TI 2019-024/V
Tinbergen Institute Discussion Paper

An introduction to Participatory Value Evaluation

Revision: December 2019

Niek Mouter[1]
Paul Koster[2]
Thijs Dekker[3]

[1] Delft University of Technology Faculty of Technology, Policy and Management, Transport and Logistics Group
[2] Vrije Universiteit Amsterdam, School of Business and Economics, Department of Spatial Economics and Faculty of Humanities, John Stuart Mill College
[3] Institute for Transport Studies Leeds and Choice Modelling Centre, University of Leeds

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at https://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

An introduction to Participatory Value Evaluation

Abstract:

Participatory Value Evaluation (PVE) is an evaluation method which infers the social welfare effects of public policies by eliciting individuals' preferences over the allocation of public budgets as well as their private income. In a PVE, individuals are asked to choose the best portfolio of projects subject to governmental and private budget constraints. This paper explains that PVE gives individuals the opportunity to express a broader range of preferences concerning (impacts of) government policies compared to existing valuation methods. The paper illustrates PVE with a case study on projects to mitigate flood risks in which 2,900 Dutch citizens participated.

Keywords:

Participatory Value Evaluation; Cost-Benefit Analysis; Benefit-Cost Analysis; Citizen Participation; Valuation Methods.

1. Introduction

In virtually all western countries, Cost-Benefit Analysis (CBA) is nowadays considered the gold standard for supporting public decision-making (e.g. Boardman et al. 2013). In these countries CBA is mandatory when national funding is required for large transport projects (Mackie et al., 2014). Furthermore, CBA is widely applied to governmental decisions on environmental, health and safety regulation (e.g. Hahn and Tetlock, 2008) and the instrument is also adopted in other policy domains, examples being energy and water management (e.g. Persky, 2001). The theoretical

foundations of CBA are rooted in welfare economics which is a branch of economics that investigates the social desirability of alternative economic outcomes (e.g. Boadway and Bruce, 1984). A CBA is built on the Kaldor-Hicks efficiency criterion (e.g. Boadway, 2006), which recommends projects where the sum of monetary gains outweigh the sum of monetary losses and winners can potentially compensate the losers. The conversion of positive and negative social impacts of government projects into monetary units relies on willingness to pay (WTP) estimates.

The WTP valuation paradigm is fiercely criticized in the literature (e.g. Ackerman and Heinzerling, 2004; Sagoff, 1988). Several scholars argue that this paradigm takes a too narrow perspective when evaluating government projects because choices individuals make with their private income might not accurately reflect their preferences towards public policy (e.g. Ackerman and Heinzerling, 2004; Sen, 1995). To resolve this issue, scholars developed so-called willingness to allocate public budget experiments (WTAPB) in which individuals make choices when faced with effects accruing from alternative allocations of *government budget* (e.g. Johansson-Stenman and Martinsen, 2008; Mouter et al., 2017). The WTAPB approach aims to infer welfare effects of government projects from individuals' preferences regarding the expenditure of public euros. However, one clear downside of the WTAPB approach is that respondents are forced to make a choice between two or three alternative allocations of public budget. Hence, preferences of individuals who believe that it is better to do nothing and/or reduce taxes instead of allocating public budget to one of the proposed projects are not respected.

Participatory Value Evaluation (PVE) is an evaluation method which expands the WTAPB valuation paradigm. The similarity between WTAPB experiments and PVE experiments is that participants are asked to express to policy makers which projects should be financed from a limited amount of public budget. The most important difference is that participants in a PVE have the

option to advise the government against allocating the budget to any (or some) of the projects that are considered in the PVE and shift the remaining budget to the next year. In addition to the experimental setting in which individuals are asked to advise the government on the allocation of an earmarked public budget (the 'fixed budget PVE' format), we develop an experimental setting in which participants are also allowed to adjust the size of the public budget (and thereby their after tax income) by changing taxes (the 'flexible budget PVE' format). The flexible budget PVE format, for instance, gives individuals the opportunity to express the view that it is better to reduce taxes and respectively increase private income instead of choosing to fund any of the government projects considered in the PVE.

The conceptual innovation of the introduction of the flexible budget PVE format is that the WTP valuation paradigm and the WTAPB valuation paradigm are integrated in one valuation framework. That is, the social welfare effects of government projects are established through the elicitation of individuals' preferences over the allocation of public budgets as well as their private income. Dekker et al. (2019) developed the econometric and microeconomic framework which allows for a welfare analysis of government projects that are included in fixed and flexible budget PVE experiments. The economic framework of PVE rests on the microeconomic principle that individuals maximise utility subject to private and public budget constraints.

The present paper aims to introduce PVE in a non-technical way to a broad readership of economists and policy makers working at local, regional or national governments. Firstly, we aim to increase understanding of why PVE was developed by discussing shortcomings of other valuation methods (Section 2). We pay specific attention to the fact that PVE gives individuals the opportunity to express a broader range of preferences compared to valuation methods that are solely grounded in the WTP or the WTAPB valuation paradigm. Secondly, we aim to familiarise

economists and policy makers with PVE by presenting a real-life case study in which the social welfare effects of projects to mitigate flood risks in the Netherlands are evaluated (Section 3). To make the paper accessible for a broad readership, we refer to the supplementary material and the Dekker et al. (2019) paper for technical details. Thirdly, Section 3 also presents experiences from participants in the PVE experiment and policy makers who commissioned the case study. Finally, we argue in Section 4 that PVE is a promising evaluation method which warrants further development in future research projects. Notably, we do wish to emphasize that the purpose of this study is not to claim that PVE is superior to other valuation approaches. More research is needed to make any inference about whether PVE is (or is not) a more appropriate method for the valuation of public goods compared to other approaches.

2. Participatory Value Evaluation compared with other valuation methods

2.1 Private WTP as the primary measure of value in CBA

A CBA expresses the social costs and benefits of government projects in monetary terms based on the private WTP of affected individuals and parties (Persky, 2001). Ideally, analysts derive such WTP estimates directly from market behaviour. Impacts of government projects are, amongst other things, often evaluated through investigating the private decisions people make when buying a house. For example, Koster and van Ommeren (2015) estimate the impact of earthquakes on housing prices and use these estimates to draw conclusions about the non-monetary costs of being exposed to the risk of earthquakes resulting from gas extraction. In cases where market behaviour is absent, monetary impacts can be inferred from hypothetical consumer choices. For instance, the conventional empirical approach used to infer the value of travel time savings accruing from

government projects relies on (hypothetical) route choice experiments (e.g. Abrantes and Wardman, 2011). In these experiments, respondents are asked to make a series of private choices between routes which differ in terms of travel time and travel costs.

The most prominent critique concerning the evaluation of government projects through the number of euros that individuals are willing to pay in (hypothetical) consumption choices (henceforth: the private WTP approach) is that individuals' consumption choices may not reflect how they want public policies to change (e.g. Ackerman and Heinzerling, 2004; Lusk and Norwood, 2011; Sen, 1995). For instance, in the case of animal welfare, people may not be willing to contribute individually to the public good because, in their view, the impact of their individual contribution is negligible. People may, however, be willing to contribute when the whole community contributes because the impacts of coordinated efforts can be substantial (Ivehammar, 2009; Sen, 1995). An example is provided by Lusk and Norwood (2011) who describe how 63% of Californians voted in favour of a ballot prohibiting battery-produced eggs, which at the time of the vote were the most popular type of eggs purchased and consumed in California. Scholars also argue that individuals' private consumer choices might not reflect their preferences towards public policy because moral considerations might be more salient in the latter context (e.g. Sagoff, 1988). Sagoff (1988, p. 48) asserts that: "*many of us are concerned, for example, that the workplace be safe and free of carcinogens; we may share this conviction, even if we are not workers. And so, we might favour laws that require very high air-quality standards in petrochemical plants. But as consumers, we may find no way to support the cause of workplace safety. Indeed, if we buy the cheapest products, we may defeat it. We may be concerned as citizens, or as members of a moral and political community, with all sorts of values – sentimental, historical, ideological, cultural, aesthetic, and ethical – that conflict with the interests we reveal as consumers, buying shoes and*

choosing tomatoes. The conflict within individuals, rather than between them, may be a very common conflict."

