TI 2019-046/VIII
Tinbergen Institute Discussion Paper

Participatory Value Evaluation versus Cost-Benefit Analysis: comparing recommendations in the context of urban mobility investments

Revision: January 2020

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] Department of Spatial Economics, School of Business and Economics and John Stuart Mill College, Vrije Universiteit Amsterdam and Tinbergen Institute Amsterdam
[3] Institute for Transport Studies Leeds and Choice Modelling Centre, University of Leeds

Participatory Value Evaluation versus Cost-Benefit Analysis: comparing recommendations in the context of urban mobility investments

Authors: Niek Mouter[1]*, Paul Koster[23], Thijs Dekker[4].

1 Delft University of Technology, Faculty of Technology, Policy and Management, Transport and Logistics Group. Email: n.mouter@tudelft.nl.
2 Vrije Universiteit Amsterdam, School of Business and Economics, Department of Spatial Economics. Email: p.r.koster@vu.nl.
3 Vrije Universiteit Amsterdam, John Stuart Mill College. Email: p.r.koster@vu.nl.
4 Institute for Transport Studies Leeds and Choice Modelling Centre, University of Leeds. Email: t.dekker@leeds.ac.uk.

* Corresponding author: Niek Mouter: n.mouter@tudelft.nl

Abstract

Cost-Benefit Analysis (CBA) is a widely applied economic appraisal tool to support the planning and decision-making process for transport projects. In a CBA, impacts of government projects are made comparable by converting them into monetary units using the number of euros individuals are willing to pay from their private income. Scholars argue that such willingness-to-pay estimates may be a poor proxy for how the same individuals believe that their governments should trade-off public budget and impacts of government projects. Participatory Value Evaluation (PVE) is a new appraisal method specifically designed to overcome this critique. PVE establishes the desirability of government projects based on an experiment in which individuals select their preferred portfolio of government projects given a constrained public budget. The present paper investigates whether CBA and PVE lead to different policy recommendations. We conducted CBAs and a PVE for 16 transport projects and find that projects which focus on improving traffic safety and improvements for cyclists/pedestrians perform relatively good in the PVE, whereas car projects perform relatively good in the CBA analysis. Moreover, this paper explains how the results of a PVE should be positioned next to the results of a CBA and it generates empirical insights into potential reasons why safety projects and cycling project perform differently in a PVE.

1. Introduction

Cost-Benefit Analysis (CBA) is a widely applied economic appraisal tool to support the planning and decision-making process for transport projects (e.g. Asplund and Eliasson, 2016; Thomopoulos et al., 2009). In many Western countries it is obligatory to assess a transport project using a CBA when a project needs (co)funding from the National Government (Mackie et al., 2014). Attitudes of policy makers and politicians towards the use of CBA in the planning and evaluation of transport projects have been analyzed in various studies (e.g. Beukers et al., 2012; Mouter et al., 2013a; Mouter, 2017; Nyborg, 1998; Vigren and Ljungberg, 2018). Most of these studies find that these actors believe that CBA is a useful appraisal tool.

In a CBA, positive and negative social impacts of government projects are quantified and where possible converted into monetary units using willingness to pay (WTP) estimates. The fact that CBA estimates the desirability of government policies by estimating the amount of euros that individuals are willing to pay from their private income for the impacts accruing from the policy is fiercely criticized in the literature (e.g. Ackerman and Heinzerling, 2004; Sagoff, 1988). Several scholars argue that such private WTP estimates may be a poor proxy for how the same individuals believe that their governments should trade-off public budget and impacts of government projects (e.g. Ackerman and Heinzerling, 2004; Hauer, 1994; Sagoff, 1988; Sunstein, 2005). Mouter et al. (2017, 2018) empirically establish that individuals indeed make different trade-offs as they assign substantially more value to safety than travel time in the context of alternative allocations of public budget in so-called willingness to allocate public budget (WTAPB) experiments compared to a classical private WTP setting. The private WTP-based valuation paradigm adopted in CBA may thus result in a too narrow perspective when evaluating transport projects that are financed with public budget.

Estimating the desirability of transport projects solely through WTAPB experiments, however, has its own downsides. The most important limitation is that respondents are forced to spend the public budget. Consequently, 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 (Mouter et al., 2019). Participatory Value Evaluation (PVE), in its narrowest form, is a valuation method which has been developed to address the shortcomings of the WTAPB valuation method. The similarity between WTAPB experiments and PVE experiments is that participants are asked to express to the government which projects should be financed from a limited amount of public budget (Mouter et al., 2019). The most important difference is that participants in a PVE also have the option to advise the government against allocating the budget to any of the projects that are considered in the PVE and shift the budget to the next year or to reduce or increase the budget by changing taxes. Another difference is that participants in a PVE allocate public budget to a portfolio of projects and, as a result, they can consider positive and negative synergies between projects and potential spatial equality concerns. This is not possible in a WTAPB experiment were respondents select only one project from a limited set of projects.

