



ELSEVIER

Contents lists available at ScienceDirect

Resource and Energy Economics

journal homepage: www.elsevier.com/locate/ree

Buying the right to do wrong – An experimental test of moral objections to trading emission permits

Ragnhild Haugli Braaten^a, Kjell Arne Brekke^{b,*}, Ole Rogeberg^c

^a Oslo Economics, Norway

^b Department of Economics, University of Oslo, Norway

^c The Ragnar Frisch Centre for Economic Research, Oslo, Norway

ARTICLE INFO

Article history:

Received 22 May 2013

Received in revised form 8 June 2015

Accepted 12 July 2015

Available online 20 July 2015

JEL classification:

C9

Q3

Keywords:

Tradable pollution rights

Experiment

Morality

ABSTRACT

Moral objections to quota trade are common in public discussion. The underlying moral reasoning is often hard to extract from the public's rhetoric, and economists often claim that the objections are due to misunderstandings. In this paper we test experimentally one possible objection to trading emission rights: that selling pollution rights is objectionable because it involves trade in the right to do something wrong. The experiment is phrased in a neutral language to avoid invoking subjects' prior attitudes to trading pollution permits. In the experiment, subjects earn money from using stickers that inflict losses on others. After the initial rounds, the use of stickers is limited and the rights to use stickers are tradable. We observe no reluctance to trade these rights compared to a control treatment with identical incentives but no externalities. In a final stage, subjects vote on whether stickers should be tradable or not. Subjects in both treatments are almost unanimous in the support to trade. A majority of third party subjects asked to judge the experimental procedures, however, indicated that the market made the use of stickers seem more moral. A post experiment survey showed that our subject pool had a negative attitude to real life trade of emission quota. The experiment finds no indication that this attitude is due to a concern over tradable quotas legitimizing morally objectionable actions.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

* Corresponding author. Tel.: +47 48295384; fax: +47 22 855035.

E-mail addresses: rbr@osloeconomics.no (R.H. Braaten), k.a.brekke@econ.uio.no (K.A. Brekke), ole.rogeberg@frisch.uio.no (O. Rogeberg).

1. Introduction

When the “environmental revolution” arrived in the 1960s, economists were ready and waiting.

- Opening statement of “The theory of environmental policy” (Baumol and Oates, 1988)

The Times was flooded with scathing letters — mostly from economists, some of them my Harvard colleagues. I failed to understand the virtue of markets, they suggested, or the efficiencies of trade, or the elementary principles of economic rationality.

-Philosopher Michael Sandel on reactions to his moral arguments against tradable emission quotas (Sandel, 2012)

Economists were quick to suggest environmental taxes and tradable emission quotas when the modern environmental movement arrived, but, as Baumol and Oates (1988) pointed out, these solutions were not well received.¹ Command and control were preferred to economic instruments like quotas and taxes, and while some schemes involving tradable emission quotas have been implemented recently, legal instruments still dominate. There was, and still is, a considerable resistance toward both taxes and tradable quotas, often laced with moral indignation. This is especially so for tradable emission quotas.

Why do many non-economists object so strongly to an idea that economists typically support? Do non-economists fail to understand the basic idea behind quota trade, or has economic training somehow made economists blind to moral issues that others find obvious? To examine this, we designed an experiment where participants could trade the rights to do wrong. The wrong involved using ‘stickers’ that brought gain to the user and harm to the other group members, with net harm overall. We examined whether participants found trade in the rights to use stickers objectionable both before and after experience with trade in a double-auction trading market. The design did not invoke a climate/environmental context, and the use of terms such as “quotas” or “emissions” was avoided so as not to trigger prior convictions on the morality of emission trading schemes.

Skepticism toward tradable emission quotas seems to be widespread. Asked in 2010 whether “Norway ought to take its share of responsibility for greenhouse gas emissions by buying emission quotas from other countries,” a nationally representative survey of the Norwegian population found 46% disagreeing with the statement.² This negative attitude to tradable emissions quotas is also apparent in official Norwegian policy. Following an intense political debate on the morality of buying quotas abroad, the final policies emphasized targets for domestic abatements at the expense of cost efficiency (Eyckmans and Kverndokk, 2010).

One explanation for this skepticism may be that politicians and the public fail to understand how a tradable emission would work and what it would achieve. Alternatively, it may be economists that are unable to see a moral argument that non-economists emphasize. Past experimental results have shown that economics students tend to act more selfishly than other students. This may be due to selection (Frey and Meier, 2005), but it may also reflect the effects of their training: economists are typically taught to focus on outcomes, leading to an emphasis on total costs and emissions in our context. If policy A yields better consequences for everyone than policy B, economists will tend to see A as the best option, even if this policy “legitimizes” something others feel should be stigmatized.

The view that markets legitimize objectionable actions appears in many contexts. We have observed it in talk shows, cartoons, and discussions with non-economists. A typical argument is to state or suggest that the logic of quota trading would extend to quotas for any number of bad or illegal activities. The satirical website *cheatneutral.com* claims to “highlight some of the issues surrounding carbon offsetting” by providing people who cheat on their spouses the opportunity to pay others to refrain from doing so – thus offsetting the sin. By “funding someone else to be faithful and NOT cheat [t]his neutralizes the pain and unhappy emotion and leaves you with a clear conscience”. Similarly, a

¹ This is not the only issue where economists and non-economists tend to disagree, see Sapienza and Zingales (2013).

² From “The large Norwegian climate- and environmental survey 2010” conducted by Ipsos MMI and CICERO Centre for Climate Studies (“Den store norske klima- og miljøundersøkelsen 2010”).

political comic strip by Ken Fisher satirized the Clean Air Act in 1992 by showing a homeowner (Ryder) surprising a burglar in his home. Using a telephone exchange called “crime brokers” he sells his right to kill the intruder to the brokers, who sell it on to a criminal (Sluggo). “Thus, Ryder over complied with the law so that Sluggo could under comply. The same number of death results, but with a more efficient allocation”.

To economists, such arguments by analogy will seem like evidence of a basic misunderstanding. Common to many of the examples we have come across, however, is the belief that markets for an activity legitimize the activity and undermine moral norms. In this light, the ethical argument has two parts: one supporting the claim that an activity X is unethical, and one supporting the claim that a market in activity X would erode ethical norms against the activity.