2.2 From private WTP to collective WTP

In an attempt to ameliorate this issue, impacts of government projects have been evaluated through collective WTP experiments. These experiments express the impacts of government projects to the entire community but are financed, for example, by a uniform tax increase or alternative but comparable payment vehicle. Nyborg (2000, pp. 311) summarises the underlying valuation question in such experiments as: *"what is the maximum amount that I believe, everybody should pay, to ensure this government project?"* Various authors recommend inferring collective WTP through referendum-style experiments, where the preferred implementation criteria would be the majority rule (e.g. Arrow et al., 1993). Since everyone is asked to contribute, the coordination problem associated with private WTP studies is resolved. Moreover, participants are facilitated to express (altruistic and moral) considerations regarding the way government should trade-off burdens and benefits of public policies (Posner and Sunstein, 2017).

2.3 From collective WTP to WTAPB

The shift from private WTP to collective WTP does not solve all criticisms regarding WTP-based valuation. A remaining critique concerns the implicit assumption within private and collective WTP-based valuation that private euros and public euros cannot have a different purpose, i.e. they are completely fungible (Hess et al., 2012; Mouter et al., 2017). This principle allows to establish the social welfare effect of a public good that is financed with government funds (public euros) through aggregating the number of euros that individuals are willing to pay from their after-tax

income (private euros). However, 'complete fungibility' does not fit with what is observed in reality (e.g. Thaler, 1999; Tversky and Kahneman, 1981).

Thaler (1999), for instance, shows that euros contained within a given budget can indeed have a specific goal or purpose. As such, they are at best imperfect substitutes for euros from other budgets, even for the same individual. From this point of view, a more defensible notion is to assume that individuals might view their private income and government funds as constituting two separate budgets. If individual A has revealed (s)he is willing to pay X euros from their own budget for government project Z, we cannot then simply conclude that this individual also believes that X euros should be taken from the government budget. When it makes a difference whether public projects are paid for by private euros (e.g. private or collective WTP) or public euros, WTP-based valuation is no longer useful for the evaluation of government projects financed from public revenues.[1] Instead, the welfare effects of such projects should be inferred from individuals' preferences regarding the expenditure of public euros.

In so-called willingness to allocate public budget experiments (WTAPB) participants make choices over alternative allocations of government budget across different government projects (e.g. Anand and Wailoo, 2000; Dolan and Tsuchiya, 2005; Johannesson and Johansson, 1997; Johansson-Stenman and Martinsen, 2008; Mouter et al., 2017). WTAPB experiments do not directly impact the respondent's disposable income but instead highlight the opportunity costs of public spending. Mouter et al. (2017) empirically observe that participants express different preferences in a *WTAPB setting* and a *collective WTP setting*. More specifically, participants assign substantially more value to safety than travel time in a WTAPB setting compared to the collective WTP setting.

[1] Importantly, government policies which only impose costs on the private sector (and not on public revenue) should still be evaluated using WTP-based valuation.

Mouter et al. (2018) highlight that the participants in the referred study argued that the government has a special duty of care when it comes to road safety. Moreover, none of the participants insisted that the government has any special duty (of care) when it comes to reducing travel times. When individuals feel that the government has a special obligation in terms of the provision of a specific public good it is plausible that they assign a high value to the provision of this public good in WTAPB experiments. They will support a (re)allocation of government funds towards road safety (at the cost of allocating public budget towards the provision of another public good such as travel time reduction). At the same time, individuals' collective WTP in a referendum style experiment for the same special duty of the government can be relatively low or even zero as they might believe the respective project should be financed from (a re-allocation of) existing governmental budget. In the most extreme case, individuals may be unwilling to substitute between the private and public budget. They may then be inclined to express 'protest votes', i.e. a zero WTP, in a private or collective WTP setting since they believe the public good should be financed from (a re-allocation of) previously collected taxes (e.g. Howley et al., 2010; Jorgenssen et al., 1999; Ovaskainen and Kniivila, 2005; Stevens et al., 1991). In a WTAPB experiment, respondents are asked to make choices when faced with effects accruing from alternative allocations of governmental budgets. Hence, the problem that respondents protest because they think that the government should pay from (a re-allocation of) previously collected taxes is non-existent.

To re-iterate: an individual may be unwilling to pay for an environmental preservation project from his after tax income in a private or collective WTP setting, and yet believe that the government should finance the very same project by (re)allocating public budgets away from other public goods, because they see the protection of nature as an important government task. Similarly, this individual may be willing to support a referendum for organizing a major sports event financed

by a one-time tax increase (collective WTP setting), whilst fiercely opposing a re-allocation of taxes from environmental preservation to organise the same event.

In summary, WTAPB-based valuation thus allows participants to express that government funds and their own money may have different purposes. Moreover, the WTAPB approach bypasses the concern that WTP-based valuation is an unsuitable way to value impacts of government projects that are incommensurable with private income (Aldred, 2006; Clark et al., 2000). Sunstein (1993) argues that values that are not traded in a real-life market setting, such as free speech, biodiversity and landscape might be valued in the wrong way when they are expressed in private income. Crucially, WTAPB does not require translation of government project impacts into private income. Instead, an impact of a government project is valued through the extent to which individuals are willing to sacrifice impacts of other government projects. For instance, in a WTAPB experiment, individuals are asked to trade-off environmental impacts against impacts of other governmental policies (e.g. reduction of mortality risk).

2.4 Participatory Value Evaluation as an extension to WTAPB and collective WTP approaches

In this subsection we introduce two versions of Participatory Value Evaluation (PVE) as extensions to the earlier introduced WTAPB and collective WTP approaches. First, Fixed budget PVE provides an extension of the WTAPB approach and overcomes the issue that individuals are forced to allocate the full public budget over a limited set of policies. Second, Flexible budget PVE provides individuals the opportunity to increase (or decrease) the public budget by levying a collective tax increase (reduction) and thereby integrates elements of collective WTP back into the welfare framework.