Mouter et al. (2019) and Dekker et al. (2019) presented the application of a PVE in the context of a flood protection scheme for the Dutch Ministry of Infrastructure and Water Management. In this PVE, 2,900 citizens were asked to allocate a budget of 700 million euros to flood protection projects in the Netherlands and other projects that fall within the remit of

the Ministry. From this application of PVE no conclusions could be drawn with regard to the extent to which a PVE provides different policy recommendations than a traditional CBA. No CBAs were available for many of the projects considered in the PVE experiment. For the cases CBAs were available, the studies were incomplete as some of the main impacts of the flood protection projects such as biodiversity and recreational opportunities were not monetized or monetized in a very rudimentary way. This was also the main reason why the Dutch Ministry of Infrastructure and Waterworks commissioned a PVE (Mouter et al., 2019).

The primary goal of the present study is to investigate whether CBA and PVE lead to different policy recommendations. More specifically, we conducted CBAs and a PVE for 16 transport projects considered by the Transport Authority Amsterdam (henceforth: TAA) and compared the results of these studies. A transport case study lends itself well for comparative purposes due to the wealth of guidance and experience of CBA-based project appraisal in this domain. In the PVE, 2,498 citizens were presented with the 16 transport projects and related societal impacts. The total costs of the 16 projects was 386.5 million euros but with only 100 million euros to spend, it was not possible for the respondents to include all projects in their portfolio. A demo version of the PVE can be found online: www.burger-begroting.nl (in Dutch) and http://burgerbegroting.tbm.tudelft.nl/participatory-value-evaluation-transport-authority-amsterdam (the English translation). Applying PVE in a transport context is interesting as both in the Netherlands and in other western countries this is the domain in which CBA has the strongest tradition in terms of guidelines and knowledge available to transfer impacts of government projects into monetary terms (Mackie et al., 2014; Romijn and Renes, 2013).

When the two evaluation approaches produce different policy recommendations it is important for policy makers to gain insight into the reasons why the approaches provide different results and to know how PVE should be positioned next to CBA. These two questions have not been addressed in previous PVE papers (Dekker et al., 2019; Mouter et al., 2019). Hence, a second goal of the present paper is to explain how PVE should be positioned next to CBA. Finally, a third objective of the paper is to generate empirical insights into potential reasons why PVE and CBA might provide different results. To achieve this third goal, we asked participants in the PVE to provide written motivations for each selected project after they submitted their preferred portfolio. We analyzed the extent to which these written motivations might provide potential explanations for differences in outcomes between PVE and CBA.

The remainder of this paper is organized as follows: Section 2 explains the main differences and similarities between the PVE method and CBA (objective 2). Section 3 lists potential reasons for why CBA and PVE might produce different results when evaluating urban mobility projects based on a literature review (objective 3). Section 4 describes the 16 transport projects of the TAA that were analyzed using both the CBA and PVE methodology and Section 5 compares the results of the appraisal of these 16 projects using both methods (objective 1). Section 6 analyzes the written motivations of participants in the PVE to generate empirical insights into potential reasons why PVE and CBA might provide different results (objective 3). Finally, Section 7 concludes and provides a discussion.

2. Participatory Value Evaluation versus Cost-Benefit Analysis

In this section we present two interpretations of PVE. Section 2.1 presents the narrow interpretation of PVE by introducing it as a valuation method addressing the limitations of

existing valuation methods. Section 2.2 provides an extended interpretation of PVE highlighting that due to the nature of the data collection method policy evaluation can take place without relying on the notion of monetary valuation. Below, we provide an in-depth discussion of the notions of valuation and evaluation in the context of CBA and PVE. We elevate this discussion using a fictive government project Z which results in five minutes of travel time savings for 2 million trips in one year and noise pollution for 300 households (63 Decibel). The construction costs of this fictive project are 1 million euro. To keep the example simple, we assume that these are the only three effects of this project. For reasons of brevity we assume a time horizon of one year and eschew discounting.

2.1 'Narrow' PVE - a valuation method feeding into CBA

CBA evaluates a given policy based on a *monetary* comparison of its associated costs and benefits. For many benefit and cost components this requires the translation of the project impacts, in the case of project Z time savings and noise pollution, into monetary terms. Valuation methods are widely applied for this conversion and these methods conventionally use the notion of the number of euros that individuals are willing to pay from their private income for having the government project in place to assign monetary value to the project and its related impacts.

Often Lancaster's (1966) interpretation of value is applied assuming that the value of a project is driven by the sum of the value of its components. In our example, the value of project Z would be entirely comprised by the components of travel time savings, construction costs and noise pollution. Revealed and stated preference methods in the form of hedonic pricing studies and stated choice surveys provide powerful methods to obtain estimates of this notion of marginal WTP. These methods largely operate in the context of private WTP. The key characteristic of the private WTP valuation approach is that the value individuals attach to a *government project's impact* is inferred from the number of euros they are willing to pay in the context of a *private decision*. For instance, the standard empirical approach used to infer the value of a minute or hour of travel time savings from government projects relies on (hypothetical) route choice experiments. 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 (e.g. Abrantes and Wardman, 2011; Batley et al. 2019; Börjesson and Eliasson, 2014; Jara-Díaz, 2007; Kouwenhoven et al., 2014). Similarly, impacts of government projects on landscape, nature and noise pollution are evaluated through investigating the private decisions people make when buying a house (e.g. Allen et al., 2015; Seo et al., 2014). If we consider our example project Z, the travel time savings are valued using hypothetical route choices and noise pollution is valued through decisions of consumers in the real estate market. Notably, such marginal WTP values are often not elicited in the context of the proposed policy context and transferred and applied across different policy projects and contexts.

