbon offsets, it is novel in several aspects. Our paper provides the first laboratory experiment to study whether carbon offsets increase welfare as well as the channel by which they improve welfare. Using incentivized subjects, we investigate how individuals with varied personal-responsibility indices behave when deciding how many carbon offsets to purchase. Our paper complements previous empirical papers by providing a mechanism to examine questions that are difficult to answer using survey or field data. Our experiment avoids endogeneity bias by randomly allocating subjects to different treatments. In addition, a laboratory set-up can provide ad-

---

[10]Note that this is not a randomized experiment, and, therefore, it is possible that people selecting into this program may have different characteristics than people who do not select into it. In addition, in the ClimateSmart Program, customers choosing to opt-in pay an extra charge per kilowatt-hour. Therefore, how much energy they consume and the volume of offset purchases they make are a joint decision and cannot be isolated from each other. In many carbon offset settings, individuals can, however, choose whether to buy offsets (as well as how much to buy).

ditional control that is not available in the field. For example, in our experiment, individuals know exactly the externalities they create on others as well as total externalities created by everyone.

3   A Simple Model of Trade and Carbon Offsets

Consider a two-stage environment. In the first stage, agents trade in a competitive market environment where trade generates a negative external effect on everyone. We utilize the double-auction with externalities employed by Plott (1983) and Harrison et al. (1987).[11] There are six buyers and six sellers. Each agent can trade up to five units. Each trade causes a damage of four tokens for each of the 12 individuals, *i.e.*, 48 tokens of damage in total.[12] The inverse demand function for buyers and private marginal cost function for sellers are given in Figure 1. Figure 1 also shows the social marginal cost function, which is 48 tokens higher than the private marginal cost function.

In the second stage, after the trading stage has concluded, agents are eligible to purchase carbon offsets at a price $\frac{1}{12} < p < 1$. A unit of carbon offset purchased reduces total damages by 1 token. Each individual benefits from this reduction equally (independent of who purchased it)–that is, the individual return from a unit of carbon offset purchased is $\frac{1}{12}$. If enough damage offsets have been purchased to eliminate all damages, individuals cannot benefit from purchasing more offsets.[13]

In stage 2, it is therefore socially efficient to purchase a damage offset since $p < 1$. Without behavioral preferences, however, purchasing damage offsets is not individually rational. Since $p > \frac{1}{12}$, each unit of damage offset purchased costs more than the benefit to an individual from the associated damage reduction.

---

[11]Double-auction experiments are very successful in creating a market environment, and they have been widely used in experimental papers on environmental regulation (*i.e.*, Plott (1983), Harrison et al. (1987), Ledyard and K. Szakaly-Moore (1994), Muller and Mestelman (1994), Godby et al (1997)).

[12]The details of the experimental design and procedures, including the conversion rate of tokens to dollars, will be explained in detail in Section 4.

[13]In this set-up, it is evident that individuals are contributing towards an environmental public good. This public good is not a standard voluntary contribution mechanism, however, since individuals themselves created damages in the first stage. In addition, contributions towards the public good can benefit individuals only up to the total damages created.

Figure 1: Demand, Supply and Social Cost
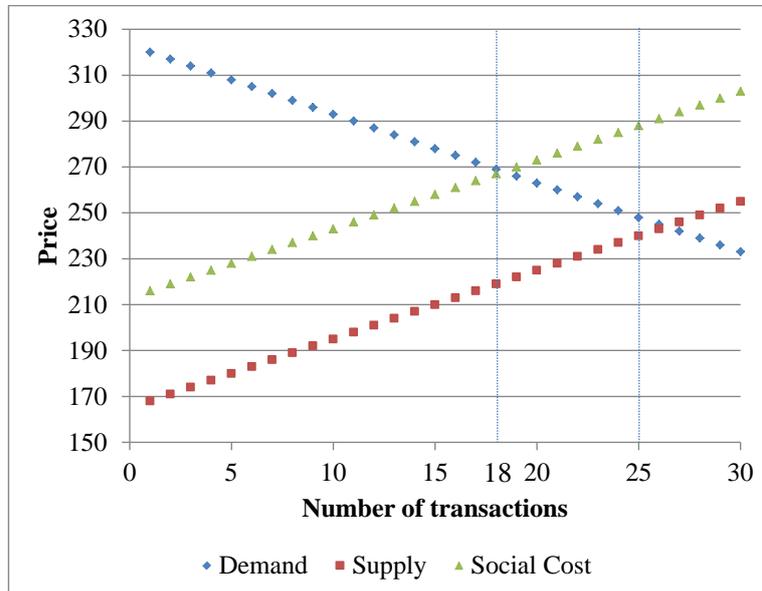
As such, in an environment in which subjects are pure profit-maximizers, we would expect to see free-riding and, hence, no impact (on prices, quantities, or total damages) owing to the availability of offsets. Individuals with purely selfish preferences do not buy any offsets for prices greater than $\frac{1}{12}$ and therefore should not contribute to ex-post damage reduction. In agreement with the results observed by Plott (1983), classical theory predicts that subjects should be trading at competitive prices in stage 1, ignoring negative externalities, with price at 244 tokens and quantity at 25 units.[14]

In the experiment we pick two different prices for carbon offsetting: $\frac{1}{2}$ and $\frac{1}{6}$. Table 1 summarizes the predictions under classical theory corresponding to these prices as well as for the case when carbon offsets are not available after trading. It also shows the efficient levels of trades, environmental damages and total carbon offset purchases as a comparison. Note that throughout the paper when we say a trade/offset is efficient, we refer to it being socially optimal according to a social planner (*i.e.*, only the units that lead to highest possible social surplus are efficient).

---

[14]While demand and supply intersect at quantity 26, because of the damage of 4 tokens per trade, agents would not trade that last unit.

Table 1: Classical Theory Predictions versus Efficient Levels

	Predicted			Efficient		
	Trades	Damages	Offsets	Trades	Damages	Offsets
No offsetting	25	1200	-	18	864	-
$p = \frac{1}{2}$	25	1200	0	22	1056	1056
$p = \frac{1}{6}$	25	1200	0	25	1200	1200

Table 1 confirms that social optimality requires that subjects offset all damages created in the first stage. The socially optimal level of trade depends, however, on the availability of carbon offsets and the price for those offsets. The socially optimal quantity of trades is 18 units when there is no offsetting opportunity, 22 units when $p = \frac{1}{2}$ and 25 units when $p = \frac{1}{6}$.[15] Therefore, by varying the price of offsets, we are able to systematically study individual behavior when socially optimal levels of trades (and damages) differ.

4   Experimental Design and Procedures

Our experiment took place at the Institute for Social Research at the University of Michigan from March-April 2012. Subjects were students at the University of Michigan from a diverse range of academic disciplines and were randomly assigned into treatments. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). Sessions lasted approximately 1 hour and 45 minutes (including payments). Subjects earned experimental currency (tokens) every period. The total earnings of each subject equalled the sum of his or her earnings from each period. The total earnings were converted at the conclusion of the session into US dollars (100 tokens = 1 US dollar). The average payment to subjects was approximately 31 dollars. The experimenter read the instructions aloud at the beginning of each session to create common knowledge. Recall that our experiment does not use explicit "environmental terminology." We refrain from using environmental keywords in order to

---

[15]When there is no offsetting, efficient level of trade is 18 units, since this corresponds to the level at which the demand and social cost functions intersect. When $p = \frac{1}{2}$, it is socially optimal to trade four extra units, since the surplus from these units would be more than enough to offset the extra damages. Trading more than 22 units would not be efficient, because it would be too costly to offset the damages later on. When $p = \frac{1}{6}$, trading at the competitive equilibrium level becomes socially optimal, since all damages created in stage 1 can be cheaply offset in stage 2.

13

mitigate framing and experimenter demand effects. We mention "damages" and "damage offsets", instead of "environmental damages" and "carbon offsets".[16]

We conducted three treatments with three sessions per treatment. Each session had 12 subjects, yielding 108 subjects in total. All three treatments consist of three parts. The first and third parts are identical across treatments.

In part 1, six sellers and six buyers trade for five consecutive periods without an opportunity to offset damages. The trading institution is chosen as a double auction and is based on Plott (1983). This part of our experiment allows subjects to learn about the procedures of the double auction.

Part 2 consists of 10 periods and adds the damage offset game as the second stage. After the conclusion of Part 1 there is a break in which new instructions are given (which depend on the treatment the experimental group is facing). We consider three treatments. In the BASELINE treatment, subjects play the double auction for another 10 periods (without any opportunity to offset damages). This treatment controls for any potential "restart effects" subsequent to the break at the conclusion of Part 1, and provides a benchmark against which the impacts of our treatments can be compared. In the HIGH and LOW treatments, subjects are provided with an opportunity to offset the damages created during the trading stage (as described in Section 3).

Part 3 of our experiment consists of the administration of a questionnaire, which is explained in detail in Section 5. The following explains the procedures of Parts 1 and 2 at a more granular level.

