Journal of Environmental Economics and Management 127 (2024) 103010

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

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

Can policy packaging help overcome Pigouvian tax aversion? A lab experiment on combining taxes and subsidies☆

Gøril L. Andreassen [a,c,*], Steffen Kallbekken [b], Knut Einar Rosendahl [a]

[a] School of Economics and Business, Norwegian University of Life Sciences, PB 5003 NMBU, 1432 Ås, Norway
[b] CICERO Center for International Climate Research, PB 1129 Blindern, 0318 Oslo, Norway
[c] Institute of Transport Economics, PB 8600 Majorstua, 0359 Oslo, Norway

ARTICLE INFO

JEL classification:
D72
H23
Q54
Q58

Keywords:
Pigouvian taxes
Policy packaging
Public support
Lab experiment
Tax aversion

ABSTRACT

Tax aversion makes it politically challenging to introduce Pigouvian taxes. One proposed solution to overcome this resistance is to package policies. Using an online lab experiment, we investigate whether combining a tax and a subsidy is perceived as more acceptable than the tax or the subsidy alone. The purpose of the policies is to reduce demand for a good with a negative externality to the socially optimal level. We find that support for a combination of a tax and a subsidy equals the simple average of support for the two instruments alone. Combining a tax and a subsidy therefore does not reduce tax aversion. We examine potential mechanisms behind the tax aversion. Participants believe they will receive a lower share of the tax revenue when the tax is implemented alone than when it is combined with a subsidy, i.e. the participants in the tax alone group hold more pessimistic beliefs about the tax revenue. We also find that the participants expect the tax to be more effective in reducing demand for the good with a negative externality than both the subsidy alone and the combinations of tax and subsidy. This belief does not, however, translate into support for the tax.

1. Introduction

Pigouvian taxes are crucial policy instruments to cost-effectively reduce negative externalities such as emissions of greenhouse gases and other types of pollution, as they internalize the external costs (Timilsina, 2022). However, public opposition towards Pigouvian taxes makes it challenging for policy makers to introduce them. The yellow vest protests against the fuel tax increase in France in 2018 is the iconic example of peoples' disapproval of carbon taxes (Douenne and Fabre, 2022). Another indication of the unpopularity and political difficulty of introducing taxes is that carbon pricing only covers 23% of global greenhouse gas emissions, with only 4% of emissions having a price sufficiently high to keep global warming below 2 °C (The World Bank, 2022). The world's inability to correctly price externalities is tremendously costly: According to Parry et al. (2021), explicit and implicit global fossil fuel subsidies amounted to 6.8 percent of global GDP in 2020, mostly due to lack of environmental and other taxes.

It is therefore essential to explore policy designs or other interventions that can increase public support for Pigouvian taxes. One idea that has been garnering growing attention in multiple fields is to create policy packages (Givoni et al., 2013; Kern et al., 2019). Put simply, the idea is to combine effective but unpopular policies with less effective but more popular policies, to use secondary

☆ This work was supported by the Research Council of Norway, grant number 308789.
* Corresponding author.
E-mail addresses: gla@toi.no (G.L. Andreassen), steffen.kallbekken@cicero.oslo.no (S. Kallbekken), knut.einar.rosendahl@nmbu.no (K.E. Rosendahl).

https://doi.org/10.1016/j.jeem.2024.103010
Received 12 July 2023
Available online 3 June 2024

policies to offset undesirable impacts of the primary policies, or that the way the instruments work together is helpful.[1] In this paper we use a lab experiment to investigate how combining a tax with a subsidy influences the relative level of support.[2] We also examine the role of beliefs about the tax, the subsidy and the combinations of the two instruments to understand the low level of public support for taxes.

Economists have labelled the opposition against taxes "tax aversion". Tax aversion can be defined as opposition towards tax schemes that would increase both individual and social economic welfare, based on incorrect and pessimistic beliefs about the properties of the tax such as its effectiveness and fairness.[3] Several of the factors shaping people's views about Pigouvian taxes are well established (Bergquist et al., 2022; Drews and Van den Bergh, 2016). From economic theory one would expect economic self-interest to play a central role. Whereas it does play a role, it cannot fully explain the opposition against Pigouvian taxes (Anderson et al., 2023; Douenne and Fabre, 2022; Heres et al., 2017; Kallbekken and Sælen, 2011; Umit and Schaffer, 2020; Dechezleprêtre et al., 2022). Several studies show that the belief that taxes are not environmentally effective is one of the most important determinants of public opinion about taxes (Bergquist et al., 2022; Douenne and Fabre, 2022). Beliefs about fairness have also persistently been shown to be important (Douenne and Fabre, 2022; Dechezleprêtre et al., 2022).[4]

Research has identified some strategies that can be helpful for overcoming tax aversion. Earmarking of the tax revenue seems to increase support (Kallbekken et al., 2011; Baranzini and Carattini, 2017; Dechezleprêtre et al., 2022).[5] What the earmarked tax revenue is spent on can have strong impact on the level of support, but varies between groups of people (Anderson et al., 2023; Dechezleprêtre et al., 2022). Heres et al. (2017) find in a lab experiment that informing the participants that the tax revenue is returned in equal proportions to them, increases their support for taxes. However, the support for both a tax and a subsidy increases when there is no uncertainty about what happens to the income from the tax and the cost of funding the subsidy.[6] Allowing people to experience positive effects of an environmental tax, can increase support (Cherry et al., 2014; Schuitema et al., 2010; Winslott-Hiselius et al., 2009). The results on providing more information about how environmental taxes work are mixed (Kallbekken et al., 2011; Douenne and Fabre, 2022; Dechezleprêtre et al., 2022). Avoiding the term "tax" can under some circumstances lead to higher support (Kallbekken et al., 2011; Hardisty et al., 2010; Baranzini and Carattini, 2017).

A scarce literature explores the impact on the level of public support of combining coercive instruments like taxes with other and more popular policies. The interesting dynamic is how preferences for different types of policies interact when they are combined. Is, for instance, the joint assessment of combined instruments (policy packages) dominated by the instrument the respondents are most (or least) averse to, or is it a simple averaging of the preferences for each instrument?

The early contribution by Eriksson et al. (2008) finds that support for a combination of two instruments is higher than for the most restrictive instrument alone, but the level of support is lower than the average of support for the two constituent parts (a fossil fuel tax combined with either improved public transport or subsidies for renewable fuels). However, when the instruments are presented in isolation, it appears that no information is provided on how the tax revenues are to be spent or how the subsidies are to be financed, whereas this information is provided for the two policy packages.

Two more recent and experimental papers indicate a positive dynamic: Milkman et al. (2012) find that bundled policies are valued more highly than the most popular policy is valued on its own.[7] The policies vary in their costs and benefits (jobs lost, acres of forest protected, etc.), and hence the results are somewhat difficult to interpret. Using a choice experiment, Fesenfeld (2022) finds that bundling policies may reduce opposition to taxation. He studies the impact of policy complexity by comparing responses to low complexity policy proposals (one goal and one instrument) with high complexity policy proposals (one goal and four policy instruments). When a large tax increase is added to a low complexity policy proposal, he finds that it decreases the probability of choosing that package (the average marginal component effect) by 15 and 27 percentage points among German and US respondents, respectively. However, when the same large tax increase is added to a high complexity policy proposal, it decreases the probability of choosing that policy package by only 9 and 18 percentage points, respectively. In this choice experiment, payoffs and policy effectiveness are not made explicit, but the subsidies offer lower consumer prices (for food and transport) at no explicit cost.

[1] Ambec and Coria (2021) find that taxes reveal information about a firm's abatement costs and this information can be used to set an emission standard. Further, Acemoglu et al. (2012) find that a combination of a tax and a subsidy is optimal to target the interaction effects between environmental externalities and intertemporal knowledge externalities.

[2] Using both taxes and subsidies to target the same externality is a common practice. One example is a congestion tax on cars and a subsidy for public transport to target pollution from cars as well as congestion. Another is energy taxes combined with subsidies for investing in energy saving, such as improved insulation or installing heat pumps. Helm and Mier (2021) investigate the optimal mix of subsidies and taxes for intermittent renewable energy and energy storage. Combining taxes and subsidies to target a negative externality will probably not be as cost-effective as having only taxes, see for instance Gugler et al. (2021).

[3] This combines and further develops the definitions of Kallbekken et al. (2011) and Douenne and Fabre (2022).

[4] Ideology or political attitudes can also be an important argument against taxes (Cherry et al., 2017). The view that the government should by principle not decide what one can and cannot do, could be a driver in the opposition against taxes. However, Douenne and Fabre (2022, p.83) find that "these results suggest that the rejection of carbon taxation does not typically result from clashing principles, such as a disinterest in the climate or a dislike of price instruments, but rather from overly pessimistic beliefs about the properties of the reform".