A systematic and explicit statement of this argument comes from Harvard philosopher Michael Sandel, who first presented the argument in a 1997 *New York Times* op-ed (Sandel, 2005), restating it in his recent book (Sandel, 2012). He argues that excessive CO₂ emissions are morally objectionable, and that this morally objectionable activity is legitimized by instituting a trading scheme. This makes the trading scheme itself immoral.

There is some evidence that market-like features can influence the use and effect of informal punishments such as social sanction. In a widely cited study, Gneezy and Rustichini (2000) found that a fine for picking up children late in a kindergarten resulted in more parents picking up their children late. They argued that “a fine is a price,” and that the introduction of a fine shifted the way parents saw their behavior: keeping your child in kindergarten after closing hours went from being an ethical issue to being a service provided at a price. The introduction of the fine led to a “crowding out” of the ethical incentives. While the study does not support Sandel’s claim that prices undermine morality more than fines do, stated behavior studied in Brekke et al. (2003) indicates that people do respond differently to fines, which are seen as symbolically supporting an obligation, than they do to prices, which are seen as fair compensation for a service available for purchase. It seems well established that prices can undermine a moral obligation (see Bowles and Polania-Reyes (2012) for a recent survey of the literature).

Viewed in this light, the strong reaction from economists that Sandel describes in the introductory quote to this article may seem surprising: textbook treatments of pollution in economics see pollution as excessive if it generates more harm than benefits. This definition of excessive seems different from Sandel’s in strictness more than in kind. Crowding out, while not yet standard textbook material, is the subject of a large literature, and would be unlikely to make economists flood the *New York Times* with letters. So where is the disagreement? As we see it, the disagreement is likely about consequence ethics versus virtue ethics.

When firm A sells a permit to firm B, A will pollute less and B more, but total pollution will remain the same. Since both firms want to trade, they must both be better off. We thus have the same pollution, lower costs for both firms and no direct negative consequences. Evaluated on these outcomes trade seems unobjectionable. Sandel, on the other hand, is inspired by virtue ethics. Polluting, and hence harming others, is not consistent with a moral character. An institution that serves to legitimize pollution is therefore itself objectionable, even if pollution levels remain the same.³

Our question is not whether Sandel is right in his moral objection to quota trade, but whether a similar moral objection could help explain the widespread opposition to tradable emission quota schemes: are experimental participants opposed to trade in the right to do something morally wrong when the trade is Pareto-improving in monetary outcomes? While surveys reveal substantial moral objection toward trading of pollution rights, attitudes may be different when the participants themselves experience the direct monetary benefits of trade.

As we see it, the main moral objection against CO₂ emissions is that marginal emissions at current levels harm others in excess of what can be justified by individual gains. In our experiment, this effect was reproduced by asking lab participants to choose how many ‘stickers’ to use. Each sticker used increased their own payoff, but inflicted a monetary loss on their co-players, with the loss exceeding

³ For a challenge to the view that markets in carbon dioxide emission are objectionable from a virtue ethical perspective, see Bruni and Sugden (2013).

their private benefit⁴. The private payoff from sticker-use was asymmetric within the group so as to generate gains from trade in stickers.

In the baseline rounds participants used most of their stickers, just as people in practice generate large CO₂-emissions outside the lab. Next, we introduced limited use-rights for the stickers along with an opportunity to trade user-rights within the group. “Burning” or retiring use-rights was explicitly ruled out to ensure that total harms were kept constant. For the treatment groups, this was framed as a tradable right to use stickers. In the control treatment, this was framed as a new and separate trading scheme involving a new good that did not impose harms on anyone (details given below). Formally, the two games were equivalent, so any differences should be due to the framing: trading the right to harm others versus trading a morally neutral good. Before the trading scheme was introduced, participants in both treatment and control groups were asked whether they thought that trading should be allowed or not. After participating in 10 trading rounds, participants were given an actual vote to decide whether trading should be allowed in a final round.

Opposition to trade could be displayed in three ways: (1) by stating opposition to trade in a survey question before experiencing trade, (2) by limiting the amount of trade (and thereby foregoing earnings), and (3) by voting against trade in a final round where the majority vote determined whether or not trading would be allowed in the final part of the experiment. If participants did see a moral objection to trading the right to use stickers that harm others, we would expect them to vote against trade both in the pre-experience hypothetical vote and in the post-experience real vote. We would expect no objections to trade in the control groups with trade in a morally neutral good. We would also expect that moral objections would show up as reluctance to trade in the treatment groups, so that fewer rights would be traded in the treatment groups compared to the control groups.

The results showed no objections to trade in any of these outcome measures for either the treatment or control groups. Various explanations of the results are possible: participants may have found the harm inflicted on others insufficiently severe to see sticker use as morally objectionable. Alternatively, they might not see trade in stickers as legitimizing sticker use. To test these hypotheses we ran a survey on Amazon Mechanical Turk. These respondents had not participated in the original experiment, but were given the instructions from the experiment and asked to assess whether sticker use was morally objectionable, and whether a market in stickers would serve to legitimize such use. Respondents tended to agree on both points: sticker use was a moral issue, and the market legitimized sticker use.

Another possibility could be that our research participants viewed markets more favorably than most. This seems implausible in light of a post experimental survey on moral judgments of different markets: our subjects stated strong moral objections to emission quota trade, rating the immorality of quota trade between two firms almost as high as that of markets in human organs and sexual services.

In sum, the results from the experiment and the survey suggest that the sticker design fulfilled both premises of Sandel’s argument, but without these features triggering strong opposition to trade in stickers. This is consistent with the economists’ hypothesis that opposition to tradable emission quotas is due to a misunderstanding of quota trading systems. Alternative explanations, however, cannot be ruled out and the results should be seen as tentative: first, the “Turkers” who evaluated the morality of sticker use and markets without directly participating in the experiment were also asked control questions to test their understanding of the experimental design. Almost none were able to answer all questions correctly, and remaining misunderstandings may have influenced their moral judgments. Second, to the extent that the ethical norms are specific to the activity involved, the case for strengthening and protecting norms against “sticker use” is likely to be weaker than the case for strengthening and protecting norms against excessive CO₂ emissions. That is, our subject may share the view that CO₂ emission is an act that should be penalized, but at the same time find sticker use within the experiment too trivial to warrant stigmatizing. If this was the case, Sandel’s argument against markets that legitimize immoral acts would not apply in the specific case studied. Third,

⁴ For most types of emissions, the private benefits exceed the social cost for low levels of emissions. By standard economic arguments, emissions are socially harmful and policy interventions required only when social costs exceed private benefits. We have thus chosen to calibrate the experiment such that this is the case for all stickers used.

the experimental instructions emphasized that subjects understood the payoff consequences of their action. This may specifically trigger consequentialist moral judgment. Paxton et al. (2011) documents that working on mathematical problems increases the probability of later moral judgments being consequentialist. One third of our participants are students of math, although recruited in their first year at university, this selection of students may be particularly prone to consequentialist behavior. In conclusion, while our experiment fails to show how moral opposition to quota trade arises, it does show that such resistance is not an automatic response triggered in any market that involves trade in immoral goods.