In each trading period each subject can trade up to five units. At any time during the trading period, any buyer is free to submit a "bid" (offer to buy a unit at the price specified in the bid). Similarly, any seller is free to submit an "ask" (offer to sell a unit at the price specified in the ask). All bids and asks pertain to trading one unit. It is not possible to sell/buy two or more units as a package. All active bids and asks are listed on a computer screen in front of each subject. The computer requires subjects to improve on the highest bid (buyers) or lowest ask (sellers) currently posted. At any point in time buyers can accept

---

[16]Instructions for our experiment are provided as supplementary material.

the lowest ask price posted on the screen and sellers can accept the highest bid price posted.

When a buyer (seller) accepts the lowest ask price (highest bid price) posted, he/she receives the value (pays the cost) corresponding to that unit. The value to the buyer or the cost to a seller depends on how many units the subject has traded previously. If a bid/ask is accepted, a binding contract has been closed for a single unit and the buyer and seller see their corresponding earnings on their computer screens. After each contract is closed, all previous bids and asks are automatically withdrawn before any new ones can be made.

Before the experiment took place, for each trading period, we randomly selected the values and costs of buyers and sellers from the values and costs listed in Figure 1. In other words, for a given session, buyers and sellers faced randomly selected values and costs for each period. These values and costs were kept the same, however, across all treatments and sessions to make sure any differences we observe across treatments can be attributed to treatment effects (and are not due to differences in the distribution of values and costs across treatments).

All buyers and sellers have anonymous identification numbers. Computers report the ID number of the buyer or seller posting a bid/ask. The roles of buyers and sellers were kept constant through out the experiment.

At the beginning of each period, subjects receive 100 tokens (to cover possible losses due to damages). The surplus, s_{ij}, from each trade is computed by taking the difference between the unit value/cost and the price of subject i when trading unit j, where $j = (1, \ldots, 5)$. The total surplus, S_i, of each subject is the sum of the surplus from all trades.

Subjects are told that every completed trade causes a damage of four tokens for everyone in the experimental session. The damages incurred by each subject are equal to $D = 4n$, where n denotes the total number of trades across all 12 subjects in a given session. The earnings, E_i, of subject i in a given period are then given by:

$$E_i = 100 + S_i - D.$$

In Part 2 of the HIGH and LOW treatments, each of the 10 periods consists of 2 stages. The first stage is exactly the same as a trading period in Part 1. In the second stage, the

15

"damage offset stage," there is an opportunity for each subject to buy damage offsets. This allows subjects to reduce the damages that have been created in the trading stage.

Buying damage offsets is entirely voluntary, and 1 unit of damage offset costs $\frac{1}{2}$ token in the HIGH treatment, and $\frac{1}{6}$ token in the LOW treatment. Any subject may buy damage offsets and reduce the damages up to the total number of damages that were created in the trading stage. Each subject decides how many damage offsets, x_i, to purchase without knowing the decisions of others. The total number of damage offsets purchased is denoted by X.

A unit of damage offset purchased (independent of who purchased it), reduces the total damages by 1 token. Each subject benefits from this equally. So, each subject receives a damage rebate of $\frac{1}{12}$ token. The same is true for all the units purchased up to the total number of damages ($TD = 12D$). If enough damage offsets have been purchased to eliminate all damages, subjects do not benefit from purchasing additional offsets.

The earnings of subject i are equal to his/her earnings from the trading stage minus his/her payments for damage offsets plus the damage rebate (when $X \leq TD$):

$$E_i = (100 + S_i - D) - px_i + \frac{1}{12}X.$$

If X is more than the total damage, TD, then the earnings of subject i are given by

$$E_i = (100 + S_i - D) - px_i + \frac{1}{12}TD = 100 + S_i - px_i,$$

We now present the hypotheses that we seek to test in our paper. We begin with hypotheses concerning behavior/outcomes in the second stage of the game, since it is the focus of the paper and, furthermore, predictions concerning stage 2 are necessary to develop hypotheses for stage 1. We start with two standard hypotheses for public goods experiments, in general, and are thus of interest to us, as well.

Based on classical model predictions with profit-maximizing agents, Hypothesis 1 states that subjects will free-ride and not buy any offsets.

Hypothesis 1: Subjects will not purchase any damage offsets.

Hypothesis 2, below, is aimed to test whether individuals react to offset prices in a rational way.

Hypothesis 2: The quantity of damage offsets purchased does not change with the price of offsets.

The remaining hypotheses constitute the core contributions of our paper. In particular, we predict that subjects will change their behavior in the second stage depending on the outcomes of the first stage. Put differently, we test whether subjects condition their purchases on own damages created, total damages created and surplus from trade.

In a standard public goods game, the only relevant information subjects have is about the total level of the public good and the cost to themselves to contribute to public good provision. For a given cost of contribution, giving therefore depends only on how much of the public good (here, total damage reduction) is achieved. In the second stage game we analyze here, individuals know the level of damages they personally generated, how much damage others have generated and how much surplus the subject obtained from the trades that generated the damage.

These different pieces of information make it possible that very different motivations generate contributions to damages reductions in contrast to the usual public goods games in the literature. To use simple terms, we call behavior that is based on the total damage "altruism", representing the same motivation that we see in typical public goods experiments. When the purchase of damage offsets is generated from the damages created by one's own trading, we call this motive "responsibility." Finally, we associate damage offset purchases explained by the surplus generated from trade as "fairness" driven, *i.e.*, subjects who make a relatively higher surplus in stage one are expected to contribute more to damage reduction due to fair sharing of trading surplus.

Furthermore, we would expect there to be considerable heterogeneity in the degree to which individuals exhibit responsibility, altruism, and fairness motives. For example, a person strongly motivated by personal responsibility may not exhibit altruism and vice versa, representing different ethical types. In this case it may be difficult to identify the average effect of total and individual damages, because effects of the two may cancel each other out on average.

17

To deal with this heterogeneity between subjects, we control for subjects' types through a personal responsibility index that captures time and effort spent towards public goods provision. This is explained in detail in section 5.

We expect individuals with a high personal responsibility index to treat damages caused by themselves versus those caused by others very differently. For these individuals, there will be a positive relationship between offsets purchased and the damages caused by their own trades. This type of personally-responsible behavior is consistent with warm-glow motives (see, for example, Blasch and Ohndorf (2015), Schwirlies and Ziegler (2016) and Lange and Ziegler (2017) in the context of damage offsets). In contrast, individuals with a low personal responsibility index would not increase offset purchases when they themselves cause higher damages. We should therefore see an increase in the response to damages caused by the individual the greater the personal responsibility index of that individual.

A low personal responsibility index rating does not automatically imply no offset purchases if there are altruists in the population. While we do not expect these subjects to condition their offset purchases on own damages created, they might still be motivated by the total amount of damages created in stage 1. In contrast, for individuals with a high personal responsibility index, we do not expect a significant relationship between offset purchases and total damages created by everyone (when their own damages created are controlled for). We should therefore see an increase in the response to total damages the smaller the personal responsibility index of a given individual.

To test these predictions, we construct hypotheses based on the classical model (see Hypotheses 3-7). Rejecting these hypotheses will give us support for our predictions.

Hypothesis 3: Damage offset purchases of an individual do not vary with the size of the trading surplus.

Hypothesis 4: Damage offset purchases of an individual do not vary with the size of the own damage created.

Hypothesis 5: Damage offset purchases of an individual do not vary with the size of the total damage created.

Hypothesis 6: There is no interaction between the level of the personal responsibility

18

index and the size of own damages on the purchase of damage offsets.

Hypothesis 7: There is no interaction between the level of the personal responsibility index and the size of total damages on the purchase of damages offsets.

Finally, in Section 6, we analyze the impact of the anticipation of second stage damage offset purchases on trading in stage 1. Different theories yield different predictions for stage 1. First, a rebound effect would lead to more trading in the LOW and HIGH treatments since buying offsets might crowd-out intrinsic motivations and lower guilt levels. Second, a saliency effect introduced through the second stage should lead to less trading in period one in the HIGH treatment.[17] Third, if trading is not affected at all by expected second stage outcomes, the result would be consistent with mental accounting (as well as the classical theory). These lead us to Hypothesis 8.

Hypothesis 8: The number of trades will be the same across all treatments.

5 The Determinants of Damage Offset Purchases

This section analyzes the behavior of subjects in the damage offset stage of the experiment. We start by investigating whether there is any demand for carbon offsets, and also examine whether demand varies with price. We then explore the potential motivating factors for the purchase of damage offsets as well as the observed heterogeneity between subjects.

The average (standard error) offset purchase in the LOW treatment is 38.09 (5.64) and 6.36 (1.38) in the HIGH treatment (see Figure 2). OLS regressions (with robust standard errors clustered at the subject-level) confirm that subjects buy positive levels of offsets in both treatments. Therefore, we can reject Hypothesis 1 (both $p-values < 0.01$).[18]

---

[17]Any effect that would make the damages caused by trading more salient during the trading period should only have an effect on trading in the HIGH treatment. Recall that trading at the competitive levels in the LOW treatment is efficient. The HIGH treatment is the only treatment in which a reduction of trade is cheaper than buying damage offsets ex-post and the introduction of offset purchases can have an impact on the saliency of damages for the trading period. In the long run, players learn that damage avoidance is cheaper through decreased trade, and less trade could occur in the HIGH treatment than when damage offsets are not offered.

[18]Regression results available from the authors upon request.