[5] We do not define earmarking of tax revenue to a specific purpose as a policy package or policy bundle. To be defined as a policy package, different instruments have to be combined.

[6] From the research design in Heres et al. (2017), we cannot disentangle whether the increased support comes from the removal of the uncertainty or that the tax revenue is shared between the participants, or both.

[7] In the article they write that "bundled legislation is valued more than the sum of its parts", but the correct wording seems to be "higher than the most popular policy on its own". For example, Bill 1 and Bill 2 has 54% and 45% support, respectively, while the combined Bill has 83% support.

The existing literature does not disentangle the mechanisms that can explain how a joint preference for combined instruments is formed: Instrument type varies together with costs and benefits, and these studies are therefore unable to pinpoint what causes the level of support for the combination to differ from the level of support for the constituent parts. Based on the diverging previous findings, we explore the dynamics of how the preferences for a tax and a subsidy interact to form the preference for a combination of the two in a setting where (1) participants' decisions are incentivized, (2) the study is sufficiently powered to detect a realistically small yet still notable percentage point difference in support for policies (see pre-analysis plan (p.7) in Appendix D for details)[8], (3) payoffs are equal across instruments (for the same behaviour), and (4) the combinations of instruments are fractional rather than additive. This means that when the tax and subsidy are combined, the tax and subsidy rates are lower than when the tax or the subsidy are implemented alone.

Further, we want to investigate the mechanisms behind the lack of support for taxes by comparing beliefs about a tax with beliefs about a subsidy and combinations of the two. Expectations about the effectiveness of the instrument in reducing the externality and about what happens with the tax revenue and the subsidy cost are interesting to shed light on in order to deepen our understanding of attitudes towards Pigouvian taxes.

In our experiment we introduce a market for a fictitious good with a negative externality where participants earn a financial reward (payoff) through the profit they make by purchasing units of the good (the difference between the gross value of the good and the price of the good). At the same time they are negatively affected by the externality from the units purchased by the other participants in their group. Participants vote on the introduction of policies that can incentivize participants to purchase the socially optimal number of units through a tax, a subsidy, or combinations of the two instruments. If implemented, the tax is charged for any units purchased, whereas the subsidy is paid for any of the units *not* purchased. The participants are randomized into five different groups: (1) 100% tax, (2) 75% tax & 25% subsidy, (3) 50% tax & 50% subsidy, (4) 25% tax & 75% subsidy, and (5) 100% subsidy.

Taxes have two core properties: First, they change the price the consumer faces so that demand decreases (as long as demand is not fully inelastic), which in turn reduces the external costs. Second, they generate revenue that can be spent by the government, such as distributing it back to the citizens. Subsidies also change the (direct or implicit) price the consumer faces, but instead of generating income for the government, subsidies need to be financed. If the tax is implemented in our lab experiment, the revenue collected from each participant is split equally between the other group members.[9] The subsidy payments received by each participant are financed through equal contributions from the other group members. The (combined) tax/subsidy rate is always equal to the marginal damages caused by purchasing the good. In this way we ensure that all policies in the experiment produce identical payoffs for the same behaviour for all group members. However, we cannot verify whether the participants actually take this payoff into account when they vote.

The first contribution of this paper is that we find that the point estimates of the public support increases linearly as the subsidy share increases and the tax share decreases in the fractional combinations of the instruments. As explained above, this is in a setting where decisions are incentivized and the payoff structures are identical across policies. Our experiment with 1641 participants thus produces results that do not align with the previous findings of Milkman et al. (2012) and Fesenfeld (2022).[10] The dynamics we observe indicate no beneficial effect on public support from combining policy instruments beyond the simple averaging of preferences for the constituent parts of the package. Thus, the increased support found in Milkman et al. (2012) might have come from the increased gain in the policy package and not from the strategy of packaging as such.

The second contribution of this paper is that we find pessimistic beliefs about what happens to the tax revenue. The 100% tax group believes that they will receive a smaller share of the tax revenues than the other groups believe. This finding is consistent with Douenne and Fabre (2022, p.83), who find that the opposition against the carbon tax comes "from overly pessimistic beliefs about the properties of the [carbon tax] reform". Revealing pessimistic beliefs in different contexts is important to understand the opposition towards Pigouvian taxes. We find more pessimistic beliefs about what share of the tax revenues the participants in the experiment will receive (a piece of information that is not clearly shared with them initially, cf. Section 2) when a tax is the only policy, than when it is combined with a subsidy. Similarly, the beliefs about whether the proposed policy will increase the payoff compared with no policy are more pessimistic when the tax is the only instrument than when the tax is combined with a subsidy or the subsidy alone. Furthermore, for the combination of a tax and a subsidy, the share expecting the policy to increase the payoff declines with the share of the tax in the proposed policy. Participants do expect to pay a substantial share of the subsidy cost, but unlike the expectations about tax revenues this share does not differ significantly across treatments. Because our design has fractional combinations of the tax and the subsidy, we can investigate what happens with a gradual decrease of the tax share of the instrument. The result shows that the pessimistic beliefs about the tax revenue are specific to the 100% tax treatment group.

The third contribution is that we find that participants expect the tax to be more effective in reducing the demand for the good causing the externality than the subsidy alone and the combinations of a tax and a subsidy. This contradicts some previous findings

[8] We designed the experiment to be able to detect a difference of 6 percentage points in public support between treatments. This number was chosen based on two considerations. First, Heres et al. (2017) found a 32 percentage points difference between a subsidy and a tax in their most relevant treatment, defining a likely upper limit for expected differences of 8 percentage points when we use a fractional design that introduces three intermediate policies. Second, differences of only a few percentage points are unlikely to be politically notable, and we therefore decided on a number somewhat below the percentage changes we could expect based on Heres et al. (2017).

[9] We do not include the revenue from the participant itself to mimic the real world setting where the revenue from the tax paid by the participant itself is a marginal contribution to the total tax revenue. As each group in our experiment consists of three members, $\frac{1}{3}$ is a large part of the total tax revenue, while in the real world the "group" typically consists of millions of people.

[10] Milkman et al. (2012) had far fewer participants (168), while Fesenfeld (2022) had 9115 participants.

Table 1
The value of each unit the participants can buy.

Number	Value of the unit
1.	130
2.	110
3.	90
4.	70
5.	50
6.	30

Table 2
Example of how the payoff is calculated in round 1. The example is based on all group members choosing five units.

	Element in the payoff	Example
	The value of the units the participant buys	130+110+90+70+50= 450
–	the price of the units	−40*5 = -200
–	the external costs from the purchase by the others in the group	−20*5*2= −200
=	Payoff	50

that people believe taxes not to be sufficiently effective in reducing demand, which has been found to be one of the main reasons why people oppose taxes (Douenne and Fabre, 2022; Bergquist et al., 2022).[11] We do not simply ask the participants whether they expect the policy to be effective. Instead we ask how many units they expect the other participants in their group to buy with and without the policy. The expectation about the effectiveness of the tax does not translate into support for the tax. By contrast, in the subsidy treatment, almost all who expect the policy to be effective also vote for the policy.

In the next section, we describe our experimental design, the theoretical predictions of what the participants will do, the experimental procedure, the sample and the balance tests. Then we analyse the findings, before the we discuss and conclude. We posted a pre-analysis plan on AEA Social Registry before the experiment started with RCT ID AEARCTR-0009099 and this can be found in Appendix D. Deviations from the plan are mentioned in the text and elaborated in Appendix E.

2. Methods

2.1. Experimental design

The experiment consists of a market round, a policy vote and then a second market round. In the market, participants decide how many units of a good to buy. Buying the good generates income for the participants, but also imposes an external cost on the other group members. In the first round there is no policy. Participants are then asked to vote for or against a policy proposal, randomly selected, that would internalize the external cost and incentivize all to limit the number of units bought. If the majority votes for the policy, the prices and the payoff structure in the second market round changes. The outcome we are interested in is whether the participants vote yes or no to the policy proposal. The experimental design is based on Kallbekken et al. (2011), Cherry et al. (2013), Cherry et al. (2014), Heres et al. (2017) and Cherry et al. (2017).

Each participant is part of a group with two other participants. All participants act as buyers in a market. Each participant can buy up to six units of the good. The price of each good is 40 tokens,[12] whereas the value of each unit differs. The first unit has the highest value and then each additional unit is worth less, mimicking declining marginal utility of consumption, see Table 1. The difference between the value and the price of each unit purchased generates income for the participant. For each unit a buyer in the group purchases, a cost of 20 tokens is imposed on each of the two other members of the group, meaning that the total external cost per unit is 40 tokens.