One clear downside of the WTAPB approach is that respondents are forced to make a choice between two or three alternative allocations of public budgets (e.g. Anand and Wailoo, 2000; Johansson-Stenman and Martinsson, 2008). In case respondents believe that it is better to do nothing instead of allocating public budget to the proposed projects, they do not have the opportunity to express this preference. Hanley et al. (2001) argue that any WTP estimate based on an experimental design in which the baseline is not present will yield inaccurate estimates of consumer welfare. Such biases will particularly arise when for some respondents, the most preferred option is the current baseline (i.e. do nothing) situation. The fact that respondents have the opportunity to express this preference in a 'fixed budget PVE' is an important benefit compared to WTAPB experiments.[2]

One potential critique with regard to the 'fixed budget PVE' format is that preferences of individuals who think that it is better to reduce taxes (and increase private income) instead of spending public budget in the current year and in the next year are not respected. Therefore, we also developed the 'flexible budget PVE' format in which participants are also allowed to adjust the public budget (and their after-tax income) by changing taxes. This set-up gives individuals who believe that it is better to reduce taxes (and increase private income) instead of choosing any of the government projects they can select in the PVE the opportunity to express their preference. Moreover, this format respects the preferences of individuals who think that taxes should be increased to facilitate the implementation of more projects than allowed by the initial public budget. The conceptual innovation of the introduction of flexible budget PVEs is that the collective WTP valuation paradigm and the WTAPB valuation paradigm are integrated in one valuation

[2] Another distinguishing feature of PVE is that respondents can delegate their vote to an expert or to other citizens (e.g. because they believe that they do not have enough knowledge to make the decision themselves). To our knowledge, WTAPB experiments that so far have been conducted do not provide respondents with the opportunity to express this preference, despite the likely presence of preference uncertainty (e.g. Dekker et al. 2016).

framework. That is, the social welfare effects of government projects are established through the elicitation of individuals' preferences over the allocation of public budgets as well as their private income. Although the experimental set-up of the flexible budget PVE re-introduces private income into the mix, the issues of WTP-based valuation which WTAPB-based valuation aims to resolve are not re-introduced. For instance, preferences of individuals who believe that government funds should be spend on different purposes than their own money can be expressed in flexible budget PVEs and not in WTP-based valuation experiments. Moreover, the problem in WTP-based approaches that respondents protest because they think that the government should pay from (a re-allocation of) expected or previously collected taxes is non-existent in a flexible budget PVE.

Table 1 illustrates the relevant differences between PVE and the other approaches addressed in this section in terms of the valuation question for the respondent, the (economic) questions that can be answered and, most importantly, the preferences of individuals that can(not) be expressed by individuals who participate in the valuation experiment. In the next section, we present a real-life case study to show how a PVE study can be conducted and how the choices of participants in a PVE can be translated into policy recommendations.

Table 1: Valuation questions, economic questions addressed and preferences that cannot be expressed by PVE and other approaches

3. Case study: A PVE for the Dutch Ministry of Infrastructure and Water Management

The first application of PVE investigates the societal costs and benefits of a flood protection scheme of the Dutch Ministry of Infrastructure and Water Management. The scheme focused on a

trade-off between two types of projects to mitigate flood risks at locations along the Dutch river 'de Waal' which do not meet the prescribed safety standards. The first type of project is simply strengthening the dikes (henceforth: 'classical project'). The second type of project involves strengthening the dikes to some extent combined with measures to give the river space to flood safely (henceforth: 'combination project'). The two types of projects have an equal impact on mitigating flood risks but differ on several societal impacts (e.g. costs, impact on biodiversity, impact on recreation and number of households that need to relocate). Combination projects increase recreation opportunities and biodiversity but are more expensive. Analysts who conducted the standard CBA analysis were not able to translate biodiversity impacts into monetary terms and for the valuation of the improvement of recreational opportunities a rough rule of thumb of 1 euro per additional recreational trip was used regardless of the quality of this recreational trip (Ecorys, 2017). The low-quality information with regard to the value of the most important goals of the combination projects urged the Ministry to look for other methods and eventually commission a PVE-study.

3.1 Experimental set-up

The main tasks of participants in the PVE concerns recommending the Dutch government on the allocation of a budget of 700 million euros. A demo version of this PVE can be found online: www.participatie-begroting.nl (in Dutch) and http://burgerbegroting.tbm.tudelft.nl/pve-flood-protection (the English translation). On four locations alongside the river 'de Waal' citizens must choose between a 'classical project' and a 'combination project' (Figure 1 depicts these four locations).

Figure 1: The four locations on which respondents have to choose between a 'classical project' and a 'combination project'.

Respondents receive information about the impacts of choosing for a 'classical project' or a 'combination project' at each location (e.g. construction costs, impact on recreation, impact on biodiversity and number of households that experience nuisance due to the project as they potentially need to relocate). In case there is governmental budget left, respondents can spend it on six other projects that fall within the responsibility of the Dutch Ministry of Infrastructure and Water Management (two road projects, two projects mitigating damage from heavy rainfall, and two projects reducing flood risks beyond current safety standards). Respondents also receive information regarding the impacts of these projects (e.g. travel time savings and reduction in the number of severe traffic accidents). In essence, respondents need to decide about the extent to which they want to sacrifice these 'other projects' in order to enable the selection of the more expensive 'combination project' on the locations alongside 'de Waal'.

The four locations alongside the river 'de Waal' were selected in close collaboration with policy makers from the Ministry who were preparing these decisions. We selected locations for which a strategic choice between a 'classical project' and a 'combination project' needed to be made in the short run (after this strategic choice various more detailed decisions needed to be made). The six other projects were selected together with policy makers from other departments within the Ministry who we asked to suggest projects that were also up for a decision in the short run.

To measure the importance of different policy impacts (i.e. the policy attributes) on the portfolio choice of respondents, we vary the project impacts of each of the projects presented to respondents using an experimental design (see the supplementary material for more information).

Across the board, the most important differences between the four locations in terms of the impact of choosing for either the 'classical project' or the 'combination project' were: 1) the difference in costs between the two options was much lower at the location Gendtse Waard compared to the other locations. However, at the other three locations choosing for the combination project had a stronger positive impact on biodiversity; 2) at the location "Oosterhout" a number of households would experience nuisance from the choice for the combination project and at the other three locations households would not experience any nuisance.

Two PVE experiments were conducted: a 'fixed budget PVE' and a 'flexible budget PVE'. In the 'fixed budget PVE' respondents were asked to recommend a portfolio of projects given a governmental budget constraint of 700 million euros. Any remaining budget was shifted forward to the next year. The only difference between the fixed and the flexible PVE is that in the latter respondents could adjust the governmental budget by increasing the tax per household or by selecting a rebate. A tax increase (decrease) of 8 million euro of the budget in the PVE resulted in an increase (decrease) of taxes of 1 euro per household in 2019. Hence, the key difference between the two formats is that individuals participating in the flexible budget PVE can influence their after-tax income, whereas participants of the fixed budget PVE do not have this opportunity.

Both the fixed and the flexible PVE were conducted in a web-based environment (see the demo version). In this environment, respondents could sort and compare the projects by one of the impacts and find out more about the (impacts of) projects through clicking on an information button. Figure 2 presents a screenshot of the page where respondents could select projects and Figure 3 presents a screenshot of a page where respondents could find more information about a project.

Figure 2: Screenshot of a part of the page of the flexible budget PVE on which respondents could select projects.

Figure 3: Screenshot of a page on which respondents could find more information about a project

Respondents could delegate their decision to an expert. The delegates also filled out the experiment. When citizens delegated, their choice was replaced by the choice of the selected delegate. After respondents selected their preferred portfolio, they were asked to motivate their choices for each project they selected. Finally, respondents were asked to evaluate several aspects of the PVE. The survey company Kantar Public was asked to draw two random samples of Dutch citizens of 18 years and older and we used a between-subjects design for our study. 2,900 respondents participated in the experiments out of which 937 respondents were specifically recruited in the areas adjacent to the river ‘de Waal’. All respondents who finished the experiment received a financial compensation from the survey company. In case respondents delegated their choice, they received a lower financial compensation (around 1 euro). Respondents were told that the difference in financial compensation would be used to pay the experts to make a choice. The use of different incentives for providing an advice or delegating might muddle the preference elicitation process. At the margin, if getting extra euros from the survey company is more important for the respondent than making the "right" choice, then he might go ahead and click through just to get the extra euro, even if, absent the monetary incentive, he thinks it is better if an expert makes the call. Notwithstanding this drawback, we decided to provide respondents who delegated with a lower financial compensation because we were afraid that many respondents would choose for ‘delegation’ when we would give them the same financial compensation. The

underlying reason for this decision is that this was the first time that we applied PVE and we wanted to learn as much as possible from the responses of the participants.