A major critique on the private WTP-based valuation approach is that the impacts of government projects are evaluated in a non-representative context (e.g. Ackerman and Heinzerling, 2004; Sen, 1995, 2000). For instance, the value individuals attach to travel time in the context of a *government decision* is inferred from the value they attach to this impact in the context of a (hypothetical) *private route choice*. Hence, travel time savings are evaluated in another context than the one in which these benefits actually occur (Mouter and Chorus, 2016).

Various scholars argue that private choices may not reflect how individuals want public policies to change and as a result crucial considerations might be lost in private WTP-based valuation (e.g. Ackerman and Heinzerling, 2004; Sagoff, 1988). For instance, Weimer (2017) argues that valuing impacts of a government project through observing individuals' consumer choices ignores that people may place a value on the way collective decisions are made. Sunstein (1993) and Anderson (1993) argue that individuals might value the same impact, differently in a private sphere and a public sphere. For instance, Sunstein (1993, p. 784) states: "distinctions among kinds of valuation are highly sensitive to the particular setting in which they operate. People do not value goods acontextually. In one setting – say, the workplace – the prevailing kinds of valuation might be quite different from what they are elsewhere – say, the home or the ballot box." Furthermore, Sunstein (1993) argues that because of the highly contextual nature of choice one should not assume that an individuals' private choices can be simply adaptable for policy use.

This critique on private WTP-based valuation approaches can be circumvented through valuing government projects in collective WTP experiments with the appropriate context (Mouter et al., 2019). These experiments express the impacts of government projects to the entire community. When we would value the impacts of our example project Z using this approach we would present the *aggregate* impacts of the project and we would ask individuals whether they agree with a one-time tax increase for the entire community to finance the construction costs of the project. Importantly, participants in collective WTP experiments consider the *overall* positive and negative impacts of a proposed government project together in a representative context: a government decision rather than a private decision (Mouter et al., in press).[1] Private WTP studies and collective WTP studies differ in the extent to which they allow individuals to express altruistic and moral considerations. Private WTP-based studies allow individuals to express moral considerations through their consumption decisions (i.e. ethical consumerism). For instance, to promote animal welfare, they can buy more expensive sustainable eggs instead of cheaper battery-produced eggs. However, in collective WTP studies 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). In the transport literature only a few collective WTP-based valuation studies have been conducted (Daniels and Hensher, 2000; Ivehammer, 2008; 2014; Mouter et al, in press).

Collective WTP studies finance public projects from private money. Many projects are, however, funded from (a re-allocation of) public budgets. One benefit of the WTAPB and PVE valuation approaches is that preferences of individuals who believe that government funds should be spend on different purposes than their own money can be expressed (Mouter et al., 2019). WTAPB experiments allow individuals to indicate which project should be funded from a given amount of *government budget* (e.g. Johansson-Stenman and Martinsen, 2008; Mouter et al., 2017). The WTAPB approach thus aims to infer welfare effects of government projects and related impacts from individuals' preferences regarding the expenditure of public euros. 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

[1] Mouter et al., (in press) coin collective WTP experiments 'social choice valuation experiments' and Ackerman and Heinzerling (2004) call it 'holistic valuation experiments'.

public budget (Mouter et al., 2019). The most important difference between WTAPB experiments and PVE 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. Participants in a PVE in which we evaluate our example project Z will be asked to provide a recommendation to the government on allocating a dedicated public budget; let's say they can allocate a budget of 5 million euro to 10 projects, including project Z, all costing 1 million euro. This implies that respondents can allocate budget to 5 or less projects. Respondents receive information regarding the overall impacts of the projects they can select. They also can decide against allocating the full public budget to these projects and shift some or all of the budget to the next year. Another difference between PVE and other valuation approaches such as collective WTP and WTAPB is that individuals in a PVE experiment select a portfolio of projects. Hence, individuals are allowed to consider positive and negative synergies between projects, and this is not possible in valuation contexts were respondents only select on the implementation of a single policy.

The four valuation methods presented, private WTP, collective WTP, WTAPB and PVE are all able to provide analysts with the necessary marginal WTP and total value (i.e. consumer surplus) measures that can be fed into a CBA. These measures may differ due to the considerations underlying their response formats. Marginal WTP estimates, such as the value of time or the value of a statistical life, have frequently been transferred across application contexts. There is no reason why this can't be done based on using PVE based marginal WTP measures.