In the first round, the payoff is the value of the units purchased minus the price of the units and minus the external costs, see Table 2. At the end of the round, participants receive feedback on all the components of this payoff equation. Note that if someone has a negative payoff at the end of the experiment, the payoff is reset to 0.[13]

In order to reduce strategic behaviour within the group, the participants only have one round to get to know the market before they vote on a policy proposal. Further, we decided to nudge[14] the participants into choosing five units (the dominant strategy, see Section 2.2) in the first round by informing them (truthfully) that in a pilot we ran for the experiment, a majority chose five units.

[11] Dechezleprêtre et al. (2022) by contrast and in line with our study find that the majority of the respondents believe that a carbon tax would result in less driving.

[12] The currency used in the experiment is tokens, where 100 tokens equal £1.

[13] Only 15 participants, less than 1 percent of the total, had negative payoffs that had to be reset to 0 before calculating their total earnings.

[14] We knew that participants in similar experiments converged on the dominant strategy after playing the game for several rounds, but as each round is costly in terms of time and attrition, we decided to use a nudge to reduce the time taken to arrive at this outcome.

After experiencing this market for one round, participants are asked to vote on the rules that will govern the next round of the market. These rules vary across treatment groups. The participants' choices are to either (1) keep the rules as they were for the previous round, or (2) introduce a new specific policy.

The specific policy can either be a tax, a subsidy, or a fractional combination of the two. The option that receives the majority of votes (i.e., policy or no policy) will be implemented. Whether or not a participant votes for the proposed policy is the outcome variable. This a between-subjects design where participants are randomly allocated to one of five treatments. The five treatments are:

1. 100% tax: A tax of 40 tokens per unit.
2. 75% tax & 25% subsidy: A tax of 30 tokens per unit and a subsidy of 10 tokens per unit not purchased.[15]
3. 50% tax & 50% subsidy: A tax of 20 tokens per unit and a subsidy of 20 tokens per unit not purchased.
4. 25% tax & 75% subsidy: A tax of 10 tokens per unit and a subsidy of 30 tokens per unit not purchased.
5. 100% subsidy: A subsidy of 40 tokens per unit not purchased.

Whereas previous studies add policies in the policy package, our design using fractional combinations of two policies (tax and subsidy) is essential for keeping monetary payoffs for choosing the same behaviour identical across treatments. All policies fully internalize the marginal external cost (40 tokens per unit). The tax revenue generated from the purchases of each participant is shared equally among the two other participants. Similarly, the cost of funding the subsidy each participant in the group receives is shared equally among the two other participants, deducting the amount from their earnings. The participants are, however, not informed fully about how the tax revenues or the cost of the subsidy will be shared among the three group members (see below).

For all policies the payoff π for a participant is given by the following expression (cf. Appendix H):

$$\pi = \sum_{i=1}^{n} V_i - (p+q)n \tag{1}$$

where V_i is the gross value of good i, n is the number of goods purchased by the participant, p is the price, and q is the sum of the tax and the subsidy level in each treatment (remember that this sum is constant across treatment groups).[16] We see that the payoff is independent of which policy is implemented, and also of the number of goods bought by the other participants. With any policy implemented, the payoff depends only on the participant's own purchase. The intuition is that an additional unit bought by another participant implies more costs inflicted on the participant, but also a corresponding increase in tax revenue received or less subsidy to finance, which benefit the participant. These two effects cancel out as the externality is perfectly internalized through the tax and/or subsidy (with q equal to the total damage inflicted on other participants). It is also straightforward to show that the payoff in the second round is non-negative irrespective of how many units the participant buys.

When voting for policy or no policy, the participants are by design not fully informed about how the tax revenues will be distributed nor how the subsidy will be financed. This is the same design as in Heres et al. (2017) and resembles real world situations where tax revenue use and subsidy funding are rarely explicit. To avoid deception, we provided the following information before participants were asked to vote: "The tax generates revenue. The group's budget will be balanced through personal transfers of tokens between the members of the group". This is a crucial design choice, one that makes the experiment less transparent for participants, but one that is essential for retaining a property of taxes that is likely a core reason for tax aversion (as incorrect beliefs would be far less likely if we provided full information).[17]

After the votes have been cast and the number of units chosen, we inform participants about the distribution of the tax revenues and the financing of the subsidy payments.

After the vote, but before participants decide how many units to purchase, we elicit expectations in order to help uncover the mechanisms behind their voting decision. First we ask how many units they expect the other members in the group to buy with and without the policy implemented. Second, we ask what share of the tax revenue from the other group members the participant expects to receive, and/or what share of the subsidy cost for the other group members the participant expects to pay (depending on the treatment they are in). Finally, we ask whether the participant expects their payoff to increase, decrease or remain the same if the policy is implemented.

After the purchasing decisions are made in the second round, participants are asked: "Imagine that to combat climate change the government proposed to increase the cost of emitting CO_2 by £100 per ton from next year. This would increase the cost of petrol by 23 pence per litre and diesel by 26 pence per litre. If there was a vote on this tax proposal today, what would you have voted?" This question is to test whether voting for the tax in the experiment is correlated with expressing a willingness to vote for a hypothetical CO_2 tax.

Fig. 1 shows a timeline of the experiment. Screenshots of each page the participants see are found in Appendix G.

One factor which has been found to be important for the level of support for Pigouvian taxes is distributional effects (see for instance Andor et al. (2022) and Dechezleprêtre et al. (2022)). We wanted to investigate other factors important for attitudes

[15] To give subsidies for not buying or producing something is e.g. done within farming and foresting. For instance, the subsidy scheme REDD+ is paying for not deforesting forests. Another example is to subsidize no-till farming.

[16] In our experiment, $p = 40$ and $q = 40$, while V_i is shown in Table 1.

[17] For the subsidy treatment we write: "The subsidy costs money. Your group's budget will be balanced through personal transfers of tokens between the members of your group". For the three combinations we write: "The tax generates revenue and the subsidy costs money. Your group's budget will be balanced through personal transfers of tokens between the members of your group".

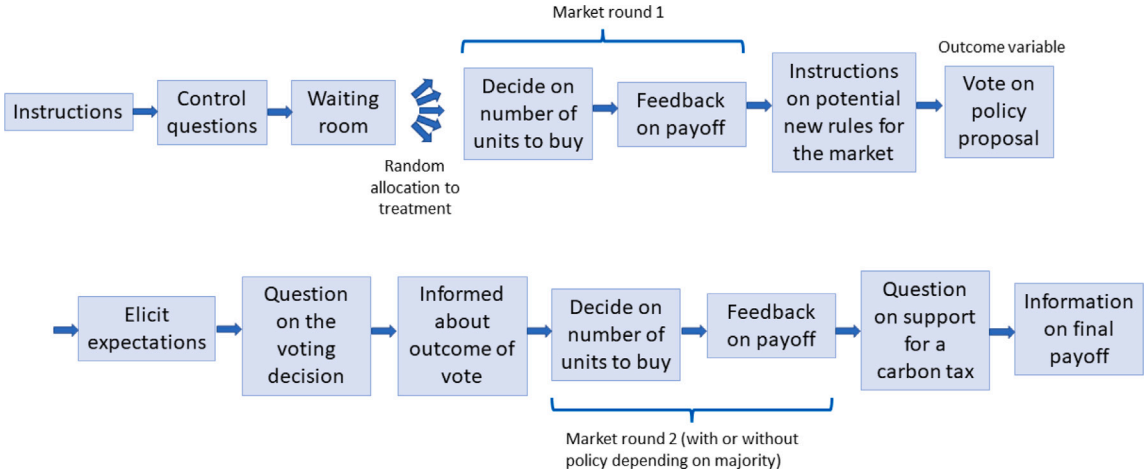

Fig. 1. The timeline of the experiment.

towards taxes and subsidies, and therefore we have designed an experiment in which there are no distributional effects (for the same behaviour). We keep the value of the units of the goods and the payoff structure identical for all participants within each group.

2.2. Theoretical predictions

For the individual participant, buying five units maximizes own payoff in the first round of the experiment, irrespective of how many units the other participants buy. We refer to this choice as the dominant strategy.[18] The socially optimal number of units to purchase is, however, three per participant.

The social optimum represents an efficiency gain over the market equilibrium if participants choose to buy fewer than five units. If all choose the social optimum of three units, total group payoff increases from 150 to 270 tokens. By reducing their purchases by a total of 6 units (2 units per person), the buyers forego profits of 120 tokens, but external costs are reduced by 240 tokens, yielding a net gain of 120 tokens.

If the policy is implemented, the payoff structure changes, and the participants are incentivized to reduce the number of units they purchase. With full information the dominant strategy is to purchase three units. The payoff structure with policies ensures that all participants receive the same payoff independent of the choices made by the other group members (see the payoff function in Section 2.1). Note, however, that participants beliefs about the policies will influence what they think is the profit-maximizing strategy, and these beliefs are explored in Section 3.3.