Figure 2: Mean Offset Purchases per Period
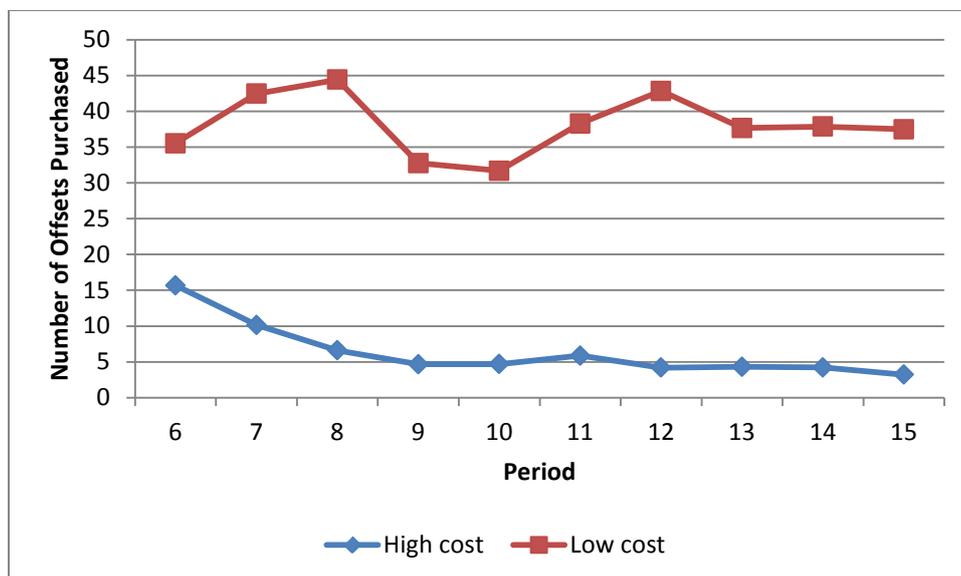
We now consider the average treatment effect of offering damage offsets at different prices. Figure 2 shows large differences in the amount of offsets purchased.[19] The difference is significant at the 1% level (OLS regression, $p-value < 0.01$), which leads us to reject Hypothesis 2. Subsequent analyses confirm that this result is robust to controlling for additional relevant variables, including trading surplus, damages created and demographic characteristics by subject (see Table 3).

Another important observation is that we see the typical decay in offset purchases over time in the HIGH treatment, replicating the well-known result from most public goods experiments. Yet, a first indication that the observed behavior in our setting is quite different from that in standard public good experiments comes from the fact that we do not observe any such decay in the LOW treatment. On average the total damages with which subjects start the second stage in the LOW treatment is 1192. The average number of total offset purchases is 457.13. The percentage of damages offset in the LOW treatment is stable over time and on average 38.35% per trading period. The regression analyses presented below confirm these observations.[20]

---

[19]Recall that the offset stage is introduced in Part 2 (periods 6-15).

[20]In addition, we also perform a simple linear regression analysis to estimate asymptotic investment be-

The dependent variable in the regression is the level of damage offsets purchased by a subject in a period (*DamageOffsets*). As explanatory variables we use the three observed features of the outcome of first stage trading: (1) the total surplus of a subject from the trading stage (*Surplus*), (2) the total damages created by the trades of each subject in a given period (*DamagesCaused*), and (3) the total damages created by *all* subjects in a given period (*TotalDamages*). These variables are observable to the subjects after stage 1, so that they can make informed choices in stage 2. Note that the variable *DamagesCaused* is defined as the total damages created by the trades of each subject in the trading stage instead of being defined as the damages created by each subject. While we could have split this variable into two and explained to subjects that they are technically responsible for half of these damages only, we chose to let subjects decide for themselves how much of damages generated by their trades are their responsibility. To see if buyers and sellers feel differently about these damages, our regression analyses will add a dummy variable for sellers as a control.

We then construct a variable from our survey through which we attempt to capture heterogeneity in the weight on the variable *DamagesCaused* between different subjects. The idea is to capture the notion of responsibility. The variable *Responsibility* is equal to the personal-responsibility index we derive using our questionnaire data. The index is a measure of how much a subject puts effort and time towards public goods provision, and it relies on three different questions from the questionnaire. Subjects are asked how often they donated blood during the last 3 years, how often they have done any kind of volunteer work in the past 12 months and how often they reuse or recycle. The personal-responsibility index is a summation of the scores from these three questions. It takes values between 3 and 15. Higher scores correspond to higher levels of dedication towards public goods. We interpret such reported behavior as a subject being more *personally* responsible in taking an action that improves public goods.

---

havior. Similar to the analysis presented in Noussair et al. (1995), we regress damage offset purchases on $\frac{1}{t}$ and $\frac{t-1}{t}$, where t represents the period number. Notice that as t gets larger, $\left(\frac{1}{t}\right)$ approaches zero, and $\left(\frac{t-1}{t}\right)$ approaches 1. Hence, the coefficient in front of $\left(\frac{t-1}{t}\right)$ gives the asymptotic estimate of investment, and the coefficient in front of $\frac{1}{t}$ gives an estimate for the initial investment. While mean offset purchases are estimated to start at 15.90 (with a robust standard error of 3.37) in the HIGH treatment, they decay over the periods and converge to 2.41, which is not significantly different from 0 ($p-value = 0.19$). In contrast, in the LOW treatment, offset purchases are estimated to start at 37.75 and converge to 38.24, with robust standard errors of 7.96 and 4.97, respectively.

As further controls, we allow for a time trend to capture the typical decaying trend in public good experiments towards smaller contributions over time. It is captured by the variable *Period* (taking values from 1 to 10). We also allow for differences between sellers and buyers using the dummy variable *Seller* that takes value 1 for a seller and 0 for a buyer.

To control for other heterogeneity in the level of damage offset purchases, we include other demographic variables from the survey: *Age* (age of the subject in years), *Female* (1 if the subject is female and 0 otherwise), *MajorEcon* (1 if the subject is an economics major and 0 otherwise), and *FamilyIncome* (a categorical variable that takes values from 1 to 4, where 1 corresponds to the lowest income class).

In addition to the variable *Responsibility*, we include variables that capture heterogeneity in normative judgements of subjects: *Conservative* (1 if a subject picked the category conservative as their political view (over moderate or liberal), 0 otherwise), *Liberal* (1 if a subject picked the category liberal as their political view (over conservative and moderate), 0 otherwise.), *Religion* (from "not important" (1) to "very important" (4)), *GiveHomeless* (donations to homeless people in the past 12 months from none (1) to more than once a month (5)), *SocialPolicy* (a measure of how much a subject cares for social policies for the provision of public goods),[21] *Unemp_vs_Env* (agreement with statement "If we want to combat unemployment in this country, we shall just have to accept environmental problems" from strong agreement (1) to strong disagreement (5)),[22] and finally *Trust* ("Most people can be trusted", strong disagreement (1) and strong agreement (5)). The summary statistics for these variables are provided in Table 2.[23]

Table 3 shows the OLS regressions of damage offset purchases on these explanatory variables.[24] The first two columns present the results for pooled data (from LOW and HIGH

---

[21]This is constructed from the summation of the scores from two different questions from the questionnaire. Subjects were asked whether they would accept higher prices to protect the environment and whether they agree with the statement "Those in need have to take care of themselves." It takes values between 2 and 10. Higher scores correspond to higher levels of agreement with social policies towards public goods provision.

[22]*Unemp_vs_Env* captures the valuation between two public goods: employment and the environment. Given that our experiment has a neutral framing, it is not surprising that this variable does not have any explanatory power.

[23]While in our experiments, subjects are randomly distributed to each treatment, there are still some imbalances observed across treatments. In particular, we have a higher proportion of females in the LOW treatment compared to the HIGH treatment ($p-value = 0.06$). Therefore, to elicit true treatment effects, it is important to control for these variables with regression analysis.

[24]Since offset purchases are censored from below, we also provide Tobit regression analysis in Appendix A

Table 2: Summary Statistics per Treatment

Variables	BASELINE			LOW			HIGH		
	mean	min	max	mean	min	max	mean	min	max
DamageOffsets	-	-	-	38.1	0	200	6.4	0	100
				(41.4)			(13.3)		
Surplus	163.8	43	327	164.0	19	327	165.1	67	339
	(55.8)			(56.7)			(47.8)		
DamagesCaused	202.1	48	240	198.7	96	240	201.3	96	240
	(38.6)			(41.1)			(36.7)		
TotalDamages	1212.8	1104	1344	1192	1104	1248	1208	1152	1344
	(52.6)			(43.1)			(51.3)		
Female	0.47	0	1	0.61	0	1	0.39	0	1
	(0.51)			(0.49)			(0.49)		
Age	21.5	18	29	21.7	18	26	21.7	19	29
	(2.3)			(1.8)			(2.2)		
MajorEcon	0.11	0	1	0.06	0	1	0.14	0	1
	(0.32)			(0.23)			(0.35)		
FamilyIncome	2.1	1	4	2.0	1	4	2.1	1	4
	(1.1)			(1.0)			(0.8)		
Conservative	0.14	0	1	0.08	0	1	0.08	0	1
	(0.35)			(0.28)			(0.28)		
Liberal	0.58	0	1	0.50	0	1	0.50	0	1
	(0.50)			(0.51)			(0.51)		
Religion	2.1	1	4	2.1	1	4	2.3	1	4
	(1.1)			(1.1)			(1.2)		
Responsibility	7.4	4	11	7.8	4	13	8.0	5	13
	(1.7)			(1.7)			(2.0)		
GiveHomeless	2.3	1	5	2.3	1	5	2.4	1	5
	(1.0)			(1.2)			(1.2)		
SocialPolicy	5.5	2	8	5.6	3	9	6.0	4	8
	(1.5)			(1.6)			(1.3)		
Unemp_vs_Env	3.5	1	5	3.8	1	5	4.0	1	5
	(1.1)			(1.1)			(0.9)		
Trust	3.3	1	5	2.8	1	5	3.0	1	5
	(1.0)			(1.2)			(1.1)		
Standard deviations in parentheses.									

treatments) with and without the additional controls for demographic and value judgements. The third and fourth (fifth and sixth) column show the results for LOW (HIGH).