2.2 "Extended PVE" – an alternative evaluation procedure to CBA's Kaldor-Hicks

Section 2.1 established how project impacts can be translated into monetary terms using alternative valuation methods, including PVE. We now discuss how the results of PVE can be used to evaluate public projects without relying on monetary valuation. This is a direct result of applying the PVE in the context of the specific policy application and enables an extended interpretation of PVE to be seen as an alternative to CBA.

The central outcome of a CBA is the net present value (NPV). A positive net present value indicates that the benefits expressed in today's money outweigh the costs and that therefore the project can achieve a *potential* Pareto improvement. The Kaldor-Hicks efficiency criterion is essential to this interpretation and recommends policy options where the sum of monetary gains outweigh the sum of monetary losses such that winners can *potentially* compensate the losers (e.g. Boadway and Bruce, 1984). The Kaldor-Hicks efficiency criterion is not concerned with whether compensation or redistribution is actually achieved, hence the emphasis on potential Pareto improvements. If we would use official numbers adopted in the Dutch practice (value of travel time savings = 9 euro; a household experiencing 63 decibel noise pollution = 333 euro per year), our example project Z would be recommended based on the Kaldor-Hicks efficiency criterion because the aggregate monetary benefits (2 million trips * 5 minutes * 9 euros per hour / 60 = 1.5 million) are higher than the aggregate monetary costs (300 people suffering from noise pollution * 333 euro + 1 million euro construction costs = 1.1 million).

An alternative interpretation of a CBA outcome is that it computes the social welfare effect of the government project under scrutiny. Welfare economics assumes that society's

welfare can be represented by a social welfare function, where the welfare of society as a whole is determined by the utilities of each member of society (Bergson, 1938; extended by Samuelson, 1947). Indeed, if the required compensation under the Kaldor-Hicks criterion is actually paid everyone is better off and a real Pareto and social welfare improvement is achieved. Since the compensation usually does not get paid, a positive outcome of a CBA can only be interpreted as an improvement of social welfare under very specific and rather strong normative assumptions (e.g. Adler, 2012; Nyborg, 2014; Nurmi and Ahtianen, 2018). Social welfare functions compare social states in terms of utility not money and since the outcome of a CBA is entirely expressed in money terms, a conversion needs to be made from money into utility. The key concept in this context is the marginal utility of income, which expresses a person's utility increase if this person becomes marginally more affluent. If the marginal utility of income is assumed constant across individuals, then a positive net present value can be interpreted as an improvement in social welfare. This is, however, a controversial assumption because the literature supports that marginal utility of income is decreasing in the income level (Nyborg, 2014). Nyborg (2014) asserts that CBA only measures social welfare under the condition that both the marginal utility of income and the welfare weights are equal for all citizens or that the weight individuals receive is inversely related to their marginal utility of income. Based on these critiques, we establish that CBA lends itself best to be interpreted in the context of economic efficiency, i.e. satisfaction of the Kaldor-Hicks criterion.

When PVE is simply considered as a valuation method, the estimated utility functions can be converted into project specific measures of WTP and consumer surplus (Dekker et al. 2019; Lloyd-Smith 2018). Consequently, the monetary costs and benefits of each policy can be aggregated and implemented in the standard CBA fashion to see if the Kaldor-Hicks efficiency criterion is satisfied (or not). The policy recommendation would simply be based on an alternative valuation method then conventionally applied in CBA. Since PVE experiments are conducted in the respective policy implementation context, the estimated utility functions can be put to good use and interpreted in the context of social welfare. In the supplementary material and Dekker et al. (2019), we set out how the notion of the social welfare function can be used to aggregate (and weight) individual utility functions and thereby contrast and rank alternative policy portfolios (i.e. social states) without relying on arbitrary assumptions regarding the marginal utility of income.[2]

As will be illustrated in Section 5, combining PVE with a social welfare interpretation directly allows the method to be used for evaluation purposes and provide policy makers with information regarding the optimal policy portfolio and expenditure levels and information on the extent to which individual projects and portfolios result in welfare improvements. Hence, PVE in its extended interpretation can be used as a full-fledged alternative to CBA. Although the results are expressed in different units (money vs. utility), the ranking of policy options can easily be compared as illustrated by Section 5.

[2] Indeed, this requires the adoption of cardinal utility functions to work around Arrow's impossibility theorem, but these cardinal utility functions are implicitly undergirding any valuation and consumer surplus studies underlying CBA.

3. Potential reasons for differences in results of a CBA and a PVE in the context of urban transport projects

Section 2.1 highlighted that PVE, in terms of being a valuation method, is distinctively different from *private* WTP-based valuation. In this section, we discuss how this difference in valuation method may lead to different recommendations between *private* WTP-based CBA and PVE in the context of the appraisal of urban transport projects. Specifically, we present critiques in the planning literature raised against the application of (private WTP-based) CBA in this domain and how PVE (or collective WTP-based) can potentially alleviate these critiques. The potential reasons we distill from this literature review will be used to structure our empirical analysis in section 6.