If the whole group chooses three units (the socially optimal number) in the first round, one could argue that there is no incentive to vote for a policy, and the incentive to vote for a policy is lower than if all group members chose 5 units in the first round. However, there is no guarantee that the other participants will continue choosing three units in the second round if a policy is *not* implemented, as the dominant strategy in the absence of a policy is to choose five units.[19]

2.3. Experimental procedures

We conducted an online interactive experiment on November 24th 2022, using the software Lioness lab (Giamattei et al., 2020). The participants were recruited from the online platform Prolific, a United Kingdom based firm that recruits participants for research.

Each participant is guaranteed to earn £1.5 if completing the experiment, in addition to the payoff. The median time to complete was just below 15 min, which means that the guaranteed payment equalled £6 per hour on average. The average payoff was £1.16, which is added to the guaranteed earnings. The payoff is based on the choices the participants make in the experiment, as explained in Section 2.1. The total average payment was thus almost £11 per hour.[20]

[18] When taking into account how choices in the first round may influence voting and the choice in the second round, it is not obvious that this is a dominant strategy. However, as we only had one round before the vote, choosing five units in the first round would most likely maximize final payoff, too.

[19] In the experiment, a majority of the participants chose five units in the first round, and the share of participants choosing five units is balanced across the five treatment groups (see Table A-1).

[20] There is a considerable literature on stake size, and a general conclusion is that stake size has little effect (see e.g. the meta-analysis by Larney et al. (2019)), and more specifically that online experiments with many participants and relatively low stakes reproduce the outcomes from classical lab experiments with fewer participants and higher stakes (Amir et al., 2012).

Table 3
Number of observations in each treatment group.

Treatment group	Number of observations
Tax	331
75% tax and 25% subsidy	323
50% tax and 50% subsidy	325
25% tax and 75% subsidy	328
Subsidy	334
Total	1641

Table 4
Descriptive statistics of the participants.

Characteristics	Average/Share	Number of observations
Age	39.5	1573
Female	50.0%	1579
Student status	11.4%	1641
Full-time employed	47.2%	1619
Part-time employed	15.7%	1619
Unemployed	4.6%	1619
Country of birth United Kingdom	80.2%	1641
Ethnicity white	83.2%	1619
Nationality United Kingdom	86.5%	1619
Language English	88.0%	1619

Notes: Some participants have missing values on some variables and therefore the total number of observations are not 1641 for all variables.

All interactions are anonymous. Before the participants can enter the experiment, they receive instructions and have to correctly answer two control questions to test that they understand the rules. They have three attempts and those that do not manage to answer correctly by the third attempt are screened out of the experiment.

Participants are assigned to a treatment group depending on the order in which they enter into the experiment after answering the control questions correctly. The first three participants entering are assigned to treatment 1, the next three participants entering are assigned to treatment 2 and so on. As several hundred participants take part simultaneously, and assignment to groups happens sequentially, the allocation to treatment is in effect random.

2.4. The sample

The sample consists of 1641 participants, all of them UK residents. Table 3 shows the number of observations in each of the five treatment groups. Table 4 displays the observational characteristics of the participants.[21]

We test whether the other treatment groups differ from the tax group (which we use as the "control" group) for the variables age and gender, in accordance with the pre-analysis plan. We also check whether the choices in round 1 of the experiment are balanced across treatments.

Table A-1 in Appendix A shows that there is balance between the treatment groups on the share of people choosing five units in the first round (the dominant strategy, see Section 2.2). This is reassuring.

Furthermore, for two of the treatment groups the average age is around 2 years younger than in the tax group. This is statistically significant, but it is not a large difference, and therefore we do not see it as a cause for concern. We check whether the age variable interacted with the treatment groups is statistically significant, see Table A-7. The coefficient is significant for the 50% tax & 50% subsidy treatment group interacted with age, but the size of the coefficient is only −0.005. For the other treatment groups the interaction term is not statistically significant. For gender there are no statistically significant differences between the tax treatment group and the other treatment groups.

We also test whether the treatment groups are different from each other, for instance if the 100% subsidy group is different from the 25% tax & 75% subsidy group on choosing five units in round 1, age and gender. They are not statistically significantly different, see Table A-2 in Appendix A.

3. Results

In this section we present the individual payoffs, the purchases the participants make, voting results, expectations about the policies, and the result of the test for external relevance.

[21] See Appendix I for a comparison between the observational characteristics of the participants and the UK population.

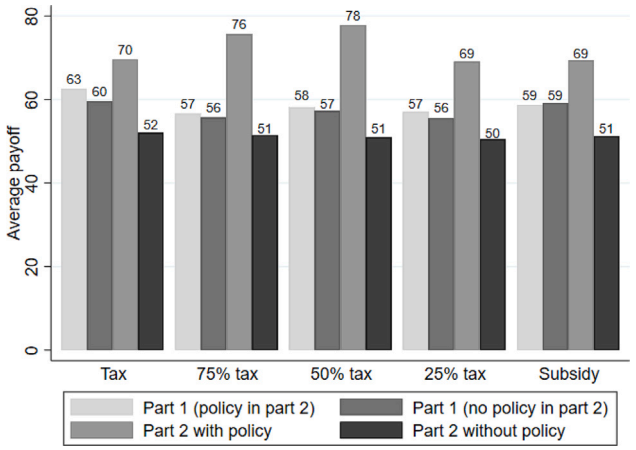

Fig. 2. The average individual payoff for the different groups in round 1 and round 2 with and without policy.
Notes: The 75% tax group also has 25% subsidy. The 50% tax group also has 50% subsidy. The 25% tax group also has 75% subsidy.

We analyse the results from the experiment using an OLS regression as generally recommended by Duflo et al. (2007). OLS coefficients are intuitive to interpret. As long as the probability is not close to 0 or 1, using OLS in combination with a binary outcome variable is regarded as unproblematic (Stock and Watson, 2015).

3.1. Individual payoffs and purchases

The average individual payoff is higher with policy than without, see Fig. 2. The payoff in round 1 for the groups where the majority voted to implement policy in round 2 is similar to the payoff in round 1 for the groups where majority voted against policy in round 2. The groups where the majority did not vote to implement policy experienced reduced payoff from round 1 to round 2, and this is consistent with learning how the game works and is also consistent with findings in previous experiments such as Cherry et al. (2013).

Participants buy on average 4.6 units in the first round, 3.5 units in the second round when a policy is implemented, and 4.8 units in the second round when no policy is implemented. See Figure A-1 in Appendix A for the average number of units by treatment and by whether the majority of the groups voted for policy. This means that all the policies reduced the demand significantly.

3.2. Support of the policy

We now turn to the main result of the experiment. The outcome variable is whether or not a participant votes for the proposed policy.[22] We test whether there is a difference in the level of support between tax, subsidy and combinations of tax and subsidy. We estimate the following regression equation:

$$v_i = \beta_0 + \beta_1 c3_i + \beta_2 c4_i + \beta_3 c5_i + \beta_4 s_i + u_i \tag{2}$$

v_i is a binary variable for whether the person voted for the policy or not, $c3_i - c5_i$ are binary variables for whether the participant is in the treatment group with respectively 75% tax & 25% subsidy, 50% tax & 50% subsidy, and 25% tax & 75% subsidy, and s_i is a binary variable for being in the treatment group with 100% subsidy.

The tax treatment group is the baseline, and all coefficients are compared to the level of support in the 100% tax treatment. The H_0-hypothesis (no difference between the treatment groups) is rejected, see Table 5.[23] We find that the support for policy increases with the subsidy share in the policy proposal.

In Table 5, column (2), we control for the payoff in the first part, as this could potentially influence voting behaviour. The differences in voting behaviour between the treatment groups change only trivially. The coefficient on the payoff is negative and statistically significant. Increasing the payoff in the first round by one unit, reduces the probability of voting for the policy by 0.1%. This is intuitive as increased payoff in the first round reduces the need for policy in the second round (if the participants expect the group members to purchase the same number of units in the second round as they did in the first round).

Fig. 3 shows the share of participants voting for the policy proposal by treatment. This illustrates visually what the coefficients of the regression analyses show: Support increases approximately linearly with the share of the subsidy in the policy proposal. The

[22] On average across treatments, only 36.2% of the participants voted for the policy, meaning that 63.8% voted to keep the rules as they were.

[23] To adjust for the fact that we test several hypotheses, we follow Fink et al. (2014) and use Benjamin–Hochberg adjusted p-values. This can be seen in Appendix C.

Table 5
Testing the difference in support for policy between the treatment groups.