3.2 Results

Before we present the results, we wish to emphasize that the main goal of the case study was to learn how PVE would work out when applied in practice and to provide a tangible example of the method. We found much room for improvement while conducting the study. We recommended the policy makers who commissioned the study to keep this in mind when interpreting the results of the study.

In the fixed budget PVE 266 out of 1,855 respondents delegated their choice to an expert (14%) and in the flexible budget PVE 223 out of 1,045 respondents delegated (21%). The fact that more individuals delegated in the flexible budget PVE may be the result of the increased complexity of the choice task in this experimental setting. In the fixed budget PVE 2.1% of the respondents fully exhausted the budget, which is partly a result of the limited number of project portfolios available that would exhaust the public budget. In the flexible budget PVE the percentage of the respondents who fully exhausted the budget was a bit higher (6.8%). Figure 4 provides more detailed information with regard to the size of the budget that was shifted forward in the two experiments.

Figure 4: Budget shifted forward to the next year in the Fixed Budget PVE and Flexible Budget PVE

620 respondents participating in the flexible budget PVE did not change the budget, 122 respondents decreased the budget and 82 respondents increased the budget (see Figure 5 for more detailed information on the extent to which respondents changed the budget). We did not find a significant correlation between the income of the respondent and their decision to change the budget. The only respondent who selected the null portfolio (choosing for the classical project at each location and a maximal reduction in budget) was an individual with a very low income.

Figure 5: Changes in budgets in the Flexible PVE

After participants of the flexible budget PVE completed the experiment we asked them why they decided to (not) change the budget. Most respondents who increased the budget stressed the importance of biodiversity and said that the relatively low tax increase that was needed to improve biodiversity urged them to increase the budget. Many respondents who decreased the budget referred to low importance of the projects among which they could choose. Moreover, a group of participants argued that they reduced the budget because they had a negative stance towards government spending in general. For instance, one respondent argued: *"I think that taxes should be as low as possible. Only the most essential tasks should be conducted and financed by the government."* Many of the respondents not changing the budget argued that the government would have a good reason for setting the budget at this level and they found it risky to overrule such a decision. Another group of respondents stated that they didn't decrease the budget because they thought that it would be good if the government had some financial reserves in case of a setback. For instance, one respondent argued: *"I didn't decrease the budget because such projects always*

face cost overruns." Various respondents argued that they didn't increase the budget because they thought that the government should respect its budget.

Figure 6 presents the market shares of the different projects for the respondents who did not delegate their decision to an expert. To check for spatial differentiation in project choices, both the market shares for the sample (The Netherlands), and for the respondents recruited in the Waal area are reported. For each project the average costs presented to the sample are also displayed.

Figure 6: Percentage of respondents which selects the classical projects, the combination projects and the six other projects.

In all four locations the majority of respondents selected the 'combination project'. In the fixed budget PVE and the flexible budget PVE respectively 39.1% and 40.9% of the respondents selected the combination project at all four locations. The classical project was selected at all four locations by 7.0% of the respondents (fixed budget PVE) and 7.6% of the respondents (flexible budget PVE). Strikingly, the results did not differ very much between the respondents living close to the river 'Waal' and random sample of Dutch citizens. A proportions test revealed that, at the 5% significance level, only for the road project Joure A6/A7 motorway and the project mitigating heavy rainfall at Hooge Boezem were selected more often by respondents from the random sample of the Dutch population than respondents living in the Waal area. Overall, the results of the fixed and flexible budget PVE are quite similar. A proportions test across the fixed and flexible budget PVE reveals that only for the Road Expansion A2 motorway we observe a significant difference in proportions between the fixed and flexible PVE samples at the 5% significance level. Although the comparisons between the two samples is not entirely fair (e.g. in the flexible budget PVE

respondents were allowed to change the budget and respondents in the fixed budget PVE did not had this option) we take this as evidence that the two samples are largely comparable.

To establish how citizens value the portfolios, the choices are quantitatively analysed using advanced discrete-continuous choice models.[3] We estimate taste parameters in order to derive the relevance of societal impacts associated with the projects. Such information (comparable to attributes in stated choice surveys) was explicitly presented to participants during the PVE experiment. This analysis revealed that participants particularly preferred the combination projects over the classical projects when the former projects would positively influence biodiversity and recreational opportunities (see the supplementary material for more detailed information). This is a surprising result as these impacts generally only play a marginal role in the Dutch CBA practice (Mouter et al., 2015). Respondents' answers to the question why they selected the combination projects also show that improved recreational opportunities and variety in biodiversity are the main reasons for choosing these projects. However, we also inferred from these qualitative motivations that participants favour combination projects because they believe that this solution to mitigate flood risk is aesthetically superior and is more 'future proof'. Hence, we estimate so-called project specific parameters for each project which captures the utility individuals derive from a project irrespective of the level of the impacts included explicitly in the PVE (comparable to alternative specific constants in stated choice surveys).

The obtained results can be used for welfare analysis (see the supplementary material and Dekker et al., 2019 for a detailed discussion of conducting such a welfare analysis). A first output of the welfare analysis is the derivation of the probability that an individual project improves societal welfare compared to shifting the required budget to the next period. As such, this reveals

[3] More details on modelling and mathematical notation can be found in the supplementary material and Dekker et al. (2019).

the probability that the project provides value for money. This is a key step in the policy evaluation since participants in the PVE always have the fallback option to choose the null portfolio. To illustrate, in case all the participants in the PVE would have selected the null portfolio thereby recommending to select the classical project at each of the four locations and shift the remaining public budget to the next year, the probability that one of the combination projects improves societal value compared to shifting the money to the next period would be approaching 0%.

Figure 7 shows that all the combination projects provide value for money. For instance, choosing the combination project at Gendtse Waard has a 86% probability to improve societal value compared to choosing for the classical project at this location and shifting the difference in costs (in this case 5 million euro) to the next year. The project desirability of the road expansion of the A2 motorway is 31% which means that this project has a 31% probability to improve societal value compared to shifting public budget to the next year. This project should not be implemented, irrespective of the available budget.

Figure 7: Probability that a project improves societal welfare

A second output of the welfare analysis is the ranking of portfolios of projects in terms of expected social welfare. When the public budget is unlimited policy makers should opt for all projects with a desirability probability of higher than 50%: the combination projects at all four locations and the projects Moerwijk, Venlo, Hooge Boezem, Driemanspolder and Joure. However, in reality policy makers are faced with limited budgets and PVE allows for the identification of the optimal selection of projects (i.e. the optimal portfolio) for a given budget. Figure 8 shows the top 10 of portfolios within a budget constraint of 688 million euros. We used the average budget

recommended in the flexible budget PVE (688 million euros) as the budget constraint. This implies a tax decrease of 1.5 euro per household in 2019.