The first critique raised against CBA in the planning literature is that the instrument corrodes and degrades the forward-looking nature of the planning proficiency (e.g. Banister, 2008; Hajer and Pelzer, 2018; Handy, 2008). Planners argue that CBAs have difficulty with considering normative ideas regarding a preferred future urban mobility system (Hickman and Dean, 2018; Nicolaisen et al., 2017). For instance, Nicolaisen et al. (2017) observe that policy makers' normative aspiration to reduce car traffic in the urban core through discouraging car use is ignored or not sufficiently reflected in a CBA even though this is their key rationale for championing projects such as Light Rapid Transit (LRT), removing roads/car lanes and lowering travel speed. Banister (2008) asserts that transport planning requires clear, innovative and strategic thinking about city futures in terms of desirability, and the role that transport can (and should) play in achieving these objectives. Handy (2008) states that the central goal of transport planning is defining the desired future for a place and then think about policies which help to move a place towards that future. Hajer and Pelzer (2018) assert that planning and evaluation need to refocus from a tradition of 'expected futures' to an approach centering on 'desirable futures' and ways to get there. The notion of anticipating the (uncertain) future by setting goals goes beyond the conventional private WTP-based valuation approach which determines the value of impacts of government projects through observing people's (hypothetical) consumer choices (e.g. hypothetical route choices and behavior in the real estate market). Implicitly, planners argue that individuals' *past* consumer choices are not necessarily a good proxy for their normative ideas concerning a *future* mobility system. They argue that the importance of this issue amplified due to the broadening of goals of transportation planning in the last decades. Manaugh et al. (2015), for instance, observe that throughout most of the 20th century the goals of transportation were almost entirely mobility-based, with a focus on congestion reduction, travel time savings and safety improvements for motorists. All effects which are relatively easy to value through observing people's (hypothetical) private consumer choices. In this era CBA was an adequate tool for the planning and evaluation of transport projects (Manaugh et al., 2015). However, prompted by concerns regarding climate change, social inequality and the scarcity of public space in urban areas, the focus of transportation planning shifted more and more to other (more normative and/or future-oriented) goals, such as long-run sustainability, quality of life, social equity and promotion of green transportation in urban regions (e.g. Banister, 2008; Ferreira et al., 2012; Handy, 2008; Manaugh et al., 2015). Banister (2008), for instance, argues that in urban areas a much wider notion of the street has been created, as it is no longer only being considered as a road but also as a space where people

meet. Therefore, nowadays, urban transport projects pursue both traditional effects (e.g. costs, travel time savings, safety and reduction of noise pollution) as well as a diverse set of non-traditional effects such as long-run sustainability, townscape, social inclusion, city image and improving the quality of urban spaces (e.g. De Bruijn and Veeneman, 2009; Hickman and Dean, 2018; Nicolaisen et al., 2017). However, several authors argue that CBA fails to appreciate many of these non-traditional effects as they are generally not included in the CBA, or are given marginal importance because they are not quantified or monetized (Beukers, 2015; Handy, 2008; Hickman and Dean, 2018; Nicolaisen et al., 2017). For instance, various scholars argue that CBAs for cycling projects have difficulty with including the impacts of a modal shift from car to bicycle such as the positive health impacts of increased physical activity and a reduction of road congestion and emissions, even though realizing these impacts are often a key goal of cycling projects (de Hartog et al., 2010; Heinen et al., 2015; Adam et al., in press; van Wee and Börjesson, 2015). Moreover, planning scholars assert that goals of cycling policies such as improvement of urban quality, space efficiency, social inclusion, improved mobility for children and social interaction potential are often ignored in CBA (de Hartog et al., 2010; te Brömmelstroet et al., 2017; van Wee and Börjesson, 2015). Planning scholars argue that the weak position for such impacts in CBA fails to acknowledge the multifaceted planning priorities in urban transport infrastructure investments (Handy, 2008; Nicolaisen et al., 2017). Handy (2008) argues that the poor consideration of broader goals of transport planning in CBA is also problematic because this may lead planners and policy makers away from those goals. In her view, this could significantly impede the achievement of broader goals of transport planning such as quality of life and long-run sustainability. Banister (2008) even claims that transport planning is at a crisis point as it underestimates the key challenges facing urban planners. We hypothesize that the above-mentioned critique can be alleviated through using valuation approaches such as collective WTP, WTAPB and PVE in which participants consider the *overall* positive and negative impacts of a proposed government project together in the context of a government decision instead of a private decision. On the one hand, such valuation approaches allow participants to express preferences that are associated with their private consumer choices. On the other hand, such approaches allow individuals to express preferences that line up with their preferred future perspectives regarding the (local) urban mobility system, broader goals of transport planning and ethical considerations. Indeed, collective WTP and WTAPB studies conducted in the transport domain observe that participants include (moral) considerations regarding the way government should trade off the costs and benefits of government projects when making choices (e.g. Mouter et al., 2018, in press).