2. Other moral objections to quota trade

To avoid misunderstanding, it is worth noting that our experiment is not intended to test whether participants have moral objections to quota-markets in general. Our aim is to test how important and widely held a very specific type of objection is. To make this point clearer it may be useful to point out some examples of alternative objections that are not tested in this study. In addition to features which are specific to quota trade as a policy instrument, there are also other ethical questions raised by such trading schemes. These concerns may involve the amount, distribution and timing of abatements, as well as the distribution of costs (see further discussion of such general ethical issues in the context of climate change in Gardiner, 2004).

While trade in emission quotas is Pareto-improving, the gains may be asymmetrically shared. If sellers of quotas are poor and the gains primarily go to rich buyers, an objection to the trade may be that it increases inequality. If the prices are sufficiently favorable to the party that is best off, theories of inequality aversion like Fehr and Schmidt (1999) would predict that participant in the lab would oppose trade. Ciccone et al. (2015) test such a trading scheme in the lab and find that participants are indeed reluctant to trade at set prices that strongly favor the rich, even if the trade is Pareto-improving in monetary outcomes. This could explain objections to emissions trading if participants think the market price will favor rich countries. It is unlikely to be relevant in our experiment, where prices are endogenous and would shift in response.

Other objections to emissions trading are more pragmatic and related to specific design details. One such sub-optimal feature is the often excessive amount of quotas issued, both in the Kyoto Protocol and within the EU ETS trading system. This makes the abatement generated by the quota system limited and insufficient, and may be a reason for opposing specific implementations of such policies. Another frequently criticized feature of the Kyoto Protocol is its Clean Development Mechanism (CDM). The CDM raises several problematic incentive effects and may in fact increase emissions rather than stabilizing them as intended (Strand and Rosendahl, 2012). Such criticism, which was also included in Sandel's 1997 op-ed, is common also amongst economists, but apply to flawed implementations of the policy rather than to the "intrinsic" morality of quota trade per se.

3. Experimental design

The experiment consisted of three stages which are outlined in Table 1. Stage 1 was a linear common pool resource game where taking from the common pool resource was framed as "using stickers." By using a sticker the participant earned a private payoff, but at the same time imposed negative externalities on the three other group members. The private earnings of sticker use was asymmetric, in order to create gains from trade in stage 2. Stage 1 of the experiment was played for 10 rounds and was identical for both control and treatment sessions.

In stage 2, the groups in the treatment sessions were issued a limited number of rights to use stickers. These rights were tradable and could not be retired (i.e., the "quotas" could not be "burnt"). This design was chosen to ensure that total harms could not be influenced by participants. If it had been possible to retire stickers, however, "strongly moral" participants would face new moral issues regarding trade: selling a sticker you would not otherwise use would run the risk that the buyer would use the sticker, increasing total harm. This would make sales morally problematic. Alternatively, the "strongly moral" participant might want to purchase additional stickers in order to retire these and reduce total harms further. While such motives are an interesting issue in themselves, these effects

Table 1
Overview of the stages in the experiment.

	Description	Treatment sessions	Control sessions
Stage 1	Common pool resource game Identical for both treatments (10 rounds)	Withdrawals from the common pool by using “stickers”, which are individually profitable (asymmetrically within the group) but harm other group members	
Stage 2	Trade game explained Hypothetical question: should trade be allowed?	Cap-and trade with harmful stickers: limited amount of user rights to the harmful stickers are introduced, which are open for trade within the group	Cap and trade with new, non-harmful stickers. Payoffs transformed to equal the payoffs per trade in the treatment groups with harmful stickers
Stage 3	Trade (10 rounds) Voting: should trade be allowed in the final round? Final round with endogenous rules	Cap on user rights equal to stage 2. Vote on trade with user rights for harmful stickers	Cap on non-harmful stickers equal to stage 2. Vote on trade with the non-harmful stickers.
		Post-experiment survey	

are not part of the discussion we wanted to focus on in this experiment as it is not central to virtue ethical argument.

A second benefit of the “no burning of stickers” rule was that it enabled us to create a formally equivalent control treatment without an externality frame: by manipulating the baseline endowment and payoffs, the exact same payoff structure could be framed as trade in a “new good” with no externalities.

Before stage 2 was implemented, participants were asked a hypothetical question of whether they thought trade should be allowed or not.

In stage 3, having directly experienced trading, participants voted on whether or not to allow trading of sticker user rights in a final round, and the majority vote was implemented. The different stages are outlined in detail below.

3.1. Stage 1 – establishing the morally suspect good

To avoid having prior attitudes toward environmental or other goods influence results, a non-environmental setting with a public bad, labeled stickers, was established. In groups of four, each group member initially received 10 stickers. Their task was to determine how many stickers, R_i , to use. For each sticker used, the participant would get a private payoff, a_i , while reducing the payoff of each remaining group member by 10 kroner.⁵ Participants were randomly assigned one of the two types, differing only in their private payoff from sticker use. The low productivity type, denoted as type 1 in the experiment, earned 5 kroner per sticker used. The high productivity type, denoted type 2, earned 25 kroner per sticker used. In addition to the stickers, each participant was given a baseline endowment of 150 kroner in the beginning of each round. Payoff of participant i in one round was determined as:

$$\pi_i = 150 + a_i R_i - 10 \sum_{j \neq i} R_j \quad \text{with } a_i = 5 \text{ or } 25$$

Note that participants of both types could earn negative payoffs in stage 1. This would occur if a participant used few stickers while the remaining group members used a lot. The low productivity type would also get negative payoff if everyone, including the low productivity type, used all their stickers. The instructions emphasized that while payoffs in stage 1 were potentially negative, the payoffs in later stages would be sufficiently positive that total earnings from participating in the experiment would nonetheless be positive.