Figure 8: 10 portfolio's which result in the highest expected social utility

Based on these results we can draw three main conclusions: 1) the large road project (road expansion A2 motorway) is not included in all the top 10 portfolio's; 2) at the locations Sleeuwijk and Werkendam the combination project is included in all the top 10 portfolio's; 3) the optimal portfolio opts for the combination project in all four locations. In order to see whether these conclusions are robust to changes in assumptions concerning the level of the social impacts included in the experiment (e.g. costs and impact on biodiversity) we performed various sensitivity analyses (see Dekker et al., 2019). These sensitivity analyses reveal that the first two conclusions are highly robust to changes in assumptions. However, when we assume a very low impact of the combination projects on biodiversity and recreational opportunities, the combination project is not included in the optimal portfolio at the location 'Oosterhout'.

3.3 Experiences of participants and policy makers

Policy makers who commissioned the PVE case study gave various presentations about the results of the study (both within and outside the Ministry). In the presentations they stated that a useful insight for them was that both the quantitative and qualitative information from the PVE indicate that citizens particularly prefer the combination projects over the classical projects when the former projects positively influence biodiversity and recreational opportunities. This insight complemented the lack of information with regard to the value of biodiversity impacts in the

standard CBA analysis. Moreover, policy makers found it interesting to learn which type of projects citizens are willing to sacrifice to enable them to choose for the more expensive combination project instead of the cheaper classical project. For instance, citizens are willing to sacrifice the large road project (road expansion A2 motorway) to make way for more expensive combination projects which foster biodiversity and recreation. Finally, policy makers acquired new insights from the qualitative motivations of respondents for their choice for either a classical project(s) or a combination project(s). For instance, they learned that citizens also favour combination projects because they believe that this solution to mitigate flood risk is aesthetically superior and is more 'future proof'.

Policy makers also argued that a strength of PVE concerns the facilitation of mass participation of citizens in the evaluation of public policies. A key benefit of PVE compared to existing approaches for citizen participation (e.g. public hearings and citizen juries) is that the entry barriers for participating are relatively low. Participants generally spend 20 to 30 minutes to submit their choice(s), and the respondents can choose themselves when and where they conduct the PVE. As a result of the low entry barriers not only the passionate proponents and opponents, but also the so-called 'silent majority' can participate in the evaluation of public policies. The socio-demographics of the respondents (see the supplementary materials) reveal that all relevant segments of the population are represented.

In the final part of the PVE experiment, participants were asked about the extent to which they agreed with four propositions. All four propositions were evaluated in a positive way: 1) *"I was convinced of my choices"* (89% agree; 18% neutral; 3% disagree); *"I thought that the experiment was realistic"* (61% agree; 28% neutral; 11% disagree); 3) *"I think it is good that the government aims to involve citizens in making choices between projects in the experiment"* (82%

agree; 12% neutral; 6% disagree); 4) *"This experiment provides the government with relevant information for making choices between projects"* (65% agree; 26% neutral; 9% disagree). The fact that more than 80% of the participants in the case study agreed that it is good that the Ministry of Transport and Water Management involved citizens in the evaluation of policy options indicates that there is a clear demand for public participation among Dutch citizens. Possibly, PVE can be a response to this demand.

Besides responding to the four statements mentioned above, participants in the PVE were asked to provide qualitative feedback with regard to the PVE-experiment in which they participated. Various citizens argued that they thought that participating in a PVE is a nice way to get involved into policymaking and to ensure that their voices are heard. They liked the fact that they are invited to make a strategic choice that is close to the dilemma policy makers face. Moreover, citizens who participated in the PVEs argued that it raised their awareness concerning the dilemmas public bodies are faced with in making complex decisions, because they have to make – consequential – choices themselves. For instance, citizens learned about scarcity of public resources (not everything is possible) and the cons and pros of the alternative policy options.

4. Discussion

The contribution of PVE particularly lies in the flexible budget PVE format in which individuals have the opportunity to simultaneously optimise (adjust) the size of the public budget – through a tax increase/decrease affecting their private income – and allocate it to their preferred portfolio of public projects. As such, flexible budget PVE experiments integrate the WTP valuation paradigm and the WTAPB valuation paradigm in a single framework. Policy recommendations from PVE

experiments are based on the notion of social welfare and allow easy ranking of optimal policy portfolios.

We argue that PVE is a promising valuation method which warrants to be further developed in future research projects. PVE has the potential to solve issues associated with other valuation methods (see section 2). Moreover, the first PVE application was positively evaluated by policy makers who commissioned the case study and by citizens who participated in this case study. Nevertheless, we wish to emphasize that the purpose of this study is not to claim that PVE is superior to other valuation approaches. Arguably, the appropriateness of using PVE or other valuation methods depends on the economic question that is being posed (see also Table 1). For instance, it is conceivable that PVE is a much more useful tool for analysing preferences over a fixed set of public projects using existing public funds than in the case of the evaluation of a public project that might be financed through voluntary private contributions of citizens.

More research needs to be done to provide any final answers to the normative question about whether PVE is (not) a more appropriate method for the valuation of public goods compared to other approaches. A practical approach to this question is to evaluate the welfare effects of (a set of) government projects through both a PVE and traditional WTP-based valuation techniques and evaluate the performance of these two studies on various criteria such as hypothetical bias, protest votes, confidence of respondents in their choices and the extent to which respondents believe that the study provides the government with relevant information for making choices between projects. Moreover, it might be interesting to compare the effort that participants in different valuation methods make to complete the choice task. Although the choice tasks in a PVE are relatively complex, we saw that 89% of respondents agreed with the statement: *"I was convinced of my choices"* which indicates that most of the respondents understood the experiment.

Our hypothesis is that respondents make a much harder effort to digest the information of a PVE because they believe that their responses are more consequential (both in terms of payment and policy consequentiality) than their responses in a stylized WTP experiment. However, we did not collect any empirical evidence to test this hypothesis so this would be an interesting avenue for further research.

Apart from normative research questions, it would be interesting to apply PVE in various other case studies to investigate the extent to which the results of this study can be generalized. For instance, it would be interesting to investigate whether the similarity of preferences between individuals who participated in the fixed and flexible budget PVEs is also detected when PVE is applied in other contexts. These case studies would also allow for further development of the method, for instance, in terms of the information provision to participants, experimenting with new delegation options (e.g. providing participants who delegate with the same financial compensation and also provide participants with the option to delegate to politicians) and tracking the behavior of participants during the experiment (how many respondents watch the whole instruction video and how many respondents used the sorting function?) to analyze whether respondents who adopt different choice strategies make different choices.

Acknowledgements: the first and second author wish to thank the Netherlands Organisation for Scientific Research for financial support (NWO Responsible Innovation grant – 313-99-333). Moreover, the authors wish to thank the Dutch Ministry of Infrastructure and Water Management for financing the data collection and the design of the PVE.

References

Abrantes, P.A.L., Wardman, M., 2011. Meta-analysis of the UK values of time: an update. Transportation Research Part A: Policy and Practice 45, 1-17.

Ackerman, F., Heinzerling, L., 2004. Priceless: on knowing the price of everything and the value of noting. The New Press. New York.

Aldred, J., 2006. Incommensurability and monetary valuation. Land Economics 82 (2), 141-161.

Anand, P., Wailoo, P., 2000. Utilities versus rights to publicly provided goods. Arguments and evidence from health care rationing. Economica 67, 543-577.

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., 1993. Report of the NOAA panel on contingent valuation. Federal Register 58, 4601-14.

Boadway, R., Bruce, M., 1984. Welfare Economics. Oxford: Basil Blackwell.

Boadway, R., 2006. Principles of cost-benefit analysis. Public Policy Rev. 2(1), 1–42.

Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L., 2013. Cost-benefit analysis concepts and practice. New Jersey: Prentice Hall.

Clark, J., Burgess, J., Harrison, C.M., 2000. I struggled with this money business: respondents' perspectives on contingent valuation. Ecological Economics 33, 45-62.

Dekker, T., Hess, S., Brouwer, R., Hofkes, M. 2016. Decision uncertainty in multi-attribute stated preference studies. Resource and Energy Economics 43, 57-73.

Dekker, T., Koster, P.R., Mouter, N., 2019. The economics of Participatory Value Evaluation. Tinbergen Institute Discussion papers 19-008/VIII.

Dolan, P., Tsuchiya, A., 2005. Health priorities and public preferences: the relative importance of past health experience and future health prospects. Journal of Health Economics 24, 703–714.

Ecorys, 2017. MKBA Rivierverruiming. Rapportage ronde 2. Opdrachtgever, Ministerie van Infrastructuur en Milieu.

Hahn, R.W., Tetlock, P.C., 2008. Has economic analysis improved regulatory decisions? J. Econ. Perspect. 22 (1), 67–84.

Hanley, N., Mourato, S., Wright, R.E. Choice modelling approaches: a superior alternative for environmental valuation? Journal of Economic Surveys 15: 435-62.

Hess, S., Orr, S., Sheldon, R., 2012. Consistency and fungibility of monetary valuations in transport. Transportation Research Part A, 46 (10), 1507–1516.

Howley, P., Hynes, S., O'Donoghue, C., 2010. The citizen versus consumer distinction: An exploration of individuals' preferences in Contingent Valuation studies. Ecological Economics 69, pp. 1524-1531.

Ivehammer, P., 2009. The payment vehicle used in CV studies of environmental goods does matter. Journal of Agricultural and Resource Economics 34 (3), 450-463.

Johannesson M., Johansson P.O., 1997. Is the valuation of a QALY gained independent of age? Some Empirical Evidence. Journal of Health Economics 16(5), 589-99.

Johansson-Stenman, O., Martinsson, P., 2008. Are some lives more valuable? An ethical preferences approach. Journal of Health Economics 27 (3), 739-752.

Jorgensen, B., Syme, G., Bishop, B., Nancarrow, B., 1999. Protest responses in contingent valuation. Environ. Resour. Econ. 14, 131–150.

Koster, H.R.A., van Ommeren, J., 2015. A shaky business: natural gas extraction, earthquakes and house prices. European Economic Review 80, 120-139.

Lusk, J.L., Norwood, F.B., 2011. Animal Welfare Economics. Applied Economic Perspectives and Policy 33 (4), 463-483.

Mackie, P.J., Worsley, T., Eliasson, J., 2014. Transport Appraisal Revisited. Research in Transportation Economics. Vol. 47, 3-18.

Mouter, N., Annema, J.A., Van Wee, B., 2015. Managing the insolvable limitations of cost-benefit analysis: results of an interview based study. Transportation 42(2), 277–302.

Mouter, N., van Cranenburgh, S., van Wee, G.P., 2017. Do individuals have different preferences as consumer and citizen? The trade-off between travel time and safety. Transportation Research Part A 106, pp. 333-349.

Mouter, N., van Cranenburgh, S., van Wee, G.P., 2018. The consumer-citizen duality: Ten reasons why citizens prefer safety and drivers desire speed. Accident Analysis & Prevention 121, 53 – 63.

Nyborg, K., 2000. Homo Economicus and Homo Politicus: interpretation and aggregation of environmental values. Journal of Economic Behavior & Organization 42, 305–322.

Ovaskainen, V., Kniivilä, M., 2005. Consumer versus citizen preferences in contingent valuation: evidence on the role of question framing. Australian Journal of Agricultural and Resource Economics 49, 379–394.

Persky, J., 2001. Retrospectives. Cost-Benefit Analysis and the classical creed. Journal of Economic Perspectives 15 (4), 199-208.

Posner, E.A., Sunstein, C.R., 2017. Moral commitments in Cost-Benefit Analysis. Coase-Sandor Institute for Law and Economic Working Papers. No. 802.

Sagoff, M., 1988. The Economy of the Earth: Philosophy, Law, and the Environment. Cambridge University Press.

Sen, A., 1995. Environmental Evaluation and Social Choice: Contingent Valuation and the Market Analogy. The Japanese Economic Review 46 (1), 23-37.

Stevens, T., Echeverria, J., Glass, R., Hager, T., More, T., 1991. Measuring the existence value of wildlife: what do CVM estimates really show? Land Economics 67, 390–400.

Sunstein, C.R., 1993. Incommensurability and Valuation in Law. Michigan Law Review 92: 779.

Thaler, R., 1999. Mental accounting matters. Journal of Behavioral Decision Making 12, 183-206.

Tversky, A. Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211, 453-458.

Table 1: Valuation questions, economic questions addressed and preferences excluded by PVE and other approaches

Approach	Valuation question	Economic question	Which preferences cannot be expressed?
Private willingness to pay	What is the amount that somebody is willing to pay from his/her private income on top of the taxes that are already paid to finance a public good?	Should we finance a public good through (voluntary) private contributions of individuals?	1) Preferences of individuals who believe that their private consumption choices do not reflect the way that they think that government should trade-off impacts of public policies 2) Preferences of individuals who believe that government funds should be spend on different purposes than their own money. 3) Preferences of individuals that want to delegate their judgment with regard to the value of a government project to an expert.
Collective willingness to pay	What is the amount that all individuals under the jurisdiction of the relevant authority should pay from their private income, on top of the taxes they already pay, to finance a public good?	Should we finance a public good through a collective tax increase?	1) Preferences of individuals who believe that government funds should be spend on different purposes than their own money. 2) Preferences of individuals that want to delegate their judgment with regard to the value of a government project to an expert.
Willingness to allocate public budget	To which extent do individuals support the allocation of public budget towards a public good at the cost of allocating public budget towards the provision of other public goods that are part of the experiment?	Given that the government decided to allocate a certain amount of public budget; which public goods should be financed?	1) Preferences of individuals who believe that it is better to do spend the public budget on other projects than the proposed projects in the experiment. 2) Preferences of individuals who believe that it is better to reduce taxes instead of allocating public budget to one of the proposed projects. 3) Preferences of individuals who believe that it is better to increase taxes to select more projects than allowed by the constrained public budget. 4) Preferences of individuals that want to delegate their judgment with regard to the value of a government project to an expert.
Fixed budget PVE	To which extent do individuals support the allocation of public budget towards a public good at the cost of allocating public budget towards the provision of other public goods that are part of the experiment or to public goods that can be financed in the next year (which are not part of the experiment)?	Given that the government decided to allocate a certain amount of public budget; which public goods should be financed in this year or should we shift budget to the next year?	1) Preferences of individuals who believe that it is better to reduce taxes instead of allocating public budget to one of the proposed projects. 2) Preferences of individuals who believe that it is better to increase taxes to select more projects than allowed by the constrained public budget.
Flexible budget PVE	To which extent do individuals support the allocation of public budget towards a public good at the cost of allocating public budget towards the provision of other public goods that are part of the experiment, at the costs of public goods that can be financed in the next year (which are not part of the experiment) or at the costs of their private income?	Should the government allocate a certain amount of public budget to a set of proposed public goods? If so, to which public goods? Or, should the government shift the budget to the next year? Or, should the government amend the public budget through a tax increase/decrease that changes private income?	

Figure 1: The four locations on which respondents have to choose between a 'classical project' and a 'combination project'.