The regressions show a general decline in damage offset purchases with repetition, but only for the HIGH treatment is there a strong and statistically significant decline. To test Hypothesis 3, we look at the coefficient of the variable *Surplus*. There is no statistically significant effect of the *Surplus* variable in either treatment. In the pooled regressions, when all controls are added, *Surplus* is positively correlated with higher offset purchases, but the effect is significant only at the 10% level. Therefore, we find weak evidence that individuals with larger surplus from trades will buy more offsets. In other words, we fail to reject Hypothesis 3 at 95% confidence level.

We now test Hypotheses 4-7. For the LOW treatment, *DamagesCaused* and *Total-Damages* are both significant, as are the interactions of these variables with *Responsibility*. First, it should be noted that the interaction between *Responsibility* and *DamagesCaused* is positive, while the interaction between *Responsibility* and *TotalDamages* is negative. This confirms the intuition that the dependence of damage offset payments on damages caused by own trades and total trades is negatively correlated between subjects. In other words, subjects are either "altruism" motivated or "responsibility" motivated. It should also be noted that, when we do additional regression analyses (not shown here), we see that interactions with *TotalDamages* and *DamagesCaused* do not exist for other control variables such as *GiveHomeless*, *Unemp_vs_Env* and *SocialPolicy*, which gives further credence to our interpretation of heterogeneity being primarily dependent on differences in the relevance of personal responsibility in decision making. Taken together, these results provide strong evidence against Hypotheses 4-7.

Although the damage offset payments in the HIGH treatment are quite low and, therefore, it is unlikely to observe significant effects of explanatory variables on offset purchases, these are not totally absent. As in the LOW treatment, the cross-effects of *Responsibility* with *TotalDamages* and *DamagesCaused* are statistically significant (although only weakly), and the signs are consistent with those in the LOW treatment regressions. The coefficients of the other controls for heterogeneity differ, however, between the two treatments. In the

_____

as a robustness check.

Table 3: OLS Regressions

Dependent var = *DamageOffsets*	ALL	ALL	LOW	LOW	HIGH	HIGH
	(1)	(2)	(3)	(4)	(5)	(6)
Surplus	0.28	0.36*	0.51	0.63	0.07	0.08
	(0.21)	(0.21)	(0.40)	(0.39)	(0.07)	(0.07)
*Resp*Surplus*	-0.03	-0.04	-0.05	-0.06	-0.01	-0.01
	(0.02)	(0.02)	(0.04)	(0.04)	(0.01)	(0.01)
DamagesCaused	-0.55**	-0.49**	-0.96**	-0.94**	-0.29	-0.26
	(0.24)	(0.22)	(0.45)	(0.37)	(0.20)	(0.17)
*Resp*DamageC*	0.07**	0.06**	0.12**	0.11**	0.03*	0.03*
	(0.03)	(0.02)	(0.05)	(0.04)	(0.02)	(0.02)
TotalDamages	0.38**	0.36**	0.94**	0.84***	0.09	0.08
	(0.16)	(0.15)	(0.41)	(0.27)	(0.07)	(0.06)
*Resp*TotalD*	-0.04**	-0.04**	-0.09*	-0.08**	-0.01	-0.01
	(0.02)	(0.02)	(0.05)	(0.03)	(0.01)	(0.01)
Responsibility	36.64*	35.80*	96.13*	78.47**	8.12	6.51
	(18.96)	(18.75)	(52.96)	(34.98)	(7.96)	(7.57)
Period	-0.7	-0.7	-0.79	-0.98	-1.03***	-1.03***
	(0.48)	(0.48)	(1.04)	(1.10)	(0.33)	(0.34)
Seller		-0.69		-0.49		0.5
		(4.74)		(7.94)		(2.52)
Female		-5.14		-19.69**		-0.15
		(5.33)		(9.02)		(2.86)
Age		0.75		2		0.56
		(1.60)		(2.71)		(0.53)
MajorEcon		-8.05		-39.07*		-0.53
		(6.91)		(20.39)		(2.93)
FamilyIncome		0.21		-2.19		-0.34
		(2.08)		(3.54)		(1.51)
Conservative		2.8		-8.42		-2.44
		(7.38)		(10.48)		(3.40)
Liberal		-4.61		-12.45		-2.13
		(5.87)		(12.05)		(2.98)
Religion		-3.69		-6.31		-1.16
		(2.37)		(4.28)		(0.83)
GiveHomeless		1.28		4.74		2.00**
		(2.11)		(3.93)		(0.83)
SocialPolicy		3		4.7		0.66
		(2.40)		(4.32)		(0.95)
Unemp_vs_Env		-2.25		-5.48		-0.56
		(3.40)		(5.55)		(0.78)
Trust		2.42		3.45		-0.93
		(2.40)		(4.57)		(1.01)
Low	31.79***	32.77***				
	(4.67)	(4.62)				
Constant	-379.72**	-397.68**	-965.45**	-892.08***	-50.4	-51.77
	(176.35)	(166.67)	(461.08)	(315.43)	(75.79)	(71.23)
Observations	720	720	360	360	360	360
Standard errors are clustered at the individual level.						
Robust standard errors in parentheses.						
*** denotes significance at 1 percent, ** denotes significance at 5 percent, * denotes significance at 10 percent						

LOW treatment, both students with a major in economics and women buy significantly fewer damage offsets. The first is a common result in public goods games, while the second result is surprising.[25] In the HIGH treatment, by contrast, only giving to homeless persons is significant. This is not surprising. Overall giving is so low in this treatment that the usual effect of an economics major is unlikely to be identifiable. Giving to the homeless is unconditional behavior that is likely to be independent of total damages or damages caused.[26]

A closer look at the regressions for the LOW treatment also shows how important it is to control for the heterogeneity between subjects in the weight they are attributing to the total damages and the damages caused by the trades of the subject. If one evaluates the regression at the mean of *Responsibility*, the effects of increasing *TotalDamage* or *DamagesCaused* are not statistically different from zero. This implies that around the mean there is no detectable effect owing to either variable. In fact, when these regressions are run without the *Responsibility* variable (or its interactions) no significant effects for *TotalDamage* and *DamagesCaused* are observed. Indeed, behavior is quite polarized. For values of the personal-responsibility index above the mean, purchase of damage offsets increase when *DamagesCaused* increases, while at lower values, offset purchases increase when *TotalDamages* increases.

This fundamentally different behavior at the extremes of the *Responsibility* variable can be seen more clearly when one divides the sample into three groups of personal-responsibility levels and repeats the regressions without the *Responsibility* variable for each group separately. Table 4 presents the analogue of regressions 1 and 2 in Table 3, where the *Responsibility* variable and its interactions have been dropped.[27] The results show that subjects characterized by a high personal-responsibility measure increase their offset purchases when their share in damages increases, but do not react to total damages. Subjects characterized by a low personal-responsibility measure do not increase their offset purchases when

---

[25]One possible reason for seeing lower contributions by females could be that, as we show in Table 6, females create lower damages by trading less (significant at the 90 percent confidence level). In Appendix B, we also study the determinants of net damages–damages done net of damage offsets.

[26]Note that concern for others' welfare in general is likely not well-captured by the variable *GiveHomeless*. A subject might be very altruistic, *i.e.*, buys more offsets when total damages are higher, yet would not give to a homeless person.

[27]The interpretation of the coefficients in front of the demographic variables (and their statistical significance) may not be reliable, since the data are now divided into three sub-groups (reducing the power of the regressions). Therefore, to simplify the presentation, we do not show these results here. They are available upon request.

their share in damages increases, but they statistically significantly increase their offset purchases as total damages increase. In addition, when demographic and other controls are included, individuals with low personal-responsibility indices increase their offset purchases as their surplus from the trading period increases. On the other hand, individuals with high personal-responsibility indices do not seem to increase their offset purchases as their surplus increases.