⁵ The exchange rate on the day of the experiment was 1 US dollar = 5.54 Norwegian kroner.

Participants played 10 rounds of the stage 1 game, and this stage was identical for both treatment and control sessions. They were explained that their realized payoff from stage 1 of the experiment would be the payoff from one of the ten rounds, determined by a random draw.

After each round, participants were given information on their earnings (endowment + payoff from own sticker use), how much they would lose from the harm imposed on them by other group members, and how much they would receive net, *should* that particular round be drawn. This information was provided in order to increase the probability of a move to a non-cooperative solution of the public good game, and to further establish the sticker as a morally suspect good.

3.2. Stage 2 (treatment) – establishing a market for harmful stickers

After stage 1 was completed, the treatment groups were told that the right to use stickers would now be constrained to avoid excessive harms. Each participant would be given an endowment of 150 kroner and two “rights” to use stickers per round. These use rights could be traded within the group. Retiring use rights was not possible: each participant would automatically use their stickers. Total harm from externalities was thus fixed, and $\sum_i R_i = 8$ for all groups.

In the stage 2 game, the participant’s decision problem was to decide on the number of user rights to buy or sell. Whatever stickers they possessed at the end of the trading round were automatically used.

The double auction market was implemented in a way that ensured that all realized trades would be Pareto-improving in monetary payoffs. Since the total number of stickers used in a four person group would always be 8, $\sum_{j \neq i} R_j = 8 - R_i$, we can rewrite the payoff of participant i :

$$\pi_i = 150 - 80 + (a_i + 10)R_i - (R_i - 2)p_R = 70 + (a_i + 10 - p_R)R_i + 2p_R,$$

where p_R is the trading price for the right to use stickers. A participant with $a_i = 5$ will thus gain from selling rights at prices of 15 kroner and above, while those with $a_i = 25$ will gain from buying at prices of 35 kroner and below. To ensure that no participant saw their payoffs reduced after trading, market prices were constrained to lie in the interval from 15 to 35 kroner. The market was open for one and a half minutes in each round.

In the main treatment, the instructions were designed to highlight the harm imposed on others by sticker use: using a printed table and a detailed run-through of a single example, participants were shown how sticker sales would affect payoffs through multiple channels. For instance: if a seller sold a right for 15 kroner, the seller would (i) receive the agreed price of 15 kroner, but also (ii) lose the revenue from using the sticker him/herself (5 kroner), and (iii) lose 10 kroner due to the increased harm the extra sticker-right would allow the buyer to impose on the seller.

Before trade started, participants were asked a hypothetical question: “Do you think it should be possible to trade the user rights?”

As in stage 1, stage 2 lasted for 10 rounds – one of which was randomly chosen to determine actual payoffs. After each round, participants were informed about the payoffs that would result from this stage of the experiment if the prior round was drawn. The information was presented in a manner highlighting the shifting of harms, including: the participant’s gross income (endowment and earnings from any stickers one held at the end of the round), the gross loss (loss due to sticker use by other group members), and net earnings.

3.3. Stage 2 (control) – establishing a market for non-harmful stickers

After completing stage 1 participants in the control treatment were told that the next stage of the experiment would introduce a new type of stickers that imposed no harm on others. In actuality, the game was the same as that played by participants in the main treatment, but with endowments and productivities altered to “hide” the similarity. By making the externality effects invisible and

presenting new private productivities, the new stickers appeared substantively different from the harmful stickers encountered in stage 1.

The low productivity type now earned 15 kroner per new sticker used and the high productivity type earned 35 kroner per sticker used. Each participant received 2 such “new” stickers in the beginning of each round, and could trade these within their four-person group. The trade was conducted in the same double auction market as the treatment groups, with prices limited to the interval between 15 and 35 kroner, in order to ensure that trade was Pareto improving.

In this way, we ensured that the game played was formally equivalent to that played by participants in the main treatment. This implied an endowment per participant of 70 kroner. The resulting payoff for participant i was then:

$$\pi_i = 70 + (a_i + 10)R_i - (R_i - 2)p_R = 70 + (a_i + 10 - p_R)R_i + 2p_R$$

Before trade started, participants were asked a hypothetical question: “Do you think it should be possible to trade the stickers?”

As in stage 1, stage 2 lasted for 10 rounds – one of which was randomly chosen to determine actual payoffs. After each round, participants were informed about the payoffs that would result from this stage of the experiment if that particular round was drawn. They received information about their net income (endowment and earnings from any stickers used and income/expenditure from traded stickers), with no reference to externalities or harms.

3.4. Stage 3 – revealing attitudes toward market

In stage 3 of the experiment, participants in both treatments were told that they would play one more round of the game they had played in stage 2, but that the presence of a market would be decided by a majority vote. The question was otherwise identical to that used in the “hypothetical” vote conducted prior to stage 2.

After casting their vote, the last round was run in accordance with the majority decision of the session. Stage 3 consisted of only one round, with the choices made in that round determining the payoffs. After the round, participants were given information on their stage 3 payoffs. They were also told which of the stage 1 and stage 2 rounds that would determine their realized payoffs from those stages, in addition to the total payoff earned from the experiment as a whole. This information was followed by a short post-experiment survey.

4. Results

Participants were recruited from large undergraduate classes in mathematics⁶ and assigned to two experimental sessions held back-to-back on Friday, September 9th, 2011. There were 24 participants in the treatment session (6 groups) and 20 participants in the control session (5 groups). At the request of one of the referees, an additional four sessions were held back-to-back on Tuesday, March 25th, 2014 in order to increase statistical power (see Section 5). These participants were recruited from classes in the philosophy of science, mandatory to all students at the University of Oslo. There were 40 participants in the new treatment sessions (10 groups) and 40 participants in the new control sessions (10 groups). In total: 64 participants (16 groups) in treatment sessions and 60 participants (15 groups) in control sessions.

4.1. Stage 1: extensive sticker use in the common pool resource game

On average over the 10 rounds, 70% of all stickers were used, with some group heterogeneity. Calculating, for each group separately, the average share of stickers used over the 10 rounds, the share used varied between groups from a minimum of 26.3% to maximum of 95.5%. There were no systematic differences in group sticker use by control and treatment sessions (see Fig. 1). Using a

⁶ These are students who plan to proceed with math, computer science or natural sciences.