	(1) Vote	(2) Vote
Subsidy	0.388***	0.387***
	(0.0339)	(0.0338)
25% tax and 75% subsidy	0.322***	0.317***
	(0.0341)	(0.0341)
50% tax and 50% subsidy	0.206***	0.203***
	(0.0334)	(0.0332)
75% tax and 25% subsidy	0.106**	0.101**
	(0.0317)	(0.0313)
Payoff in the first round		−0.001***
		(0.0004)
Constant	0.157***	0.246***
	(0.0200)	(0.0318)
Observations	1641	1641
R^2	0.086	0.095

Tax treatment group is the baseline. Robust standard errors in parentheses.
Column 2 was not mentioned in the pre-analysis plan, but requested by a reviewer.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

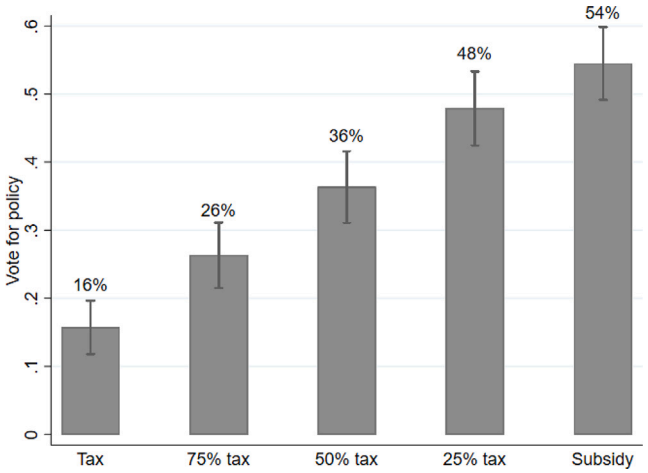

Fig. 3. The share voting for the policy in each treatment group.
Notes: The 75% tax group also has 25% subsidy. The 50% tax group also has 50% subsidy. The 25% tax group also has 75% subsidy. The lines are 95% confidence intervals.

effect of combining policy instruments is, in our experiment, close to a perfectly linear combination of the support for a tax alone and a subsidy alone.

Before the experiment was conducted, our hypothesis was that the support for the combinations of instruments would be somewhere between the support for the tax alone and the subsidy alone (see p. 4–5 in the pre-analysis plan in Appendix D). We did not have an a priori belief about whether the support would be different between having a 25% tax & 75% subsidy, 50% tax & 75% tax or a 75% tax & 25% subsidy. What we find is a clear difference in support between the combinations and that the support for the policy is increasing in the subsidy share. Further, both the support for policy of the 50% tax & 50% subsidy and the 75% tax & 25% subsidy groups are statistically significantly different from both the 100% tax and the 100% subsidy group.

3.3. Expectations about the policy

We elicited four different expectations about consequences of the policies. We use these to investigate mechanisms that can potentially explain why participants voted for or against the policy proposals.[24] We first test whether the treatment groups differ

[24] There is a slight deviation from the pre-analysis plan, see Appendix D. We have reformulated the regression equations. This is done to obtain a more relevant comparison group when testing whether the difference is statistically significant. The topics we test follow the pre-analysis plan, and all are included.

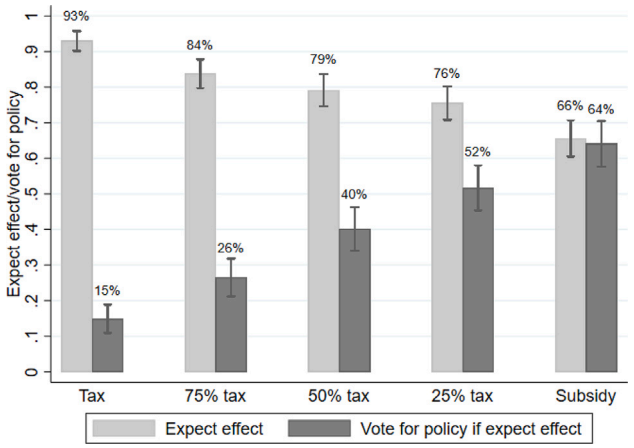

Fig. 4. The share expecting effect in each treatment group and the share voting for policy if they expect an effect.
Notes: The 75% tax group also has 25% subsidy. The 50% tax group also has 50% subsidy. The 25% tax group also has 75% subsidy. The lines display 95% confidence intervals. The voting of those that do not expect an effect can be seen in Table A-3.

in their expectations about how the policy works. Then we test whether the support for the policy differs between the treatment groups among those that hold the specific expectation and those that do not. At the end of the section, we do an exploratory analysis where we simulate the expected payoff based on the answers the participants gave.

3.3.1. Mechanisms: Expectations about the effect of the policy

To what extent participants expect the policy to reduce the demand differs across treatments. Results are shown in Fig. 4 and Table A-3, column 1. In the tax treatment group, 93% of participants expect a reduction in demand. In the subsidy group, 27 percentage points fewer (66%) expect a reduction in demand. In the combination treatments, expectations are in between those for the tax and subsidy treatments, and are all statistically different from the tax group.[25] This indicates that the type of policy instrument influences expectations about policy effectiveness. Keep in mind that the experiment is designed so that all policies provide the same incentives and should be equally effective in reducing demand.

Further, Fig. 4 and Table A-3, column 2 shows that the probability of voting for the policy among those who expect the policy to reduce demand increases with the share of subsidies in the policy proposal. Those who expect the policy to reduce demand have 49 percentage points higher probability of voting for the policy in the case of the subsidy compared to the tax. The results for the treatments with combinations of tax and subsidy lie in between the levels for the tax and subsidy treatments. The higher the tax share in the proposal, the higher the share of people expecting the policy to be effective, but also the lower the support for the policy. This indicates that expectations about effectiveness is not what drives policy support in our experiment.

For those that do not expect an effect, there is no statistically significant difference between the tax treatment group and the other treatment groups in voting behaviour, see Table A-3, column 3.

3.3.2. Mechanisms: Expectations about the tax revenue

Participants in the combination treatments expect to receive a higher share of the tax revenues than those in the 100% tax treatment, see Fig. 5 and Table A-4, column 1. The difference in expected share of revenues is not statistically different between the three combination treatments, but the combination treatments are all significantly different from the 100% tax treatment group. The experimental design and instructions should not give participants any reason to hold different expectations regarding the share of revenues shared across the tax and combination treatments. This shows pessimistic beliefs about how the tax works when it is implemented alone, compared to when the tax is combined with subsidies.

In addition, the share of voting for the policy among those who expect to receive a share of the revenue is increasing with the subsidy share in the policy proposal, from 48% voting for policy in the 25% tax & 75% subsidy treatment group to 18% voting for the policy in the 100% tax treatment (see Table A-4, column 2 and Figure A-2).

For those that do not expect to receive a share of the tax revenue, the combination treatment groups have a higher voting share for the policy than the 100% tax treatment groups, but the voting pattern is not linear (see Table A-4, column 3 and Figure A-2).

The reformulation of the regression equation follows the same pattern for all topics. This is explained in detail in Appendix D. The results from the regression equation in the pre-analysis plan can be seen in Appendix F.

[25] The 50% tax & 50% subsidy group is not significantly different from neither the 25% tax & 75% subsidy group nor the 75% tax & 25% subsidy group. The 25% tax & 75% subsidy group and 75% tax & 25% subsidy group are statistically different from each other (tests not shown).

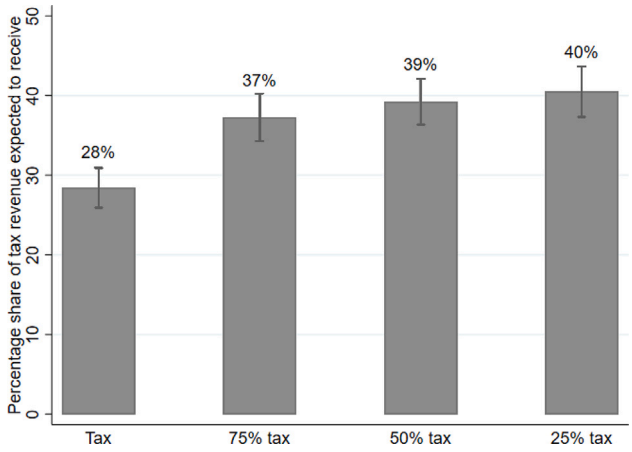

Fig. 5. The share of the tax revenue expected to receive.
Notes: The 75% tax group also has 25% subsidy. The 50% tax group also has 50% subsidy. The 25% tax group also has 75% subsidy. The lines display 95% confidence intervals. The difference between the tax group and the other groups is statistically significant.