Table 4: OLS Regressions By Different Levels of Responsibility Index

Dependent var = *DamageOffsets*	Low Responsibility		Average Responsibility		High Responsibility	
	(1)	(2)	(3)	(4)	(5)	(6)
Surplus	0.08	0.13**	-0.02	0.04	0.01	0.01
	(0.08)	(0.06)	(0.04)	(0.04)	(0.07)	(0.04)
DamagesCaused	-0.09	-0.03	0.01	0.04	0.25***	0.24***
	(0.08)	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)
TotalDamages	0.12**	0.16***	-0.03	-0.03	-0.03	0.04
	(0.06)	(0.04)	(0.05)	(0.04)	(0.06)	(0.06)
Period	-0.49	-0.58	-0.17	-0.27	-1.14	-1.28
	(0.74)	(0.75)	(0.52)	(0.57)	(1.14)	(1.14)
Low	38.49***	39.98***	15.32**	15.19***	43.40***	47.89***
	(7.45)	(4.91)	(7.00)	(4.74)	(8.39)	(9.13)
Constant	-124.89*	-129.05**	40.03	-89.8	-8.28	-28.62
	(63.01)	(57.06)	(70.34)	(109.81)	(65.01)	(58.28)
Controls added	No	Yes	No	Yes	No	Yes
Observations	290	290	220	220	210	210
Standard errors are clustered at the individual level. Robust standard errors in parentheses. *** denotes significance at 1 percent, ** denotes significance at 5 percent, * denotes significance at 10 percent						

To summarize, we see clear evidence that there are individuals for whom the purchase of damage offsets is driven primarily by the damages caused by their own trading. We interpret these subjects as being driven by personal responsibility concerns. Other subjects primarily react to the total damages arising as the result of trading, which we interpret as the typical altruistic behavior observed in standard public goods games. We show that this behavior is highly correlated with a variable that measures the degree to which subjects report that they contribute their time, but not necessarily their money, to public goods provision. We interpret this as a measure of personal responsibility, since it can statistically identify individuals that respond more strongly in their damage offset payments to the damages they directly caused.[28]

---

[28]Harding and Rapson (2014) also identify the extent to which a household is involved in the community

6 Does the Introduction of Damage Offsets Alter Trading Decisions?

In this section we analyze whether the introduction of damage offset purchases in the second stage of the experiment affects behavior in the double-auction stage. In particular, we test Hypothesis 8, which states that subjects will trade the same number of units across all treatments. Rejection of Hypothesis 8 will provide support for either a rebound effect or a saliency effect. Recall that the existence of a rebound effect would predict that more trades will occur in the HIGH and LOW treatments compared to the BASELINE. In contrast, the presence of a saliency effect would predict a lower number of trades in the HIGH treatment (especially in the later periods) compared to the BASELINE, since the introduction of the damage offsets increases the saliency of negative externalities. This effect is not present in the LOW treatment, since trading at the competitive levels and then offsetting the total damage is efficient.

We again investigate the average treatment effects first. The average number of trades is given in Table 5. We use session averages as our (independent) observations. A Wilcoxon rank-sum test indicates that there are no differences in the number of trades across treatments in Part 1 ($p-values$ are larger than 0.50).[29] This result is anticipated, since all treatments were identical in Part 1. The important question is, are there differences across treatments in Part 2?

Table 5: The Average Number of Trades in Parts 1 and 2

|	Average # of trades in Part 1	Average # of trades in Part 2
BASELINE	24.00	25.27
	(0.31)	(0.43)
HIGH	24.20	25.17
	(0.80)	(0.48)
LOW	24.27	24.83
	(0.53)	(0.35)
Standard errors in parentheses. 3 independent observations per cell.		

---

and local charities as a strong predictor of selection into for a green program (PG&E's ClimateSmart Program for carbon offsets).

[29]Throughout the paper, unless otherwise noted, all reported tests are two-sided.

In all treatments the mean number of trades is higher in Part 2. This appears to be the case because the learning process about trading in the double auction has not converged in Part 1 (see Figure 3). The slight increase in trading volume, however, is significant only for the BASELINE treatment ($p-value = 0.05$), but not for the other two treatments ($p = 0.27$ for HIGH and $p-value = 0.83$ for LOW). While we observe a lower number of trades in the HIGH treatment relative to the BASELINE and even lower number of trades in the LOW treatment, the number of trades is not significantly different from 25 in any of the treatments (Wilcoxon signed-rank tests, $p-values$ are larger than 0.59). More importantly, a pairwise comparison of treatments in Part 2 shows that the differences in the number of trades across treatments are not significant (all $p-values$ are larger than 0.51). As such, we fail to reject Hypothesis 8.

Figure 3: Number of Trades per Period
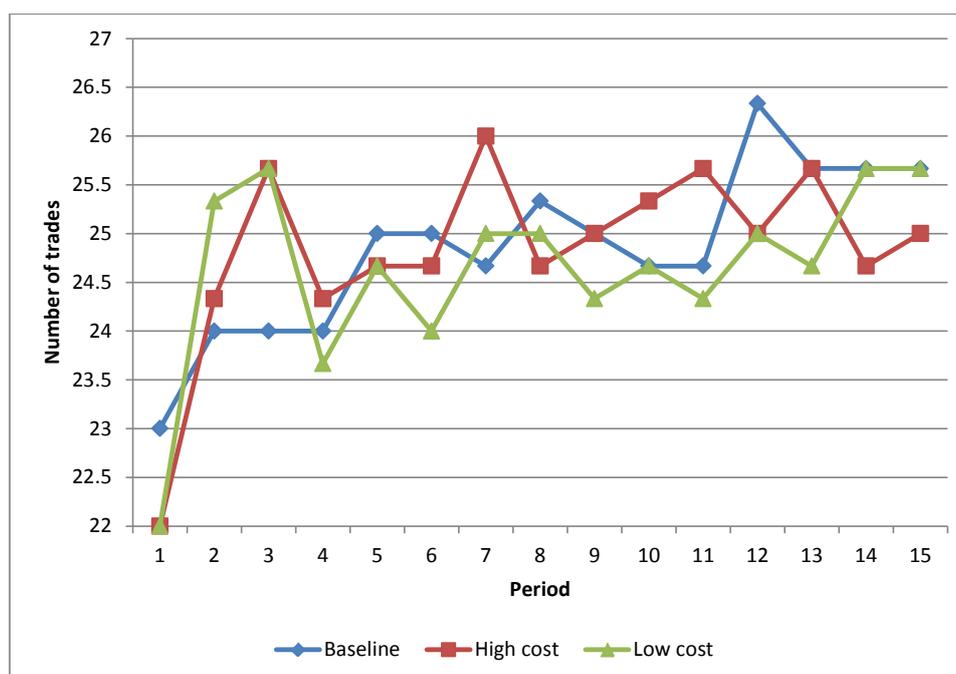
Before providing a regression analysis to strengthen these results, we first provide a quick look at the effect of treatments on trading prices. Figure 4 shows the average price of a traded unit per period for each treatment. We find that prices increase from Part 1 to Part 2, although the effect is only statistically significant for the BASELINE (Mann-Whitney tests, $p-value = 0.05$ for the BASELINE, $p-value = 0.13$ for the HIGH, and

29

$p-value = 0.28$ for LOW). More importantly, pairwise comparisons confirm that prices are not different across treatments (rank-sum tests; all $p-values$ are larger than 0.512 for part 1 and all $p-values$ are larger than 0.275 for part 2).

Figure 4: Mean Price per Period
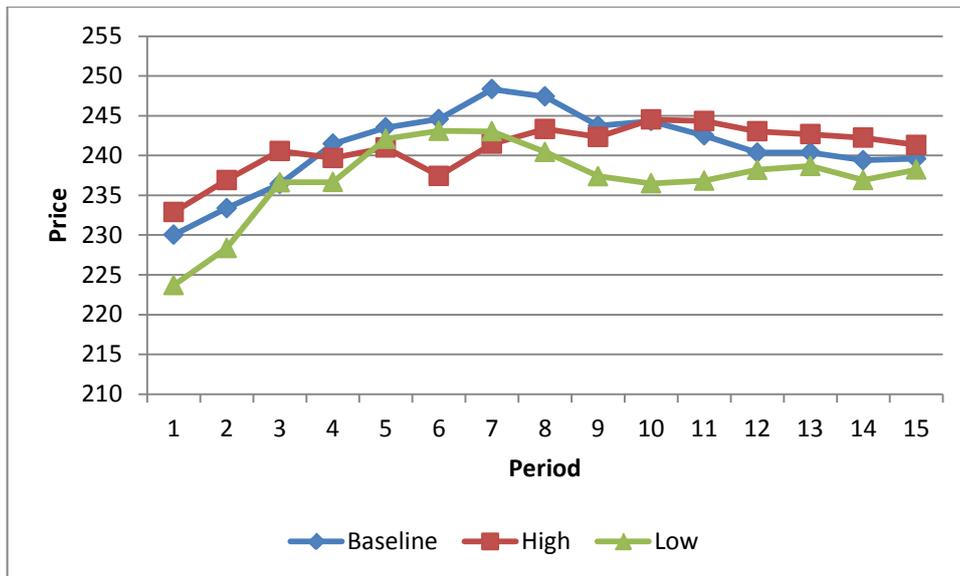
The analysis so far suggests that, on average, there is no impact on the trading period whatsoever when damage offsets are introduced–neither in the form of norm activation nor in the form of a rebound effect.[30] We now attempt to look more carefully into trading decisions using additional regression analyses incorporating control variables. First, trades would obviously be affected by the achievable surplus. For example, a unit that has a production cost of 100 is more likely to be produced (sold) compared to a unit that has a cost of 300. We create a variable, *GainMargin*, that gives the difference between value (or cost for a seller) and the competitive equilibrium price.[31] We also create a dummy variable, *Traded*, as our dependent variable and run Logit regressions. Specification 1 of Table 6 shows

---

[30]We cannot rule out the possibility that some subjects are affected by a saliency effect while others are affected by a rebound effect, but on average these two effects cancel each other out. In our between-subjects analysis, we compare the second stages of all treatments with each other to control for the re-start and/or order effects. While one could possibly look at trading behavior before and after the introduction of carbon offsets for a given individual (and test Hypothesis 8 at the individual level), this would suffer from order effects.