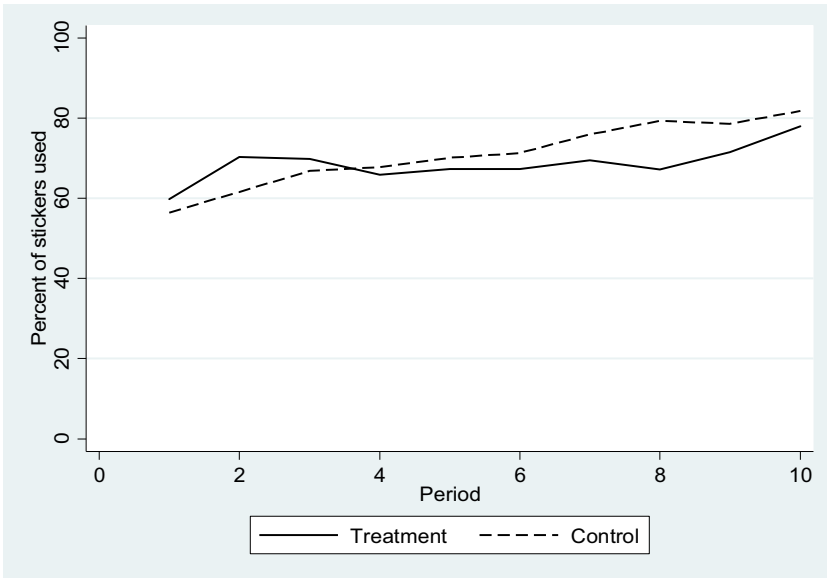


Fig. 1. Stickers use by period and treatment in Stage 1.

t-test at the group level: calculating the mean sticker use in the first five and last five periods for each group separately, we found no significant differences between treatment and control groups in either set of periods. In the first five periods, treatment-groups used an average of 66.7% of their stickers against 64.6% in the control groups (p -value = 0.7). In the last five periods, the numbers were 70.7% for treatment and 77.4% for control groups (p -value = 0.27).

In the 10 rounds of stage 1, average sticker use increased from 58% in the first round to 80% in the last. A regression of group-level sticker-use on the round-number (with group fixed effects) gives a statistically significant positive trend of 1.8 percentage points per round ($p < 0.01$).

On average across the 10 rounds, low productivity players used 1.2 fewer stickers than high productivity players ($p = 0.01$). Recall that while all participants imposed a loss of 30 kroner on the other members of the group for each sticker they used, the private payoffs for imposing this harm was only 5 kroner per sticker for low earners as compared to 25 kroner for high earners. If all the members in a group used 7 stickers each, the high productivity members would earn 105 kroner while the low productivity members would lose 35 kroner. If no one in the group used stickers, each member's net payoff would equal the endowment of 140 kroner and this would clearly be a Pareto improvement.

4.2. Stage 2: no difference in market attitudes between the treatment and the control

A cap on sticker use equal to two stickers per person gives a payoff of 90 and 130 (to low and high productivity type respectively), and is thus a Pareto improvement relative to the “laissez faire” sticker use in stage 1 of the experiment. The results from stage 2 are shown in Table 2.

Participants were first asked to vote hypothetically for or against trade, given the cap on sticker-use per player. In the treatment sessions 80% of the participants supported trade in the hypothetical vote, while 83% of the control sessions participants did the same. The difference was not significant ($p = 0.61$). However, when we look at the separate types in the treatment sessions only, we see that the low productivity types are significantly more negative to trade than the high productivity types in this pre-experience hypothetical vote (69% of low productivity types versus 91% of high productivity types oppose trade in the treatment session, $p = 0.03$). This difference is not found in the control sessions. Low productivity types in the treatment sessions are also more reluctant to trade in this pre-experience vote than low productivity types in the control sessions (a weakly significant difference, $p = 0.09$).

Table 2
Market attitudes in stage 2 and stage 3 of the experiment.

	Treatment	Control	t-test
Pre-experience hypothetical vote	80% ($n = 64$)	83% ($n = 60$)	$p = 0.61$
% in favor of trade (n)			
Low productivity types	69% ($n = 32$)	87% ($n = 30$)	$p = 0.09$
High productivity types	91% ($n = 32$)	80% ($n = 30$)	$p = 0.24$
Realized trade	93% ($n = 16$)	88% ($n = 15$)	$p = 0.13$
In % of possible trades			
Average price	27.8 kroner ($n = 16$)	26.9 kroner ($n = 15$)	$p = 0.42$
Price (n)			
Post-experience real vote	64% ($n = 64$)	77% ($n = 60$)	$p = 0.13$
% in favor of trade (n)			
Low productivity types	72% ($n = 32$)	87% ($n = 30$)	$p = 0.16$
High productivity types	56% ($n = 32$)	67% ($n = 30$)	$p = 0.41$

In the treatment sessions, 93% of all user rights were traded, with an average price of 27.8 kroner per right. As a price of 25 kroner would give equal gains to both players, an average of 27.8 kroner is somewhat skewed in favor of the low productivity types (sellers) who were also worse off before trade. In the control sessions, 88% of all stickers were traded, with an average price of 26.9 kroner. There was no significant difference in the amount of realized trades or the average prices between the treatment and control.

4.3. Stage 3: no difference in voting behavior between the treatment and the control

Finally, when actually voting over trade, the support for trade was still substantial. 64% voted in favor of trade in the treatment sessions, while 77% voted for trade in the control session. While we expected the control group to have fewer moral objections to trade, the overall difference was not significant ($p = 0.13$). Comparing the post-experience vote with the pre-experience hypothetical vote, the support for trade is relatively constant among the low productivity types but has fallen considerably among the high productivity types. In the treatment sessions, the support for trade among high productivity types fell from 91% to 56%, a significant reduction ($p = 0.0015$). The support for trade among high productivity types fell in the control sessions as well, however not significantly.

A possible reason why the support for trade fell is that it took time in the lab. As each market was open for 90 s, 10 trading rounds took at least 15 min. Subjects eager to leave the lab may thus have voted against more trading. Actually, the last stage had only one trading round, but this may not have been sufficiently clear from the instructions.

4.4. Potential reasons for no market opposition

The results fail to support the hypothesis that participants resist trade *because* this would legitimize an objectionable action. This does not in any way undermine Sandel's moral argument, but it does show that an objection along the lines of his argument is not triggered in this specific experimental setting. There are several possible reasons for this, some of which can be ruled out.