Another critique relates to the fact that CBAs generally use standardized transport models to establish the impacts of a transport project and transfer these impacts into monetary terms using *generic* price tags such as the 'Value of Time' and the 'Value of a Statistical Life'. The transport models do seldom operate on the detailed level necessary to accurately estimate impacts of specific transport projects (van Wee and Börjesson, 2015) and it is highly questionable whether applying generic price tags leads to correct valuations in specific contexts (Mouter et al., 2013b). Hence, planners argue that this approach is unrefined and fails to recognize the special (local) conditions of the problem which the transport project aspires to address (Beukers et al., 2012; Handy, 2008; Mouter et al., 2013b). This critique relates to the dichotomy between formal assessment and informal assessment (Pesch et al., 2017). Formal

assessment methods include institutionally established methods, such as CBA and environmental impact assessment (EIA). Apart from these formal assessment trajectories, transport projects are also assessed by local citizens, local businesses and other actors that are not part of established institutions. This so-called informal assessment trajectory particularly focusses on the specific characteristics, needs and concerns of the local communities that are affected by the transport project and/or problem. In case the informal assessment provides new insights, this may lead to adaptations in the formal trajectory, which Pesch et al. (2017) refer to as 'backflowing'. However, problems can emerge when the formal assessment methods do not respond sufficiently to the values and concerns that emerge from the informal assessment (Cuppen et al., 2016; Rip, 1986). When local populations feel that their arguments and sentiments are excluded in the formal assessment this can give rise to protests and growing distrust of citizens (Pesch et al., 2017). In short, planners contest CBA's 'backflowing capacity' as its generic approach does not properly account for the insights of citizens regarding the specific characteristics of the problem/project at hand. In Section 2, we discussed that this 'backflowing capacity' is built into PVE which ask individuals to assess (the impacts of) a government project with the appropriate context. This allows individuals to assess the projects under scrutiny based on the impacts which are provided by the analyst (formal assessment) or the consequences of the projects that emerge from their personal experience (informal assessment).

4. Selecting transport projects which will be assessed through a CBA and PVE

Together with the program managers of the 'car', 'public transport', 'cycling' and 'safety' departments of the TAA, we selected 16 transport projects that were considered for inclusion in a transport investment scheme. Table 1 provides a brief verbal description of the 16 projects. A more elaborate description of the projects is included in the demo versions of the PVE (www.burger-begroting.nl (in Dutch) and http://burgerbegroting.tbm.tudelft.nl/participatory-value-evaluation-transport-authority-amsterdam (the English translation). Where available, we used the project descriptions of the TAA to compose the verbal descriptions. Table 1 (column 2) presents whether the projects were suggested by the car, public transport (PT), cycling or safety department. We also received documents to determine seven types of societal impacts of the projects: 1) costs; 2) number of travelers who experience travel time savings during an average working day; 3) average number of minutes of travel time savings per traveler; 4) change in traffic deaths per year; 5) change in severe traffic injuries per year; 6) additional households affected by noise pollution; 7) number of trees that have to be chopped. We asked civil servants of the TAA to provide information regarding the impacts in bandwidths because we needed to differentiate among the participants in the PVE in terms of the attribute levels of the impacts to estimate people's sensitivity for these impacts (see the supplementary material for more detailed information). Table 1 describes the bandwidths of the impacts for each project.

Table 1: Impacts of the 16 projects

	Type	Costs	Travellers affected (thousands)	Minutes time savings	Change traffic deaths	Change severe injuries	Households affected by noise	Trees cut
1) Faster connection to the provincial road N516 (Zaandam) at the Poelenburg/Achtersluispolder will decrease travel time for car/bus traffic	Car	40/60	50/70	2/4	0	0	20/100	0
2) Fly-over on the A10 at the junction Amsterdam Noord will decrease travel time for car and bus traffic.	Car	30/50	60/80	2/4	0	0	50/200	40 /200
3) Extending Mac Gillavrylaan to the Middenweg improves accessibility of the Science Park and reduces noise pollution for citizens living at the Middenweg.	Car	7/13	30/40	3/6	0	0	-50/-150	0
4) Extra lane on Bovenkerkerweg decreases travel time for car users.	Car	7/13	25/40	2/6	0/0.2	0/2	0/20	20/40
5) New bus connection IJburg – Bijlmer Arena will improve public transport between IJburg, Amstelveen and Schiphol Airport.	PT	40/60	3/6	4/11	0	0	0	0
6) Route of busses that run between Amsterdam CS and Zaandam will be shortened through the realization of an extra entrance and exit ramp.	PT	3/7	3/7	1/2	0	0	0	0
7) The tram connection between Diemen and the Linnaeusstraat will be accelerated through a more efficient allocation of stops and traffic lights.	PT	11/19	4/10	3/5	0	0	0	0
8) A comfortable cycling path (cycling highway) will be realized between Hoofddorp – Schiphol and Aalsmeer.	Bike	5/11	2.5/4	3/6	0	0	0	0
9) A cycling highway will be realized between the sports facilities at the Amstelveenseweg (Amsterdam)	Bike	4/8	8/15	2/4	0	0	0	20 /100
10) New bridge for cyclists/pedestrians at Hoornselaan (Purmerend).	Bike	3/6	6/10	2/4	0/-0.1	0/-2	0	0
11) Bike tunnel will be built at the Guisweg (Zaandam) where cyclists now cross the railroad.	Bike	30/50	5/8	1/3	0/-0.2	0/-3	0	0
12) A new bridge for cyclists will be built between Borneo-Eiland and Zeeburgereiland (Amsterdam).	Bike	25/45	6/8	5/8	0	0	0	0
13) Ilpendam pedestrian tunnel will diminishing travel time for car traffic and bus traffic and improve safety for pedestrians.	Safe	2/4	15/25	1/2	0/-0.1	0/-2	0	0
14) The Stadhouderskade will be tunnelled for car users at the entrance of the Vondelpark. Cyclists/pedestrians and car traffic will be separated.	Safe	30/50	35/40	1/2	0/-0.8	-2/-6	0	0
15) Traffic education for children in the age group 4 – 18 will prevent traffic accidents through improving awareness of children.	Safe	40/60	0	0	0/-1	-2/-15	0	0
16) Five additional police officers will be hired who will specifically focus on enforcing traffic laws.	Safe	15/25	0	0	0/-1	-3/-10	0	0