[31]Here, the competitive equilibrium price is only for normalization. The important point is that the units are ranked according to their values/costs.

that valuations/costs are important. In addition, the odds of a unit being traded increases significantly over time. Specification 2 shows that there is no aggregate effect of introducing the possibility to offset damages (*i.e.*, the coefficients of LOW and HIGH are not significant) and confirms results from the nonparametric tests.[32]

In specification 3, there is weak evidence that some impact on the trading period may be generated from the introduction of damage offsets: females trade less in Stage 1 (at the 10% level). We have also seen that females buy fewer damage offsets. It is, however, unclear from the regressions whether there really is a causal link between the two observations. Furthermore, we also see that the index of *SocialPolicy* is significant at the 10 percent level.

Note that individuals should only contribute to the public good of damage offsets in stage 2 if there is some component of their preference structure that values damage reduction (either in proportion to damages personally caused or damages generally). This does not generally mean that subjects should reduce trading in stage 1 if the possibility of buying damages offsets triggers awareness of the external effect of trading. Only in the HIGH treatment is it privately (and socially) more efficient to refrain from trading the last units if damage offset purchases are anticipated.

An outcome at the competitive equilibrium in the first stage and the purchase of damage offsets in the second stage observed here is indicative of some form of mental accounting: individuals cannot think about trading and damage avoidance at the same time in the double auction. This can lead to inefficiency because the cost of ex-post damage reduction is greater in the HIGH treatment than for ex-ante damage avoidance through foregone trade (at least for the last units transacted at a competitive equilibrium). Subjects do trade at competitive equilibrium levels and then try to offset the damages in the early periods of Part 2 in the HIGH treatment. This damage offset demand disappears over the experimental periods but is not compensated by reductions in trading in the trading period.[33]

Finally, we also investigate the determinants of the *number* of trades using OLS analyses

---

[32] As shown in Section 3, the socially optimal number of trades differs across treatments. Our regression results do not change when we control for efficiency.

[33] When we consider only the last five periods, we still find the number of trades in the HIGH treatment to be not significantly different from 25 ($p - value = 0.59$). Moreover, there are no statistically significant differences between the number of trades in the BASELINE and the HIGH treatments in the last five periods either ($p - value = 0.38$), which is not consistent with a saliency effect.

Table 6: Determinants of Trade

Dependent var = *Traded*	(1)	(2)	(3)
GainMargin	0.18***	0.18***	0.19***
	(0.02)	(0.02)	(0.02)
Period	0.04**	0.04**	0.05**
	(0.02)	(0.02)	(0.02)
Low		-0.25	-0.21
		(0.38)	(0.37)
High		-0.06	-0.1
		(0.32)	(0.30)
Seller			0.01
			(0.29)
Female			-0.52*
			(0.27)
Age			-0.02
			(0.09)
MajorEcon			-0.1
			(0.23)
FamilyIncome			0.16
			(0.14)
Conservative			0.04
			(0.47)
Liberal			-0.12
			(0.32)
Religion			0.01
			(0.13)
Responsibility			-0.05
			(0.09)
GiveHomeless			0.04
			(0.10)
SocialPolicy			-0.15*
			(0.09)
Unemp_vs_Env			0.16
			(0.18)
Trust			-0.08
			(0.15)
Constant	-0.57***	-0.47	0.81
	(0.16)	(0.29)	(1.70)
Observations	5,400	5,400	5,400

Standard errors are clustered at the individual level. Robust standard errors in parentheses.
*** denotes significance at 1 percent, ** denotes significance at 5 percent, * denotes significance at 10 percent

(not shown).[34] We do not find any statistically significant effect of LOW and HIGH treatments on the number of trades ($p-values$ are larger than or equal to 0.50). We also do not find any correlation between offset purchases and number of trades ($p-values$ are larger than or equal to 0.28). Regarding the control variables, the only statistically significant control is being a female. Being a female decreases the number of trades by 0.14 (which is significant at the 10% level).

7 Discussion

What do these results suggest about the possible effectiveness of voluntary damage offset programs? In the BASELINE treatment subjects earned 162.7 tokens on average per period (averaged over 360 data points). Subjects earned 4.9 tokens more in the HIGH treatment but this is not statistically significant ($p-value = 0.45$). In contrast, subjects earn 33.6 tokens more in the LOW treatment compared to the BASELINE treatment and this difference is significant at the 99% confidence level ($p-value = 0.00$). This is an economically significant 20.7% increase relative to the BASELINE treatment.

While our results suggest that the total damage done to the environment ex-ante (via trades) does not change with the presence of damage offsets, individuals persistently reach higher payoff levels with the presence of offset markets when offset technology provides a reasonably cheap way to offset damages. In contrast to other public goods games, this difference in motivation also appears to lead to much more persistent contributions to the environmental public good. It is clearly far less effective than a general Pigouvian tax (as Plott (1983) has shown), but such campaigns can still have a substantial effect where it is politically difficult to implement measures like a general carbon tax.

Though the introduction of damage offsets in our experiment led to improved earnings in the LOW treatment, it is not clear how this translates into efficiency.[35] Subjects in the HIGH and LOW treatments have an opportunity to increase their earnings to a great extent due to

---

[34]Our results are robust to different specifications, such as ordered logit analysis.

[35]We thank an anonymous referee and Charles Noussair for pointing this out and encouraging us to investigate further.

the possibility of offsetting, yet they buy very little in the HIGH treatment and only about 38% of the damages are offset in the LOW treatment. The earnings of subjects could have been much higher. This suggests that overall efficiency in the HIGH and LOW treatments might actually be lower–not higher–than the efficiency in BASELINE treatment.

One common way to measure efficiency is to calculate the ratio of the actual payoffs to the maximum possible payoffs.[36] Table 7 presents mean actual total payoffs realized (relative to no trading), maximum total payoff levels possible (relative to no trading), as well as computed mean efficiency measures for each treatment. Results confirm that higher payoffs in the LOW treatment do not translate into higher efficiency levels. In fact, efficiencies drop sharply in the HIGH and LOW treatments compared to the BASELINE. Note that these declines in efficiency do not come from different behavior in stage 1, but are rather driven by stage 2.[37] The fact that the offsetting opportunity offers large increases in the payoffs, yet is not taken up by subjects, drops the efficiencies as shown in Table 7. The interpretation of this efficiency measure should be undertaken with caution. This standard measure of efficiency does not take into account the possible welfare gains due to contributions toward damage reduction. It is not unrealistic to think that responsibility driven individuals (possibly due to warm-glow motives) and altruistic individuals enjoy positive utility when they purchase damage offsets, yet these welfare gains are not captured by this measure. In a way, this standard measure of efficiency might be giving us a lower-bound of actual efficiency levels for the HIGH and LOW treatments.

8   Conclusion

The State of Voluntary Carbon Markets 2017 report from Forest Trends' Ecosystem Marketplace shows that in 2016 offset demand fell short of supply. Ecosystem Marketplace Senior

---

[36]There are three possible ways of losing efficiency. First, the wrong number of units might be traded in stage 1. Second, even if the correct number of units are traded, the wrong individual units might be traded (*i.e.*, high cost or low value units are traded). Third, the trading market in stage 1 might function efficiently but the offset market might not. See Suter et al. (2008) for an excellent investigation of efficiency in the context of ambient tax mechanism.

[37]In fact, if one ignores stage 2 and computes efficiency in stage 1, the mean efficiency levels of stage 1 in the HIGH and LOW treatments would both be equal to 0.81 with standard deviations 0.06 and 0.03, respectively.

Table 7: Efficiency Comparison

	BASELINE	HIGH	LOW
Mean Realized (Net) Earnings	752.93 (51.68)	811.78 (64.23)	1156.61 (97.49)
Maximum Possible (Net) Earnings	954	1430	1800
Mean Efficiency	0.79 (0.05)	0.57 (0.04)	0.64 (0.05)
Standard deviations are in parentheses. Realized earnings and maximum earnings presented in this table are net earnings (earnings – status quo), where status quo earnings are determined based on no trades. There are 30 observations (10 per session) for each cell in the first and the third rows.			

Associate Kelley Hamrick said, "Lagging demand continues to be the bottleneck that prevents project developers from expanding innovative activities to support both the climate and a host of other sustainable development benefits." Our results also show that individuals could benefit from offsetting dramatically, but, unfortunately, the free-riding incentives that are associated with offsetting cause individuals not to buy enough offsets which leads to low efficiency levels.