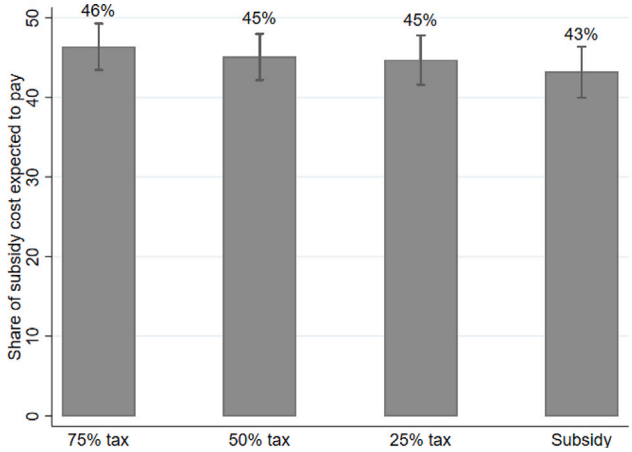

Fig. 6. The share of the subsidy cost expected to pay in the different treatment groups.
Notes: The 75% tax group also has 25% subsidy. The 50% tax group also has 50% subsidy. The 25% tax group also has 75% subsidy. The lines are 95% confidence intervals. The differences between the groups are not statistically significant.

Asking participants what share of the revenue they expect to receive might induce some people who otherwise would not consider that they might receive any of the tax revenue, to believe that they may do so. As a starting point, we placed the slider handle they use to indicate share of revenues they expect to receive in the middle (50%). This could influence participants to keep the slider closer to the middle than they would otherwise have done. However, asking this question cannot influence voting as it is asked after the votes have been cast. This caveat applies equally to all treatment groups, so even if the point estimates might be influenced by the starting point of the slider, the difference between the treatments should not be influenced.

3.3.3. Mechanisms: Expecting to pay for the subsidy cost

The average share of the subsidy cost participants expect to pay is 43% in the subsidy treatment group, and expectations are not significantly different from this in the combination treatments, see Fig. 6 and in Table A-5, column 1.

In addition, 80% of those in the subsidy group who do *not* expect to pay for the subsidy, voted for the subsidy, see Figure A-3 and Table A-5, column 2. For the combination groups, the probability of voting for the policy declines with the tax share going up, and the difference compared to the subsidy is statistically significant for the 75% tax & 25% subsidy and the 50% tax & 50% subsidy groups, see Table A-5, column 2. For those expecting to pay a share of the subsidy cost, the support for the policy is lower the higher the tax share in the policy, see Table A-5, column 3 and Figure A-3.

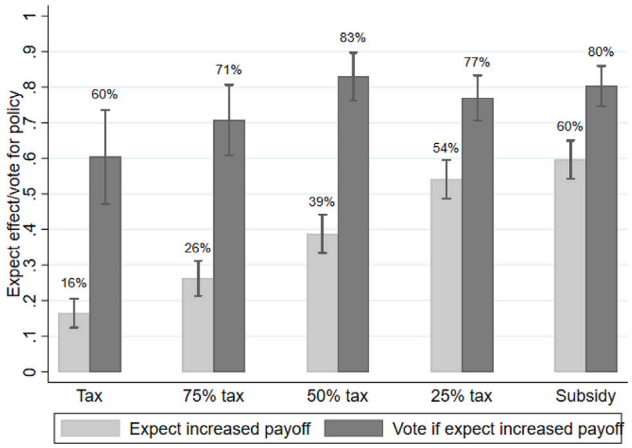

Fig. 7. The share expecting the payoff to increase if policy is implemented in the different treatment groups and the voting behaviour for those that expect increased payoff.
Notes: The 75% tax group also has 25% subsidy. The 50% tax group also has 50% subsidy. The 25% tax group also has 75% subsidy. The lines are 95% confidence intervals. The voting of those that do not expect the payoff to increase can be seen in Table A-6.

3.3.4. Mechanisms: Expectations about the payoff

Only 17% of participants in the tax treatment group expect that the policy will increase their payoff, whereas 60% in the subsidy treatment group do so, see Fig. 7 and Table A-6, column 1.[26] The share of participants in the combination treatment groups who expect that the policy will increase their payoff, is increasing in the subsidy share in the policy proposal. This shows again pessimistic beliefs for the 100% tax group.

Among those who expect the policy proposal to increase payoffs, the share voting in favour of the policy is higher for the combination treatment groups than for the tax treatment group, see Fig. 7 and Table A-6, column 2. The support for policy does not, however, increase linearly with the subsidy share, as we have seen earlier, see Fig. 7. The share supporting the policy among those who expect the policy proposal to increase payoffs, does not significantly differ between the tax treatment group and the 75% tax & 25% subsidy group, but for the two other combination groups the support is statistically significantly different, see Fig. 7 and Table A-6, column 2.

In general, expecting the policy to increase payoff seems to be a clear predictor of voting behaviour, but only 77% of those who believe the payoff will increase, vote for the policy (not shown). This is not even across treatments, as can be seen in Fig. 7. Furthermore, only 39% of the participants expect that the policy will increase their payoff.[27] This result leads us to investigate the relationship between expectations about tax revenue and subsidy cost and expectations about payoff, see Section 3.3.5.

3.3.5. Exploratory analysis: Simulating the expected payoff

We use participants' expectations about policy effectiveness, tax revenue and the subsidy cost to explore if participants expect the payoff to be higher with the policy. As this was not part of our pre-analysis plan, it is an exploratory analysis. This differs from simply asking the participants whether they expect the payoff to increase (which we also did, see Section 3.3.4). Here, we instead use the expectations about how many units the others in the groups would buy with or without the policy, what share of the tax revenue one expects to receive, and how much of the subsidy cost one expects to pay, to calculate the expected payoff with and without the proposed policy.[28]

Fig. 8 shows the share expecting increased payoff with policy based on our calculation, and we can compare it with the share voting for the policy, i.e., the same as in Fig. 3. We see that participants in the tax group seem to vote according to their payoff expectations, and as noted above these expectations are pessimistic compared to the groups where tax and subsidy are combined.

For the combination groups the number of participants who expect the payoff to increase with policy is much higher than the number of participants that actually voted for the policy proposal. This may indicate either that the expectations elicited after the

[26] For both the subsidy and the 25% subsidy & 75% tax group, a higher share of participants expect their payoff to increase if policy is implemented (60% for the subsidy group, see Table A-6) than the share of participants who actually voted for the subsidy (55% for the subsidy group, see Table 5). The reason for this may be that asking questions about the expectations can change participants' thinking about the policies by making certain aspects more salient, or through experimenter demand effects.

[27] 8% believe the policy will not change the payoff and 53% believe the payoff will be reduced with policy. In Figure A-4 we show the share expecting the payoff to decrease if policy is implemented in the different treatment groups and their voting behaviour. In Figure A-5 the share expecting no change in the payoff if policy is implemented in the different treatment groups and their voting behaviour are shown.

[28] We need to make some assumptions to do this simulation. When policy is not implemented in the second round, we do not know how many units the participant would have chosen to buy with policy. Then we assume that the participant would buy as many units as (s)he expect his/her group members to choose with policy. To calculate the payoff without policy, we use the purchases in the first round. Then we calculate the difference in expected payoff with and without policy.

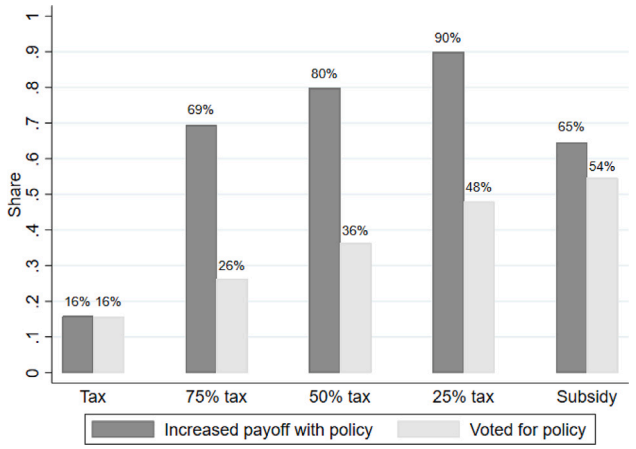

Fig. 8. The simulated share expecting increased payoff based on the elements in the payoff (reduced externality because of reduced demand, a share of the tax revenue and a share of the subsidy cost) compared to the voting behaviour in the different treatment groups.
Notes: The 75% tax group also has 25% subsidy. The 50% tax group also has 50% subsidy. The 25% tax group also has 75% subsidy.

Table 6
Relationship between voting for a tax in the experiment and a hypothetical carbon tax.

	(1) Voting for tax
Carbon tax	0.0836
	(0.0470)
Constant	0.130***
	(0.0239)
Observations	303
R^2	0.012

The sample is only the tax treatment group.
Robust standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

voting were not clear for the participants when they voted, or that something other than the payoff expectations were driving the voting behaviour.

3.4. Test for external relevance

We test whether voting for the tax in the experiment is correlated with expressing a willingness to vote for a hypothetical CO_2 tax. The result can be seen in Table 6.