4.5. Have we recruited pro-trade participants?

One possible reason why our participants are so positive to trade could be that we recruited a particularly pro-trade sample. While we took care not to recruit among economics students to avoid the doctrinal thinking Sandel claims characterizes economists, one third of our sample was recruited from mathematics students, who may also hold particularly favorable view of markets. A post-experimental survey elicited participants' attitudes toward a number of morally questionable real world markets.

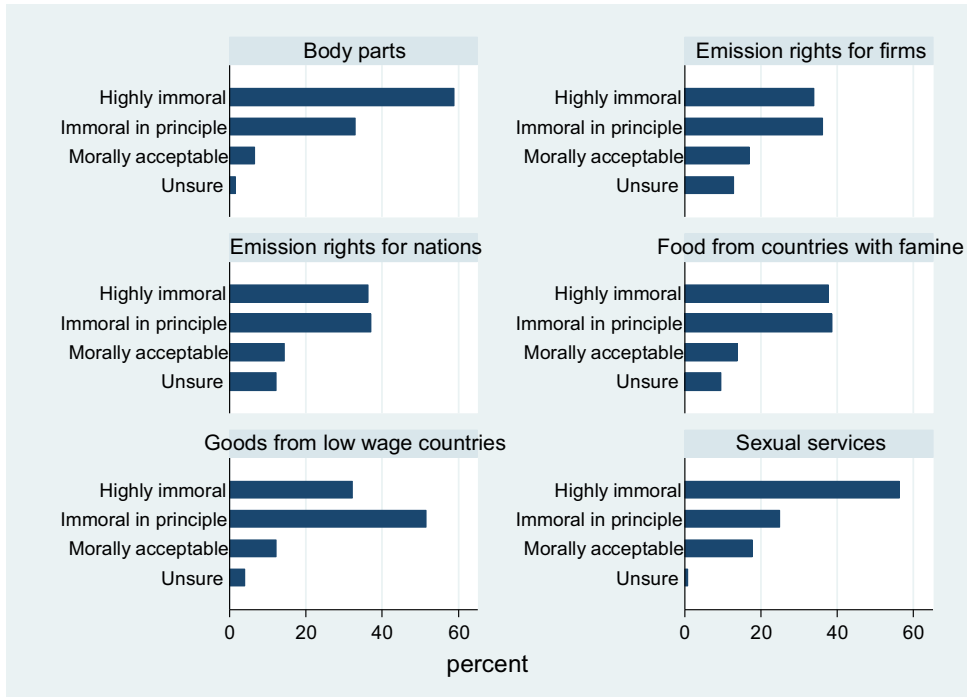


Fig. 2. Post-experiment attitudes toward markets, both sessions ($n = 124$).

The sample exhibits attitudes toward markets that are, to economists, rather astonishing in their anti-trade sentiments.

The attitudes toward specific markets were elicited by this question:

For each of the following areas, to what extent do you think that it is morally acceptable to trade in these goods/services for money? 1-Very immoral, 2-In principle immoral, but acceptable in some cases, 3-Morally acceptable, 4-Don't know.

The goods/services were: (a) Body parts, (b) Sexual services, (c) Emissions rights for firms, (d) Emissions rights for nations, (e) Goods produced in poor countries by workers on a very low salary, (f) Food from countries where a large share of the population is suffering from hunger and malnutrition.

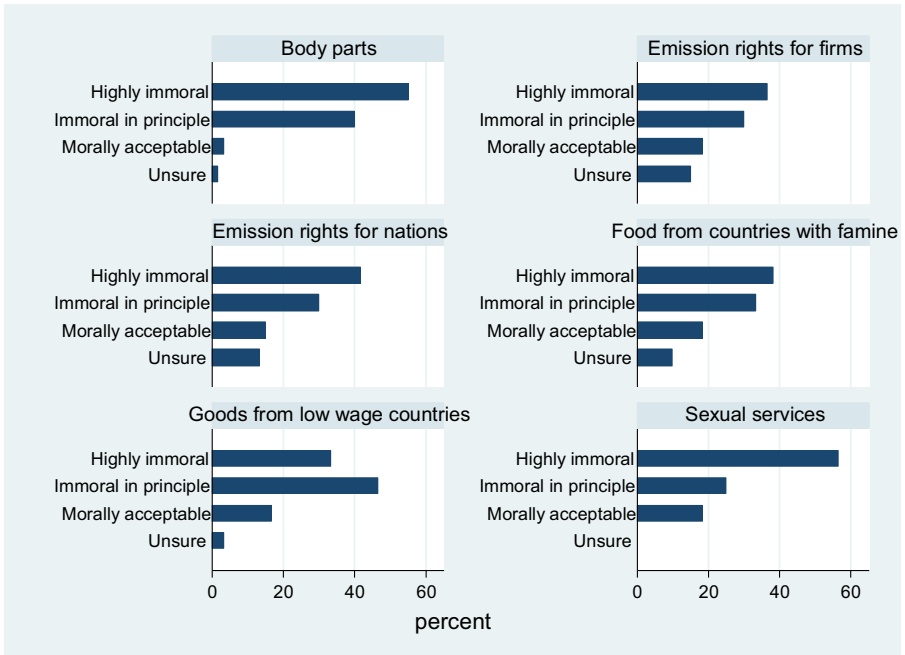
The results for all sessions combined are shown in Fig. 2. All the markets are judged to be immoral by a vast majority of participants, and Mann Whitney tests revealed no significant differences between the treatment and control sessions in their post-experiment attitudes to such markets. Fig. 3 shows the response distributions for control and treatment sessions separately.

4.6. Statistical power

Our experiment is an attempt to examine whether the virtue ethical reasoning discussed above can explain the widespread opposition to quota-trade. We argue that this is not the case, as we find no significant difference between the different treatments. We initially reached this conclusion after running the experiment on 44 participants in September 2011. A referee suggested that this sample might be too small to identify an effect, and recommended a power analysis and further sessions to increase the number of participants.