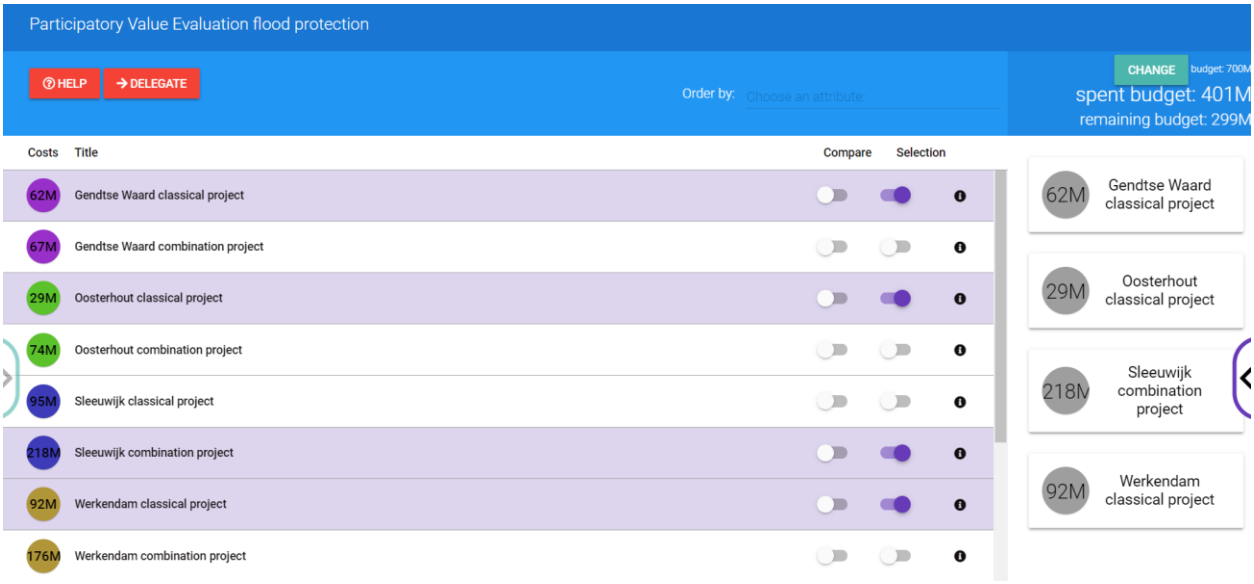

Figure 2: Screenshot of a part of the page of the flexible budget PVE on which respondents could select projects.

Oosterhout combination project

Total costs for this project: 74M

At present, the prescribed safety standards for preventing flood risk are not met at this location. This project ensures that the safety standards are met through: (1) Broadening and elevating the dikes; (2) Giving the river more space which ensures that the river can discharge more water. This results in a reduction of flood risks in the case of high water. The dike will be relocated to give the river more space. This will have some consequences for a campsite and a couple of dwellings that are located close to the current dike. Possibly these households and the campsite should be relocated. The combination project will result in an increase of the nature reserve alongside the river Waal. This will boost nature and recreational opportunities for residents of Oosterhout and Nijmegen.

Protection against flooding:	In line with standards
Prevent water damage:	No impact
Number of households nuisance:	3
Size nature/recreational area:	85
Improvement biodiversity:	Substantial increase
Change recreational area:	Improvement
Number of travelers:	0
Minutes travel time savings:	0
Decrease in severe traffic injuries:	0

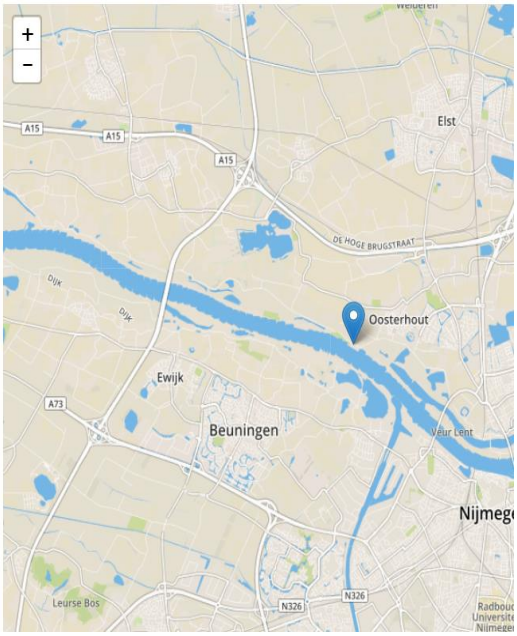

Figure 3: Screenshot of a page on which respondents could find more information about a project

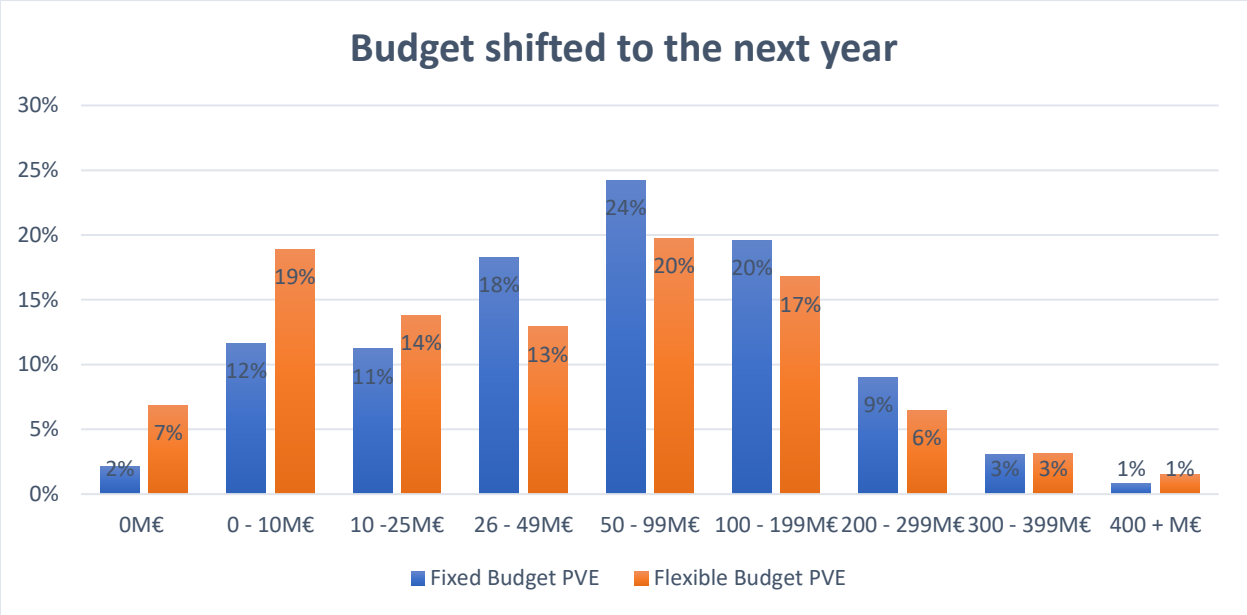

Figure 4: Budget shifted forward to the next year in the Fixed Budget PVE and Flexible Budget PVE