Support for the tax proposal in the experiment correlates with support for the hypothetical CO_2 tax on a 10% significance level. Note, however, that only 48 participants (16%) voted for the tax in the experiment. The support for a hypothetical CO_2 tax is higher (35% of all participants and 34% of the tax group) than the support for a tax in the experiment (16%).

4. Discussion

We find that support for a combination of tax and subsidy approximately equals the simple average of support for the two instruments alone. This main result appears to contradict Milkman et al. (2012) and Fesenfeld (2022), who find that policy packaging increases support (beyond the averaging of support for its constituent parts). Our result also stand in contrast to Eriksson et al. (2008), who find a negative effect of combining policies. It is, however, not possible to make clean comparisons because of important differences in methods and design between the studies. First, our study is incentivized, i.e., participants' choices have real financial consequences for them, whereas none of the three other studies are.

When comparing our results with Milkman et al. (2012), it is important to note that our experimental design keeps payoffs equal across treatments when a policy is implemented, and independent of the choices made by the other group members. Further, we use fractional combinations where we reduce the tax rate to the same extent as we increase the subsidy rate, whereas in Milkman et al. (2012) the stated gain varies across the policies. The gain from the policy in their study is higher when the two bills are combined

than when each bill is considered separately. This difference can potentially explain the difference in voting outcome between our study and Milkman et al. (2012). Another important difference is that whereas we study policy instruments in a non-contextualized lab experiment rather than attitudes towards potential public policies. Still, both studies consider bundling of policies that are often viewed as respectively desirable and undesirable. Milkman et al. (2012) find that the reason for the increased support for the policy bundle is that "policy bundling reduces the salience of losses (...) and heightens the salience of gains".

When comparing our results with Fesenfeld (2022), it is important to note that the payoffs from the policies in his study are neither stated (as in Milkman et al. (2012)), nor set to be equal by design (as in our experiment, albeit not known to the participants when they vote), but left open for respondents to consider themselves. The core idea explored is how greater complexity influences policy perceptions (including public support). The choice experiment in Fesenfeld (2022) has four policy attributes that vary simultaneously and target different behaviours, e.g., taxes for consumers combined with emissions standards for producers. This differs crucially from our experiment where the two instruments target the same behaviour.

In Eriksson et al. (2008) all participants were presented with all policy options (single instruments first and then the packages), and the costs and benefits are not clearly stated nor kept the same across instruments: For the instruments by themselves the specific tax or subsidy rates were not provided, whereas the rates were stated for the policy packages. This latter difference could to some extent explain the result that public support for the packages is closest to the level of support for the least popular instrument: Making policy proposals more specific, e.g., by stating the tax rate, may reduce support and this is only done for the policy packages in Eriksson et al. (2008).

Our findings are generally consistent with Heres et al. (2017), another lab experiment with a market with negative externalities, whose main finding is that subsidies are substantially more popular than taxes, even when payoff is kept constant across policies (if participants choose the dominant strategy). Further, Heres et al. (2017) state that this can in part "be explained by the participants' expectation that the subsidy will increase their own payoffs more than a tax, but not because it is expected to be more effective in changing behavior", which is similar to our findings. We introduced the same vagueness regarding the distribution of tax revenues and subsidy funding as Heres et al. (2017), and the (intentional) asymmetry this creates may be an important explanation for the differences in support between taxes and subsidies: The uncertainty for a participant regarding what share of the tax revenue (s)he will receive can be seen as a potential loss (e.g., expecting to receive no share of the revenues, or a smaller share than one's own tax payment), whereas the uncertainty related to paying the subsidy cost can be seen as a potential gain (e.g., expecting not to have to help fund the subsidy, or to fund less than one receives). Thus, the difference in support to the different instruments that both Heres et al. (2017) and we find, can potentially be related to loss aversion. Their results are also useful for understanding what would likely happen had we included full information about what happens to tax revenues and how the subsidy is funded; it likely would have increased support for the tax more than for the subsidy.

Status quo bias might explain why many do not support the proposed policies even when they are designed to increase individual and group payoffs (Kahneman et al., 1991; Samuelson and Zeckhauser, 1988), but it does not explain the difference in support between the policies. There seems to be a "broader aversion to market intervention", in line with the findings in Cherry et al. (2012).

Another explanation might be that participants view taxes as a more coercive instrument that reduces their own freedom to buy a "dirty" good, rather than an instrument aiming to reduce others' incentives to buy the same "dirty" good (Cherry et al., 2012). Even though the latter effect may be more significant in terms of total welfare effect, the former effect may be more visible or salient to the participant. This, however, needs further investigation. Interestingly, Dechezleprêtre et al. (2022) find that respondents rank a carbon tax as the most costly climate policy, followed by investments in green infrastructure and a ban on combustion-engine vehicles.

Do participants understand the incentive structure fully? One reason participants vote as they do could be that they do not fully understand or take into account the payoff structure. Kallbekken et al. (2011) investigated how much the participants understand and whether more information about how Pigouvian taxes work influenced support. They find that with more information the participants understand more, but it does not change the support for taxes by much. Still, the participants might not take the revenue from the tax and the financing of the subsidy sufficiently into account. Our experimental design does not make it possible to disentangle whether the participants understand the whole payoff structure. Instead, we build on the findings of Kallbekken et al. (2011), where they test the understanding of the payoff structure, and on Heres et al. (2017) where the role of budgetary information is investigated.

5. Conclusion

In this study, we have conducted an online, non-contextualized and interactive lab experiment to explore support for taxes, subsidies, and combinations of the two instruments. We find that support increases approximately linearly with the share of subsidies in the policy proposal. This finding questions the claim, based on findings in previous studies, that policy packaging can increase public support for unpopular policies. However, given the design of the previous studies, where instrument type and benefits vary together, the findings in those studies might relate to the policy package increasing gains and not the act of packaging as such.

We find that people hold pessimistic beliefs regarding taxes, specifically that they expect to receive a smaller share of the tax revenue and to a lesser degree expect taxes to increase payoffs compared to the other policies in the experiment. This pessimistic belief can *explain* the tax aversion, but our findings imply that combining a Pigouvian tax with a subsidy does not help *reduce* tax aversion as such: Support follows the share of the tax in the fractional combinations; the share of participants who expect the payoff to increase with the policy increases only in linear proportion with the subsidy (non-tax) share of the policy, and the belief that taxes

are (more) effective does not translate into policy support. The only aspect where combining instruments can be said to influence (an aspect of) tax aversion, is that when combined with a subsidy, respondents expect a larger share of the revenues to be returned to themselves.

The discrepancy between our finding and previous findings on taxes and beliefs about effectiveness raises a question about the causal direction: Do people oppose taxes because they think they are not effective, or do people answer in surveys that taxes are not effective because they do not want taxes? Our findings contrast with what is common in the literature as we find that people both oppose taxes *and* think they are effective. However, the elasticity of demand for the good in this experiment is probably higher than for instance fuel in real-life and this can therefore also partly explain the difference in findings. Further investigation on the causal direction between attitudes towards taxes and beliefs about the effectiveness is clearly warranted.

Several extensions of the experiment would be valuable. We chose a design where the packages are fractional combinations of a tax and a subsidy. It would be informative to compare this design to an additive design where the tax remains constant but different policies are added to it in a policy package.[29] Another refinement of our design would be to ask participants if they want to change their vote after eliciting their expectations, as the act of eliciting the preferences may change their thinking about the policies. In addition, it could be that combining Pigouvian taxes with other kinds of instruments than subsidies in a policy package would yield different results.

Whether the results hold outside of lab experiments is a question for further investigation. Levitt and List (2007, p.168) underline that "many real-world markets operate in ways that make pro-social behaviour much less likely" than in a lab experiment. It has been investigated to what degree some types of lab experiments find the same results in real-world settings, such as reaction to competition (Buser and Yuan, 2019). The type of lab experiment we are using has not been investigated for validity outside of the lab, and this is a point for further investigation.

Policy packaging may make sense for a number of reasons, including enhancing effectiveness (Van den Bergh et al., 2021), addressing distributional concerns (Bouma et al., 2019), eliciting information about firms' abatement cost (Ambec and Coria, 2021), or directing technological change (Acemoglu et al., 2012). However, unlike what some previous studies have indicated, this study does not find that packaging policies increases the level of public support beyond the simple averaging of support for the constituent parts of the package.

CRediT authorship contribution statement

Gøril L. Andreassen: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Steffen Kallbekken:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Knut Einar Rosendahl:** Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Gøril L. Andreassen has been employed by the Norwegian environmental NGO ZERO from 2006-2014, working with climate policy. Since 2014, she has had no formal relationship with the NGO. From 2014-2016 she was employed by the Norwegian Rainforest Foundation, also working with climate policy. Since 2016, she has had no formal relationship with the NGO. From 30th of May 2022-14 th of February 2023 she was Board of Director of the investment company Saga Pure. This company invested in green technology companies. She declares that she has no other relevant or material financial interest that relate to the research described in this paper. No party had the right to review this paper prior to the submission.