Our research question was whether Sandel's moral argument could explain a substantial share of the opposition to quota trade. As noted above, a nationally representative sample found 46%

a) Control (n=60)



b) Treatment (n=64)

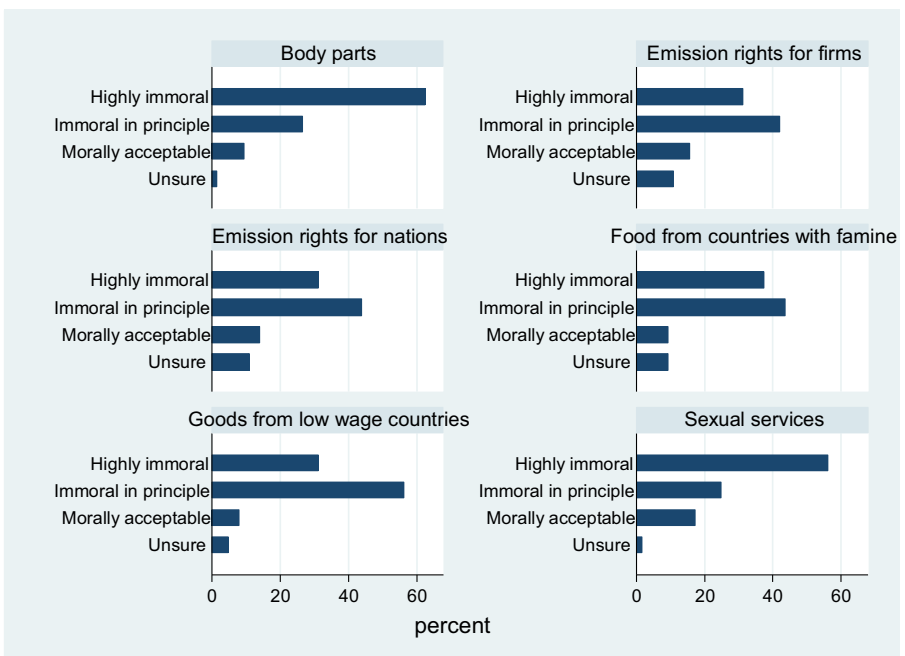


Fig. 3. Post-experiment attitudes toward markets, by treatment.

of the population opposing tradable emission quotas. Opposition amongst our participants seems even higher: trading of emission rights between nations was judged as “highly immoral” in the post-experiment survey by 36%, with an additional 37% seeing it as immoral in principle—but acceptable in some cases. Thus the opposition in our subject pool in the range 36–73%. Counting at least a fraction of those who answered “immoral in principle” as opposing emission rights trade, the opposition seems to be at least as high as the 46% in the national sample. If virtue ethics is an important explanation of this opposition, we would think that it should account for at least 15–20 percentage points of the trade opposition. We would thus want a sample size in which we would with high probability detect a 20% treatment effect, and with considerable probability also detect a 10% treatment effect. An effect of less than 10% would be of limited interest.

Under the null hypothesis, opposition to trade is the same in both treatment and control groups. Based on the 2011 sessions with 44 participants this gives us an 11% baseline opposition (10% vote against trade in the control session and 12.5% vote against trade in the treatment session). As discussed above subjects may have objected to trade to save time in the lab, or it may reflect all kinds of opposition to trade. Even if we assume that this opposition revealed by voting in our sample is not related to Sandel’s argument and subtract this opposition from the 46% opposition we want to explain, this leaves us with 35 percentage points of the sample opposing trade for reasons that may be related to Sandel’s argument. The 20% and 10% benchmarks remain reasonable in light of this.

What is the probability that we would reject the null hypothesis at conventional significance levels if the true average effect of the treatment is a 10 or 20 percentage point increase in trade opposition?

Since the mean of a binomial distribution is approximately normal, the probability that the means will be at least t^* standard deviations apart can be shown to be approximately $\Phi\left(\frac{p-q}{\sigma} - t^*\right)$ where p is the opposition to trade in the treatment group and q the opposition in the control group. The standard deviation σ is a function of the number of observations and the probabilities. (We ignore the probability of falsely finding significantly more opposition in the control group.) Following the tradition of using the stricter two sided test even with a one sided hypothesis, the relevant threshold for a 10% significance test is $t^* = 1.645$.

To reach a power of⁷ 80% if the true effect is 20% we would need 28 more observations in both control and treatment. We ran two new sessions adding 40 subjects in both control and treatment. The power is 88% when using the initial base rate, but the general opposition to trade was higher in the last sessions,⁸ reducing the power to 84%, but still above the standard 80% threshold. With a smaller true effect of 10% we would get a significant effect with 45% probability. It would take a sample size of more than 200 in each group to reach a power of 80% in this case. We can thus be rather confident that if there was a strong effect we would find it, while with a small true effect (10%) we may have failed to find it just by bad luck.

4.7. External assessment of morality of experimental choices

Sandel’s argument has two premises: an activity X is immoral, and trade legitimizes the use of the traded goods or services. One reason for our lacking results could be that sticker use in the first part of the experiment was not seen as immoral.

To test this possibility, we conducted a survey on Amazon Mechanical Turk. This is an online “labor market” where individuals from around the world do short, simple, web-based tasks for real money. The “Turkers” were presented with an English translation of the experimental instructions, somewhat simplified as the mechanics and implementation of the double auction trade were irrelevant for the question. They next had to answer comprehension tests (with monetary incentives for correct understanding). Finally, they were asked to evaluate the morality of sticker use in the first part, and for the second part of the experiment, to evaluate the morality both of establishing a market, and of trading in the market. 56 persons answered the survey which was conducted on June 1st 2012.

⁷ The 80% threshold is considered standard in the literature. See e.g. [Bellemare et al. \(2014\)](#) who also find that the required sample size greatly exceeds sizes used in the experimental economics literature.

⁸ That is based on the initial results we use $p = 0.31$ and $q = 0.11$, while the updated opposition would imply $p = 0.38$ and $q = 0.18$ with a 20% effect.

Table 3

Survey attitudes toward sticker use in the experiment from Mechanical Turk.

	Moral	Neutral	Immoral
Sticker use by the low earning type	29%	29%	43%
Sticker use by the high earning type	41%	30%	29%
Allow a market for sticker use	35%	21%	43%
Buy sticker-use rights	36%	37%	27%
Sell sticker-use rights	32%	30%	38%
	Disagree	Neutral	Agree
Market will make sticker use seem more moral	20%	30%	50%

The respondents claimed to see the experimental choices as morally relevant. More participants responded that stickers use was immoral than those who responded that it was moral. The results are presented in Table 3. Note that the categories “strongly moral” and “somewhat moral” have been merged, and similarly the categories “strongly immoral” and “somewhat immoral”.