We ensured that the projects were to some extent distributed between the six sub-regions that fall under the jurisdiction of the TAA. Figure 1 shows the locations of the projects. The projects 'traffic education' and 'five police officers' are not attached to a specific location.

Figure 1: Locations of the projects

For making the CBA calculations we translated the impacts of the projects into monetary terms using the values enumerated in Dutch CBA guidelines (Rijkswaterstaat, 2018). For these computations we used the averages for each of the impacts.

For conducting the PVE it was necessary to conduct an experiment. We developed a web-based environment (see the demo version) in which respondents received the following instruction:

> *"On the next page we present 16 transport projects that the Transport Authority Amsterdam could implement. The Transport Authority Amsterdam can only spend 100 million euro on these projects. Hence, there is not enough budget to finance all projects. The Transport Authority Amsterdam decided to consult a large number of citizens to*

provide an advice for this choice situation. You are one of the citizens that we selected. More specifically, we ask you to select the projects you advise to the Transport Authority through clicking on the 'selection button'. Please note that any remaining budget will be shifted forward to the next year which would imply that the Transport Authority Amsterdam will be able to spend more money on projects that fall within their remit in the next year."

After reading this text, respondents saw an instruction video and then they were guided to a web-tool in which they could, amongst other things, sort and compare the projects by one of the impacts and find out more about the goals and the impacts of the projects through clicking on an information button. We communicated to the respondents that they could assume that the impacts will materialize in a period of 50 years. Participants were not forced to make a choice but had the option to delegate their choice to an expert. The delegates in turn also completed the experiment. We conducted the experiment in four waves (June 2017, October 2017, January 2018, March 2018). In two waves, respondents were also allowed to adjust the governmental budget by increasing the tax per household or by selecting a rebate. The survey company Kantar Public was asked to draw four random samples from the population of the TAA of 18 years of age and older. The company was not explicitly requested to draw representative samples, but it was important that all relevant demographic segments (e.g. income, education, age and gender) were present. Respondents who completed the experiment received a monetary compensation. In case respondents delegated their choice, they received a lower financial compensation from the survey company. In total 9,607 individuals were recruited and 2,498 respondents completed the full PVE experiment (26%). Table 2 provides information about the socio-demographic characteristics of the respondents as well as their political affiliation. The sample is not representative for the population of the TAA in several respects. Males, older inhabitants and individuals with a higher income are overrepresented. In Section 5.2 we will show that this is not a problem as we can correct for this in the social welfare analysis.

Table 2: Socio-demographic characteristics of participants in the PVEs.

	Wave 1	Wave 2	Wave 3	Wave 4
Number of respondents	742	803	381	301
Gender				
Female	44%	47%	53%	50%
Male	56%	53%	47%	50%
Age				
18 - 25	4%	5%	12%	11%
26 - 35	10%	11%	13%	20%
36 - 45	14%	16%	10%	16%
46 - 66	23%	23%	19%	18%
56 - 65	22%	22%	24%	16%
65 +	27%	23%	22%	18%
Education				
Lower education	35%	35%	30%	37%
Higher education	43%	43%	47%	42%
University (of applied sciences)	21%	22%	24%	21%
Household gross income				
Less than 15,000	6%	6%	15%	8%
15,000 - 30,000	15%	14%	28%	12%
30,000 - 60,000	40%	40%	41%	38%
More than 60,000	39%	40%	16%	42%
Voted previous election				
VVD (Conservative-Liberal)	20%	24%	13%	18%
PVV, Forum for Democracy (Nationalis	8%	7%	7%	6%
CDA, CU, SGP (Christian)	10%	8%	6%	9%
D66 (Social-Liberal)	14%	14%	17%	15%
GL, PvdD (Green Parties)	22%	22%	23%	21%
SP (Socialist)	11%	10%	14%	10%
PVDA (Labor)	11%	12%	8%	8%
I didn't vote	4%	3%	11%	11%