Our paper demonstrates the importance of the connection between carbon offset purchases and personal responsibility. We find that subjects with a high personal-responsibility index care about the damages caused by their own trading, and buy more offsets as their own damages increase. This has important policy implications for carbon markets. Carbon offset programs should link the offset purchases directly to own carbon emissions producing activities, to the extent possible.[38] But not all subjects appear to be motivated by personal responsibility. Individuals with lower personal-responsibility indices do not meaningfully alter their offset purchases in response to changes in their own damages. Rather, they tend to increase their offset purchases as total damages increase. Further research could explore whether appeals to personal responsibility reduce the motivation for giving by such individuals and whether there are trade-offs for damages contributions that target different sources of individual motivation.

---

[38]The literature on the topic, for example, Vandenbergh and Steinemann, 2007, emphasizes that the personal responsibility for reducing own damages could improve the behavioral demand for carbon offsets.

References

[1] Andreoni, James (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. The Journal of Political Economy, 97(6): 1447-1458.

[2] Andreoni, James (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. The Economic Journal, 100: 464-477.

bibitemandreoni95 Andreoni, James (1995) Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. Quarterly Journal of Economics, 121.

[3] Andreoni, James, and John Miller (2002). Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. Econometrica, 70 (2): 737-753.

[4] Arana, J. E. and C. J. Leon (2013) Can Defaults Save the Climate? Evidence from a Field Experimetn on Carbon Offsetting Programs, Environmental and Resouce Economics, 54: 613-626.

[5] Blasch, J. and M. Farsi (2014) Context effects and heterogeneity in voluntary carbon offsetting–a choice experiment in Switzerland, Journal of Environmental Economics and Policy, Vol.3, Issue 1, Pages 1-24.

[6] Blasch, J. and Markus Ohndorf (2015) ALtruism, moral norms and social approval: Joint determinants of offset behavior. Ecological Economics, 116: 251-260.

[7] Bolton, G. E., and A. Ockenfels (2000): A Theory of Equity, Reciprocity and Competition, American Economic Review, Vol. 90, 166-193.

[8] Brouwer, R., L. Brander and P. Van Beukering (2008) "A convenient truth": air travel passengers' willingness to pay to offset their CO_2 emissions. Climatic Change, 90: 299-313.

[9] Capoor, K. and Ambrosi, P. 2009 State and Trends of the Carbon Market 2009 World Bank, Washington DC.

[10] Cason T. 2010 What Can Laboratory Experiments Teach Us About Emissions Permit Market Design? Agricultural and Resource Economics Review, Vol. 39, No. 2, pp. 151-161.

[11] Charness, Gary, and Matthew Rabin (2002), Understanding social preferences with simple tests. Quarterly Journal of Economics, 117: 81769.

[12] Chaudhuri, A. 2011. Sustaining Cooperation in Laboratory Public Good Experiments: A Selective Survey of the Literature. Experimental Economics, 14:47-83.

[13] Conte, M. N. and M. J. Kotchen (2010) Explaining the price of voluntary carbon offsets, Climate Change Economics, Vol. 1, No. 2, 93-111.

[14] Diederich, J. and T. Goeschl (2014) Willingness to Pay for Voluntary Climate Action and Its Determinants: Field-Experimental Evidence, Environmental Resouce Economics, 57, 405-429.

[15] Elster, Jon (1989), Social Norms and Economic Theory, The Journal of Economic Perspectives, Vol. 3, No. 4, pp. 99-117

[16] Ehmke, M. and J. F. Shogren. 2008. Experimental Methods for Environment and Development Economics, Environment and Development Economics, Vol. 00, pp. 1-38.

[17] Falk, A. and N. Szech (2013) Morals and Markets, Science, 340, 707-711.

[18] Fehr E., and Schmidt K.M. (1999), A Theory Of Fairness, Competition, and Cooperation, Quarterly Journal of Economics, Volume 114, Number 3, 817-868.

[19] Fischbacher, Urs. 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments, *Experimental Economics* 10(2): 171–178.

[20] Friesen, L. and Lata Gangadharan, Environmental Markets: What Do We Learn From the Lab? Journal of Economic Surveys, Vol. 27, No. 3, pp 515-535, 2013.

[21] Gillingham, Kenneth, Matthew Kotchen, David Rapson, and Gernot Wagner. 2013. The Rebound Effect is Over-played. Nature 493: 475-476.

[22] Grether, David M., R. Mark Isaac, Charles R. Plott (1981) " The Allocation of Landing Rights by Unanimity Among Competitors" The American Economic Review, Vol. 71, No. 2, Papers and Proceedings of the Ninety-Third Annual Meeting of the American Economic Association (May, 1981), pp. 166-171

[23] Godby, Robert W., Stuart Mestelman, R. Andrew Muller, and J. Douglas Welland (1997) Emissions Trading with Shares and Coupons when Control over Discharges Is Uncertain, Journal of Environmental Economics and Management, 32, 359-381

[24] Hamrick, Kelley and Melissa Gallant (2017) Unlocking Potential, State of the Voluntary Carbon Markets, Forest Trends Ecosystem Marketplace.

[25] Harding M. and D. Rapson (2014) Do Voluntary Carbon Offsets Induce Energy Rebound? A Conservationist's Dilemma, working paper.

[26] Harrison, Glenn, Elizabeth Hoffman, E.E. Rutstrom, Matthew L. Spitzer (1987) Coasian Solutions to the Exteernality Problem in Experimental Markets, Vol. 97, No. 386, 388-402

[27] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. AEA Papers and Proceedings 91(2), 73-78.

[28] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich., N.S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F.W., Patton, J.Q., Tracer, D., (2005). Economic Man in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. Behavioral and Brain Sciences 28, 795-855.

[29] Jacobsen, G.D., Kotchen, M.J. and M.P. Vandenbergh (2012) The Behavioral Response to Voluntary Provision of an Environmental Public Good: Evidence from Residential Electricity Demand. European Economic Review, 56, 946-960.

[30] Kotchen, M. J. and M. R. Moore (2008) "Conservation: From Voluntary Restraint to a Voluntary Price Premium" Environmental Resource Economics, 40: 195-215.

[31] Kotchen, M. J. (2009a) "Offsetting Green Guilt" Stanford Social Innovation Review, 7(2, 26-31.

[32] Kotchen, M. J. (2009b) "Voluntary Provision of Public Goods for Bads: A Theory of Environmental Offsets" Economic Journal, 119 (537), 883-899.

[33] Kube, S., M. Marechal and C. Puppe. 2012 The Currency of Reciprocity: Gift Exchange in the Workplace. American Economic Review, 102, 1644-1662.

[34] On the interrelation between the consumptionof impure public goods and the provision of direct donations: Theory and empirical evidence. Resource and Energy Economics, 47: 72-88.

[35] Lange, A. and A. Ziegler (2017) Offsetting Versus Mitigation Activities to Reduce CO_2 Emissions: A Theoretical and Empirical Analysis for the U.S. and Germany. Environmental Resource Economics, 66: 113-133.

[36] Ledyard, J. O. and K. Szakaly-Moore (1994), Designing organizations for trading pollution rights, J. Economic Behavior and Organization. 25, 167-196.

[37] Ledyard, John (1995) Public Goods: A Survey of Experimental Research, in The Handbook of Experimental Economics. Roth and Kagel, eds. Princeton, NJ: Princeton University Press.

[38] Lofgren A., Martinsson P, Hennlock M, and Sterner T (2012) Are experienced people affected by a pre-set default option–Results from a field experiment. Journal of Environmental Economics and Management, 63 (1): 66-72.

[39] Loschel, A., B. Sturm and C. Vogt (2013) The demand for climate protection–Empirical evidence from Germany, Economics Letters, 118, 415-418.

[40] Revealed preferences for voluntary climate change mitigation when the purely individual perspective is relaxed–evidence from a field experiment, Journal of Behavioral and Experimental Economics, 67, 149-160.

[41] Stuart Mestelman, 2000. "Environmental Policy: Lessons from the Laboratory," McMaster Experimental Economics Laboratory Publications 2000-01, McMaster University.

[42] Suter, J. F., C. A. Vossler, G. L. Poe and K. Segerson (2008) "Experiments on Damage-Based Ambient Taxes for Nonpoint Source Polluters" American Journal of Agricultural Economics, 90(1): 86-102.

[43] Muller R. Andrew and Stuart Mestelman (1994) Emission Trading with Shares and Coupons: A Laboratory Experiment, The Energy Journal, 15(2), 195-211

[44] Noussair, Charles N. and Plott, Charles R. and Riezman, Raymond G. (1995) An Experimental Investigation of the Patterns of International Trade. American Economic Review, 85 (3). pp. 462-491.

[45] Noussair, Charles N. and Daan P. van Soest. 2014. Economic Experiments and Environmental Policy, Annual Review Resource Economics, Vol. 6, pp. 319-37.

[46] Peters-Stanley and Yin (2013), Maneuvering the Mosaic, State of the Voluntary Carbon Markets

[47] Plott, Charles (1983) Externalities and Corrective Policies in Experimental Markets, The Economic Journal, Vol. 93, No. 369, 106-127

[48] Plott, Charles and Hong (1982) Rate Filling Policies for Inland Water Transportation: An Experimental Approach. Bell Journal of Economics, 13, 1-19.

[49] Sandel, M. 2012. What Money Can't Buy. New York: Farrar, Strausand Giroux.