A considerable share of the Turkers thus considered sticker-use to be immoral, and saw markets in user rights as undermining this. Turkers also answered the same questions regarding the morality of the particular markets from the post-experimental survey. In these questions, the Turkers expressed less anti-trade attitudes than the experimental participants. This supports the previous conclusion that our experimental sample was not generally more pro-trade than others.

Still, we have several caveats about the responses from Mechanical Turk. First, with no incentives to answer truthfully, random responses cannot be ruled out. The answers are indeed rather evenly distributed with a stronger emphasis on the middle alternatives, a similar pattern to what random responses would produce. Note, however, that the different assessments of stickers use by high and low earners make good sense. There is a stronger tendency to agree to the claim that markets can legitimize an action, which is also consistent with the findings on crowding out in the economics literature.

Moreover, the majority of respondents were unable to answer all comprehension questions correctly. Probably this is partly due to complicated instructions, lack of practical experience with the experiment (as opposed to the participants in the experiment) and insufficient time used by Turkers to read instructions. Although Turkers were given a bonus if all comprehension tests were correctly answered, they did not have the same incentives to thoroughly understand the instructions as participants in the experiment. Participating Turkers on average spent nearly 17 min on the survey, which should be enough time to read the 2 pages of instructions thoroughly, answer 7 comprehension tests and answer 6 ethical judgments. Since this is only the average, a large share may have used considerably less time and hence misunderstood.

5. Discussion and conclusion

Our experiment does not support the hypothesis that people oppose trade in immoral goods because markets legitimize actions that deserve to be stigmatized. This does not imply that such a moral argument is invalid, only that our experimental participants did not apply this logic in this specific context. The results show that moral opposition is not necessarily generated in markets trading immoral goods. However, as with any non-result the lack of result could be due to poor experimental design. While we have tried to test for some of the possible problems that could arise, we may have overlooked others.

One reviewer point out that “people often view (perhaps incorrectly) pollution as being done by “others” (firms) and victims as “us” whereas in [our] experiment everyone is both victim and polluter.” In the lab this could correspond to having the externality hit some passive participants possibly in another room. While an economist’s response to this claim would be that the externality is unaffected by the trade, we cannot rule out that it would have an impact on how subjects assess the moral standing of trade in the given setting. [Ciccone et al. \(2015\)](#) do find that distributional impact of trade matters,

and our subjects are negative to buying inexpensive goods produced by low-paid workers. Still, as Turkers found the stickers use wrong in our case too, we doubt that this is the essential element of the moral evaluation.

A different counterargument would be that check-mating your opponent in a game of chess is moral, even though it imposes a loss on your opponent. The experiment may similarly be viewed by participants as a game to be “won.” With such a mindset, sticker use may not be seen as morally wrong. This, however, should then apply to many experiments on moral behavior. In our case, it would imply that participants will act as if they do not have any moral objections to a market in the right to do wrong, while outside the experimental context they do. The assessment of the Turkers, however, is more consistent with sticker use being morally wrong, even within the context of the experiment.

Another possibility, related to the different mindset triggered in the field and in the experiment, is that experiments in economics “train” participants to focus on consequences. By highlighting and focusing on consequences, the set-up and instructions may prime for a consequentialist ethic. The argument that trade legitimizes morally objectionable behavior seems to build on non-consequentialist ethics: the *intention* (to be moral) is important in itself, independently of the behavior it generates. If so, precise instructions and training rounds to teach payoff consequences may temporarily erase or weaken the reasoning that ordinarily drives judgments of markets.

Acknowledgement

We are grateful to participants at the Norwegian Annual Meeting for Economists 2012, to the Editor and two anonymous referees for helpful comments and suggestions. We are further grateful to the Research Council of Norway for funding through the *Miljø2015* program and project 193703 – Improving international cooperation on emission abatement. Brekke and Rogeberg are associated with and the project is partly financed by CREE – the Oslo Centre for Research on Environmentally Friendly Energy – which is supported by the Research Council of Norway (Grant number 209698). Brekke also acknowledge the support of the Centre for Equality, Social Organization and Performance (ESOP).

References

- Baumol, W.J., Oates, W.E., 1988. *The Theory of Environmental Policy*, 2nd ed. Cambridge University Press.
- Bellemare, C., Bissonnette, L., Kröger, S., 2014. *Statistical Power of Within and Between Subject Design in Economic Experiment*. IZA Discussion Paper No. 8583.
- Bowles, S., Polania-Reyes, S., 2012. Economic incentives and social preferences: substitutes or complements? *J. Econ. Lit.* 50 (2), 368–425.
- Brekke, K.A., Kverndokk, S., Nyborg, K., 2003. An economic model of moral motivation. *J. Public Econ.* 87 (9–10), 1967–1983.
- Bruni, L., Sugden, R., 2013. Reclaiming virtue ethics for economics. *J. Econ. Perspect.* 27 (4), 141–164.
- Ciccone, A., Braaten, R., Rogeberg, O., 2015. Fairness preferences in a bilateral trade experiment. CREE Work. Paper Ser. 10, 2015.
- Eyckmans, J., Kverndokk, S., 2010. Moral concerns on tradable pollution permits in international environmental agreements. *Ecol. Econ.* 69 (9), 1814–1823, July 15.
- Fehr, E.K., Schmidt, M., 1999. A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114 (3), 817–868.
- Frey, B.S., Meier, S., 2005. Selfish and indoctrinated economists? *Eur J Law Econ* 19, 165–171.
- Gardiner, S.M., 2004. Ethics and global climate change. *Ethics* 114 (3), 555–600.
- Gneezy, U., Rustichini, A., 2000. A fine is a price. *J. Legal Stud.* 29 (1), 1–17.
- Paxton, J.M., Ungar, L., Greene, J.D., 2011. Reflection and reasoning in moral judgment. *Cognit. Sci.* 36 (1), 163–177.
- Sandel, M.J., 2005. *Public Philosophy: Essays on Morality in Politics*. Harvard University Press.
- Sandel, M.J., 2012. *What Money Can't Buy: The Moral Limits of Markets*. Farrar, Straus and Giroux.
- Sapienza, P., Zingales, L., 2013. *Economic Experts vs. Average Americans* (Unpublished note).
- Strand, J., Rosendahl, K.E., 2012. Global emissions effects of CDM projects with relative baselines. *Resour. Energy Econ.* 34 (4), 533–548.