Steffen Kallbekken holds a 30% position for the Ministry of Climate and Environment in Norway, drafting text on climate policy for the 2050 Climate Commission – an independent expert committee appointed by the government of Norway.

Knut Einar Rosendahl declares no known interests related to the submitted manuscript.

Acknowledgments

We would like to thank Frode Alfnes, Arild Angelsen, Kjetil Bjorvatn, Julie Brun Bjørkheim, Askill Harkjerr Halse, Karen Evelyn Hauge, Julia Del Carmen Naime Sanchez Henkel, Tora Knutsen, Andreas Kotsadam, Magnus Merkle, Marit Sandstad, Siv-Elisabeth Skjelbred, Håkon Sælen, and the research group KREM at the School of Economics and Business at NMBU for valuable input and comments.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeem.2024.103010.

[29] However, then the payoff across treatments are more difficult to keep constant.

References

Acemoglu, D., Aghion, P., Bursztyn, L., Hemous, D., 2012. The environment and directed technical change. Am. Econ. Rev. 102 (1), 131–166.

Ambec, S., Coria, J., 2021. The informational value of environmental taxes. J. Public Econ. 199, 104439.

Amir, O., Rand, D.G., Gal, Y.K., 2012. Economic games on the internet: The effect of 1 stakes. PLoS One 7 (2), e31461.

Anderson, S.T., Marinescu, I., Shor, B., 2023. Can Pigou at the Polls Stop Us Melting the Poles? J. Assoc. Environ. Resour. Econ..

Andor, M.A., Lange, A., Sommer, S., 2022. Fairness and the support of redistributive environmental policies. J. Environ. Econ. Manag. 114, 102682.

Baranzini, A., Carattini, S., 2017. Effectiveness, earmarking and labeling: testing the acceptability of carbon taxes with survey data. Environ. Econ. Policy Stud. 19, 197–227.

Bergquist, M., Nilsson, A., Harring, N., Jagers, S.C., 2022. Meta-analyses of fifteen determinants of public opinion about climate change taxes and laws. Nature Clim. Change 12 (3), 235–240.

Bouma, J., Verbraak, M., Dietz, F., Brouwer, R., 2019. Policy mix: mess or merit? J. Environ. Econ. Policy 8 (1), 32–47.

Buser, T., Yuan, H., 2019. Do women give up competing more easily? Evidence from the lab and the Dutch math olympiad. Am. Econ. J.: Appl. Econ. 11 (3), 225–252.

Cherry, T.L., Kallbekken, S., Kroll, S., 2012. The acceptability of efficiency-enhancing environmental taxes, subsidies and regulation: An experimental investigation. Environ. Sci. Policy 16, 90–96.

Cherry, T.L., Kallbekken, S., Kroll, S., 2014. The impact of trial runs on the acceptability of environmental taxes: Experimental evidence. Resour. Energy Econ. 38, 84–95.

Cherry, T.L., Kallbekken, S., Kroll, S., 2017. Accepting market failure: Cultural worldviews and the opposition to corrective environmental policies. J. Environ. Econ. Manag. 85, 193–204.

Cherry, T.L., Kallbekken, S., Kroll, S., McEvoy, D.M., 2013. Cooperation in and out of markets: An experimental comparison of public good games and markets with externalities. Econom. Lett. 120 (1), 93–96.

Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Chico, A.S., Stantcheva, S., 2022. Fighting Climate Change: International Attitudes Toward Climate Policies. National Bureau of Economic Research.

Douenne, T., Fabre, A., 2022. Yellow vests, pessimistic beliefs, and carbon tax aversion. Am. Econ. J.: Econ. Policy 14 (1), 81–110.

Drews, S., Van den Bergh, J.C., 2016. What explains public support for climate policies? A review of empirical and experimental studies. Clim. Policy 16 (7), 855–876.

Duflo, E., Glennerster, R., Kremer, M., 2007. Using randomization in development economics research: A toolkit. Handb. Dev. Econ. 4, 3895–3962.

Eriksson, L., Garvill, J., Nordlund, A.M., 2008. Acceptability of single and combined transport policy measures: The importance of environmental and policy specific beliefs. Transp. Res. Part A: Policy Pract. 42 (8), 1117–1128.

Fesenfeld, L.P., 2022. The effects of policy design complexity on public support for climate policy. Behav. Public Policy 1–26.

Fink, G., McConnell, M., Vollmer, S., 2014. Testing for heterogeneous treatment effects in experimental data: false discovery risks and correction procedures. J. Dev. Effect. 6 (1), 44–57.

Giamattei, M., Yahosseini, K.S., Gächter, S., Molleman, L., 2020. LIONESS lab: A free web-based platform for conducting interactive experiments online. J. Econ. Sci. Assoc. http://dx.doi.org/10.1007/s40881-020-00087-0.

Givoni, M., Macmillen, J., Banister, D., Feitelson, E., 2013. From policy measures to policy packages. Transp. Rev. 33 (1), 1–20.

Gugler, K., Haxhimusa, A., Liebensteiner, M., 2021. Effectiveness of climate policies: Carbon pricing vs. subsidizing renewables. J. Environ. Econ. Manag. 106, 102405.

Hardisty, D.J., Johnson, E.J., Weber, E.U., 2010. A dirty word or a dirty world? Attribute framing, political affiliation, and query theory. Psychol. Sci. 21 (1), 86–92.

Helm, C., Mier, M., 2021. Steering the energy transition in a world of intermittent electricity supply: Optimal subsidies and taxes for renewables and storage. J. Environ. Econ. Manag. 109, 102497.

Heres, D.R., Kallbekken, S., Galarraga, I., 2017. The role of budgetary information in the preference for externality-correcting subsidies over taxes: A lab experiment on public support. Environ. Resour. Econ. 66 (1), 1–15.

Kahneman, D., Knetsch, J.L., Thaler, R.H., 1991. The endowment effect, loss aversion, and status quo bias. J. Econ. Perspect. 5 (1), 193–206.

Kallbekken, S., Kroll, S., Cherry, T.L., 2011. Do you not like Pigou, or do you not understand him? Tax aversion and revenue recycling in the lab. J. Environ. Econ. Manag. 62 (1), 53–64.

Kallbekken, S., Sælen, H., 2011. Public acceptance for environmental taxes: Self-interest, environmental and distributional concerns. Energy Policy 39 (5), 2966–2973.

Kern, F., Rogge, K.S., Howlett, M., 2019. Policy mixes for sustainability transitions: New approaches and insights through bridging innovation and policy studies. Res. Policy 48 (10), 103832.

Larney, A., Rotella, A., Barclay, P., 2019. Stake size effects in ultimatum game and dictator game offers: A meta-analysis. Organ. Behav. Hum. Decis. Processes 151, 61–72.

Levitt, S.D., List, J.A., 2007. What do laboratory experiments measuring social preferences reveal about the real world? J. Econ. Perspect. 21 (2), 153–174.

Milkman, K.L., Mazza, M.C., Shu, L.L., Tsay, C.-J., Bazerman, M.H., 2012. Policy bundling to overcome loss aversion: A method for improving legislative outcomes. Organ. Behav. Hum. Decis. Processes 117 (1), 158–167.

Parry, I.W., Black, S., Vernon, N., 2021. Still not getting energy prices right: A global and country update of fossil fuel subsidies. IMF. Working Paper No. 2021/236.

Samuelson, W., Zeckhauser, R., 1988. Status quo bias in decision making. J. Risk Uncertain. 1, 7–59.

Schuitema, G., Steg, L., Forward, S., 2010. Explaining differences in acceptability before and acceptance after the implementation of a congestion charge in Stockholm. Transp. Res. Part A: Policy Pract. 44 (2), 99–109.

Stock, J.H., Watson, M.W., 2015. Introduction to Econometrics, 3rd ed. Pearson Education.

The World Bank, 2022. Carbon pricing dashboard. URL https://carbonpricingdashboard.worldbank.org/. (Accessed 10 January 2023).

Timilsina, G.R., 2022. Carbon taxes. J. Econ. Lit. 60 (4), 1456–1502.

Umit, R., Schaffer, L.M., 2020. Attitudes towards carbon taxes across Europe: The role of perceived uncertainty and self-interest. Energy Policy 140, 111385.

Van den Bergh, J., Castro, J., Drews, S., Exadaktylos, F., Foramitti, J., Klein, F., Konc, T., Savin, I., 2021. Designing an effective climate-policy mix: accounting for instrument synergy. Clim. Policy 21 (6), 745–764.

Winslott-Hiselius, L., Brundell-Freij, K., Vagland, Å., Byström, C., 2009. The development of public attitudes towards the stockholm congestion trial. Transp. Res. Part A: Policy Pract. 43 (3), 269–282.