[50] Claudia Schwirplies and Andreas Ziegler (2016) Offset carbon emissions or pay a price premium for avoiding them? A cross-country analysis of motives for climate protection activities, Applied Economics, 48:9, 746-758.

[51] Sturm, B. and Joachim Weimann, Experiments in Environmental Economics and Some Close Relatives, Journal of Economic Surveys, Vol. 20, No. 3, 2006

[52] Vandenbergh, Michael P. and Anne Steinemann (2007), "The Carbon-Neutral Individual, New York University Law Review, 82, pp. 1673 " 1745.

Appendix A

Table 8: Tobit Regressions

Dependent var = DamageOffsets	ALL	ALL	LOW	LOW	HIGH	HIGH
Surplus	0.4	0.52	0.48	0.6	0.21	0.32
	(0.30)	(0.31)	(0.43)	(0.42)	(0.19)	(0.21)
Resp*Surplus	-0.04	-0.05	-0.05	-0.06	-0.02	-0.03
	(0.03)	(0.03)	(0.05)	(0.04)	(0.02)	(0.02)
DamagesCaused	-0.76**	-0.62*	-1.02*	-0.96**	-0.5	-0.42
	(0.37)	(0.34)	(0.52)	(0.41)	(0.38)	(0.35)
Resp*DamageC	0.09**	0.07*	0.12**	0.11**	0.05	0.04
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
TotalDamages	0.58**	0.53**	0.96**	0.84***	0.19	0.05
	(0.23)	(0.22)	(0.44)	(0.29)	(0.16)	(0.15)
Resp*TotalD	-0.06**	-0.05**	-0.09*	-0.07**	-0.02	-0.01
	(0.03)	(0.02)	(0.05)	(0.03)	(0.02)	(0.01)
Responsibility	59.67**	54.53*	93.65	75.33*	21.66	5.27
	(29.11)	(27.91)	(58.57)	(39.53)	(19.87)	(17.77)
Period	-1.55**	-1.56*	-1.15	-1.31	-2.15***	-2.28***
	(0.79)	(0.79)	(1.23)	(1.29)	(0.79)	(0.85)
Seller		3.32		1.27		6.01
		(7.28)		(10.36)		(6.34)
Female		-6		-19.35*		3.96
		(7.37)		(10.40)		(6.49)
Age		2.57		1.93		2.84*
		(2.07)		(3.22)		(1.66)
MajorEcon		-23.71**		-39.98*		-10.44
		(11.17)		(23.99)		(9.60)
FamilyIncome		-1.79		-2.75		-3.59
		(2.89)		(4.05)		(3.31)
Conservative		-3.5		-6.8		-15.48
		(10.29)		(11.48)		(13.02)
Liberal		-7.01		-14.27		-0.16
		(8.69)		(14.84)		(6.69)
Religion		-5.34		-7.02		-3.27
		(3.38)		(5.02)		(2.53)
GiveHomeless		4.29		5.14		4.46**
		(3.06)		(4.38)		(1.97)
SocialPolicy		2.1		4.55		-0.83
		(3.42)		(5.17)		(2.29)
Unemp_vs_Env		-1.71		-5.12		-1.83
		(5.11)		(7.70)		(2.59)
Trust		3.74		4.07		-1.75
		(3.61)		(5.23)		(3.07)
Low	52.82***	54.25***				
	(7.91)	(7.27)				
Constant	-617.47**	-652.84**	-976.08*	-898.31**	-159.76	-62.57
	(265.22)	(256.56)	(503.33)	(361.25)	(185.66)	(166.43)
Observations	720	720	360	360	360	360

Standard errors are clustered at the individual level.
Robust standard errors in parentheses.
*** denotes significance at 1 percent, ** denotes significance at 5 percent, * denotes significance at 10 percent

Table 9: Tobit Regressions for Different Responsibility Levels

Dependent var = *DamageOffsets*	Low Responsibility		Average Responsibility		High Responsibility	
	(1)	(2)	(3)	(4)	(5)	(6)
Surplus	0.08	0.20**	-0.03	0.10*	0	0.04
	(0.11)	(0.09)	(0.07)	(0.05)	(0.08)	(0.05)
DamagesCaused	-0.16	-0.09	-0.04	0	0.29***	0.27***
	(0.13)	(0.09)	(0.10)	(0.08)	(0.08)	(0.08)
TotalDamages	0.23**	0.24***	-0.09	-0.05	-0.01	0.09
	(0.09)	(0.06)	(0.08)	(0.06)	(0.08)	(0.07)
Period	-1.71	-1.58	-0.14	-0.47	-2.31	-2.76*
	(1.58)	(1.54)	(0.76)	(0.81)	(1.56)	(1.61)
Low	67.17***	69.52***	28.99**	31.40***	62.48***	59.85***
	(13.60)	(11.66)	(12.29)	(6.77)	(11.75)	(10.95)
Constant	-274.73**	-310.75***	116.8	-205.87	-46.43	-115.23
	(108.66)	(101.84)	(101.21)	(153.73)	(82.26)	(73.31)
Controls added	No	Yes	No	Yes	No	Yes
Observations	290	290	220	220	210	210

Standard errors are clustered at the individual level.
Robust standard errors in parentheses.
*** denotes significance at 1 percent, ** denotes significance at 5 percent, * denotes significance at 10 percent

Appendix B

In this section we briefly investigate the damages done by subjects net of their offset purchases. We define the variable *NetDamages* as the difference between *DamagesCaused*/2 and the offset purchases of a given subject in a given period. The variable *NetDamages*, therefore, gives a measure of whether a subject completely offsets his/her share of the damages, more than offsets or under offsets, assuming an equal share of damages caused by the trades of subjects among buyers and sellers. Table 10 shows the number of people in each treatment that belong to these different categories, as well as the number of people who never offset throughout the experiment.

Table 10: Number of subjects corresponding to different levels of net damage

	HIGH	LOW
$NetDamages \leq -1$	-	1
$-1 < NetDamages \leq 1$	-	1
$1 < NetDamages \leq 50$	1	12
$50 < NetDamages$	27	20
$Never\ Offsets$	8	2
The categories are not overlapping. In particular, individuals that never offset their damages are a separate category than the individuals with net damages greater than 50.		

To study the correlates of *NetDamages*, we run an additional regression analyses (see Table 11). It is surprising to observe that female subjects have higher net damages. Our analysis in Section 6 shows that female trade less. Perhaps female subjects cut down their trade volume and did not buy as many offsets. Unfortunately, Table 11 shows they did not cut their trades enough (their net damages is still above male subjects). In addition, we see being an economics major is marginally significant in the LOW treatment, and being conservative (over being moderate) is significant in the HIGH treatment.[39]

---

[39]Note that the *Responsibility* variable is not significant in these regressions. This does not conflict with our previous results since these regressions do not control for either own damages caused or its interaction with the responsibility variable.

Table 11: OLS Regressions for Net Damages

Dependent var = *NetDamages*	ALL	ALL	LOW	LOW	HIGH	HIGH
	(1)	(2)	(3)	(4)	(5)	(6)
Surplus	0.11***	0.13***	0.08	0.1	0.17***	0.18***
	(0.04)	(0.03)	(0.07)	(0.06)	(0.03)	(0.03)
TotalDamages	0.03	0.03	-0.05	-0.12	0.09***	0.10***
	(0.04)	(0.04)	(0.09)	(0.09)	(0.03)	(0.03)
Period	0.66	0.67	0.83	1.19	1.00**	1.01**
	(0.47)	(0.47)	(1.01)	(1.08)	(0.38)	(0.38)
Seller		5.64		6.17		4.33
		(5.60)		(8.68)		(5.53)
Female		4.31		18.98**		-1.38
		(5.96)		(9.09)		(3.98)
Age		-0.01		-0.72		0.39
		(1.85)		(3.08)		(0.90)
MajorEcon		9.9		45.26*		-3.61
		(8.71)		(24.18)		(4.67)
FamilyIncome		2.6		7.11		2.76
		(2.47)		(4.22)		(1.67)
Conservative		2.38		8.23		13.53**
		(8.15)		(10.65)		(6.55)
Liberal		6.56		15.28		5.05
		(7.11)		(14.50)		(5.67)
Religion		3.07		5.74		1.24
		(2.63)		(5.07)		(1.46)
Responsibility		0.92		0.19		0.6
		(2.14)		(3.03)		(1.42)
GiveHomeless		0.13		-2.82		-0.85
		(2.47)		(4.56)		(1.88)
SocialPolicy		-1.94		-4.22		0.17
		(2.53)		(4.46)		(1.44)
Unemp_vs_Env		2.04		6.63		-1.33
		(3.99)		(6.58)		(1.58)
Trust		-1.83		-1.73		-0.82
		(2.67)		(4.49)		(1.53)
Low	-32.41***	-32.37***				
	(5.50)	(5.39)				
Constant	31.91	12.13	106.27	148.45	-49.24	-81.36*
	(43.46)	(61.69)	(99.14)	(119.76)	(34.50)	(40.55)
Observations	720	720	360	360	360	360

Standard errors are clustered at the individual level.
Robust standard errors in parentheses.
*** denotes significance at 1 percent, ** denotes significance at 5 percent, * denotes significance at 10 percent