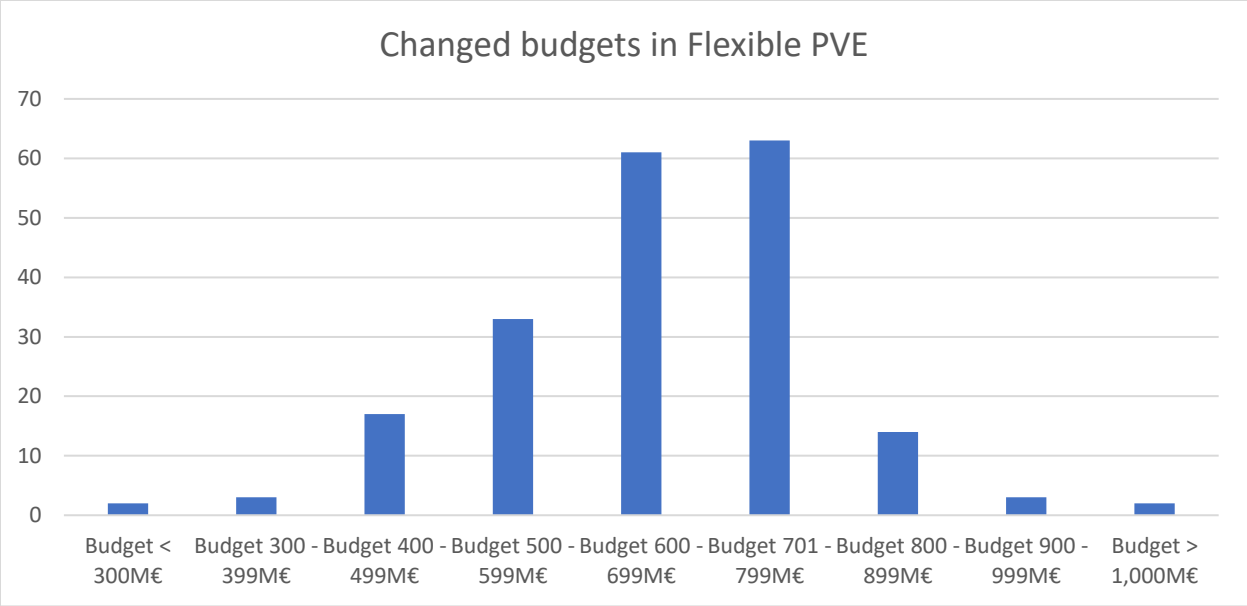

Figure 5: Changes in budget in the Flexible PVE

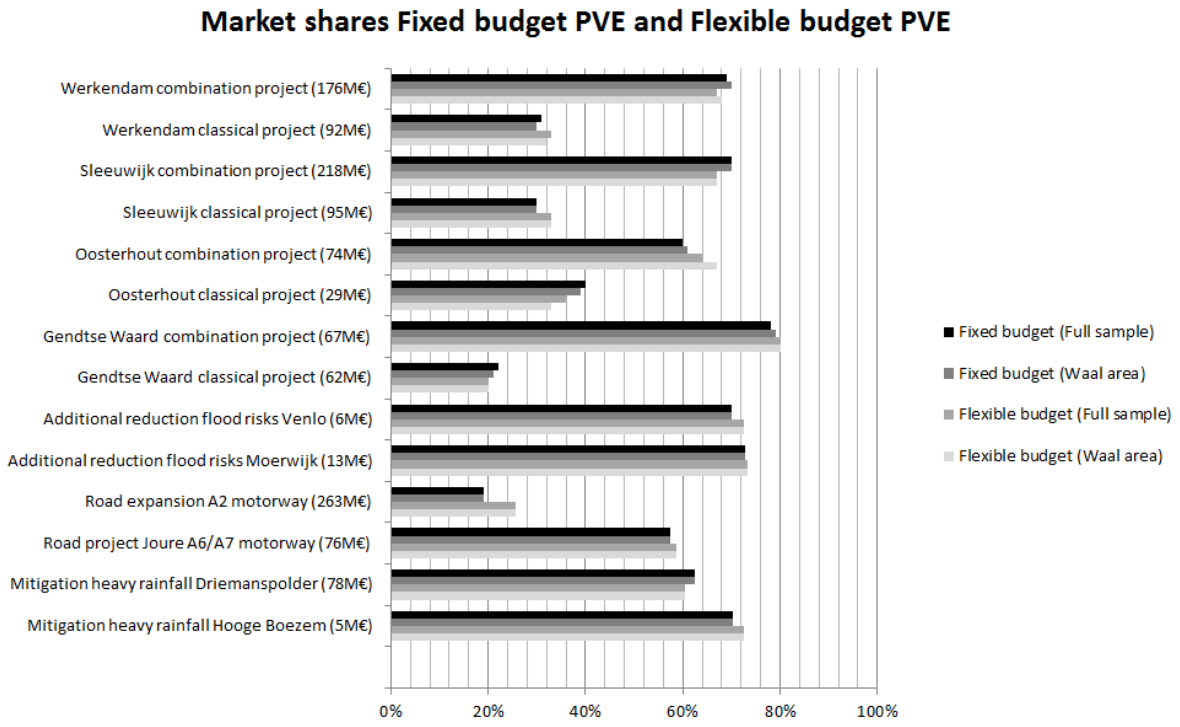

Figure 6: Percentage of respondents which selects the classical projects, the combination projects and the six other projects.

Combination project versus classical project	Project desirability	
Gendtse Waard combination project (67M€)	86%	✓
Gendtse Waard classical project (62M€)	--	X
Sleeuwijk combination project (218M€)	75%	✓
Sleeuwijk classical project (95M€)	--	X
Werkendam combination project (176M€)	75%	✓
Werkendam classical project (92M€)	--	X
Oosterhout combination project (74M€)	67%	✓
Oosterhout classical project (29M€)	--	X
Other projects		
Additional reduction flood risks Moerwijk (13M€)	75%	✓
Additional reduction flood risks Venlo (6M€)	74%	✓
Mitigation heavy rainfall Hooge Boezem (5M€)	74%	✓
Mitigation heavy rainfall Driemanspolder (78M€)	70%	✓
Road project Joure A6/A7 (76M€)	61%	✓
Road expansion A2 motorway (263M€)	36%	X

Figure 7: Probability that a project improves societal welfare

	Top 10 portfolio's									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Gendtse Waard classical project (62M€)	0	0	0	0	0	0	0	1	0	0
Gendtse Waard combination project (67M€)	1	1	1	1	1	1	1	0	1	1
Oosterhout classical project (29M€)	0	1	0	0	1	1	0	0	1	0
Oosterhout combination project (74M€)	1	0	1	1	0	0	1	1	0	1
Sleeuwijk classical project (95M€)	0	0	0	0	0	0	0	0	0	0
Sleeuwijk combination project (218M€)	1	1	1	1	1	1	1	1	1	1
Werkendam classical project (92M€)	0	0	0	0	0	0	0	0	0	0
Werkendam combination project (176M€)	1	1	1	1	1	1	1	1	1	1
Mitigation heavy rainfall Hooge Boezem (5M€)	1	1	1	0	1	0	0	1	0	1
Mitigation heavy rainfall Driemanspolder (78M€)	1	1	1	1	1	1	1	1	1	1
Road project Joure A6/A7 (76M€)	0	1	0	0	1	1	0	0	1	0
Road expansion A2 motorway (263M€)	0	0	0	0	0	0	0	0	0	0
Additional reduction flood risks Moerwijk (13M€)	1	1	1	1	1	1	1	1	1	0
Additional reduction flood risks Venlo (6M€)	1	1	0	1	0	1	0	1	0	0
Costs (in millions of euros)	**641**	**671**	**635**	**635**	**665**	**665**	**629**	**636**	**659**	**627**

Figure 8: 10 portfolio's which result in the highest expected social utility