After respondents made a portfolio choice, they received some questions about how they experienced their participation in the PVE and they were asked to provide arguments for why they selected the projects. The primary reason to ask participants in the PVE to provide written motivations was to generate empirical insights into potential reasons why PVE and CBA might provide different results which is the third objective of this paper. In the next section, we explain that a secondary reason is that the written motivations provide a qualitative explanation of some of the outcomes of the quantitative outcomes of the PVE. A large group of respondents took the time to (thoroughly) explain their choices. Respondents could also mention multiple motivations. The written motivations were manually coded using content analysis. Content

analysis is a systematic, replicable technique for compressing many words of text into fewer content categories based on explicit rules of coding and categorizing (Weber, 1990). In the field of transport policy, content analysis has been increasingly applied by researchers because the method is regarded as a structured approach for decoding texts as free from inferences as possible (Ardic et al., 2013; Vonk-Noordegraaf et al., 2014). A content analysis starts with a theory or relevant insights from the literature as a starting point for initial categories of codes. In this study, we started our content analysis with a list of 29 initial codes that were associated with the impacts of the 16 projects that were explicitly described in the experiment as well as the goals of urban transport projects that emerged from the literature review discussed in Section 3. To make the process transparent and replicable, we developed a 'coding and categorizing protocol' which is available upon request. Next, we analyzed the data in two rounds. The goal of the first round was to identify new categories of motivations which added to the initial list of categories which resulted in 85 categories. In the second round, the 9,920 motivations were divided across these categories.

5. Comparing the results of the CBAs and the PVE for the 16 projects

Section 5.1 describes the results of the CBAs that we conducted for the 16 projects. Section 5.2 presents the results of the PVE which includes all these 16 projects and compares the results of the CBAs and the PVE.

5.1 Results of the CBA

We conducted the CBAs for the 16 transport projects based on the average impacts presented in Table 1. For instance, to compute the costs and benefits of the first project – faster connection of bus and car traffic Zaandam – we used 50 million as the costs (average between 40 million and 60 million), 60,000 travelers who experience travel time savings during an average working day, 3 minutes number of minutes of travel time savings per travelers and 60 additional households affected by noise pollution. Next, we converted these impacts into money metrics using the recommended values enumerated in Dutch CBA guidelines (Rijkswaterstaat, 2018): value of travel time savings: 9 euros per hour; value of a statistical life: 2.6 million; value of statistical severe traffic injury: 0.3 million; value of one additional household facing noise pollution: 250 euro per year. The Dutch Guidelines do not prescribe standard numbers for converting chopped trees into monetary terms. However, in a previous CBA a dedicated study was conducted to value replacing and/or replanting a similar number of trees that have to be chopped in the projects that are part of our study (Decisio, 2014). This study estimated the monetary costs at 2 million euros, and we decided to use this figure in our study. Subsequently, we computed the costs and benefits of the projects using the prescribed discount rate in the Netherlands of 4.5% (Mouter, 2018) over a time horizon of 50 years. Table 3 presents the monetary values for the impacts of the 16 projects as well as the final indicators (net present value and benefit-cost ratio). The Dutch CBA Guideline (Romijn and Renes, 2013) prescribes the net present value as the final indicator that should be presented in a CBA because the benefit-cost ratio (BCR) can be easily manipulated. For this reason, we ranked the projects based on their net present value.

Table 3: Cost-Benefit computations for the 16 transport projects (net present value is in millions of euros)

	Type	Costs	Travel time savings	Traffic death	Traffic injuries	Noise pollution	Chopped trees	Net present value	BCR
Fly-over A10 at the junction Amsterdam Noord	Car	-40	149.4	0	0	-0.6	-2	106.9	3.67
Extending the MacGillavrylaan to the Middenweg	Car	-10	112	0	0	0.5		102.5	11.25
Faster connection Poelenburg/provincial road Zaandam	Car	-50	128	0	0	-0.3	0	77.7	2.55
Extra lane on Bovenkerkerweg	Car	-10	92.5	-5.1	-5.9	-0.1	-2	75.3	7.94
Five police officers which sanction violation of traffic regulations	Safety	-20	0	51.4	38.5	0		69.9	4.50
Stadhouderskade car tunnel at the entrance of the Vondelpark	Safety	-40	42.7	20.5	23.7	0		46.9	2.17
Ilpendam pedestrian tunnel	Safety	-3	21.3	2.6	5.9	0		26.8	9.93
Traffic safety education for children in the age group 4 -18	Safety	-50	0	25.7	50.4	0		26.1	1.52
New bridge for cyclists and pedestrians Purmerend (Hoornselaan)	Cycling	-4.5	17	2.6	5.9	0		21.1	5.67
Cycling highway Amstelveenseweg	Cycling	-6	24.5	0	0	0	-2	16.5	3.75
Improvement tram connection Diemen – Linnaeusstraat	PT	-15	20	0	0	0		5	1.33
Cycling highway Hoofddorp – Schiphol – Aalsmeer	Cycling	-8	10.4	0	0	0		2.4	1.30
Acceleration of the bus connection Amsterdam CS - Zaandam	PT	-5	5.3	0	0	0		0.3	1.06
New cycling bridge Zeeburgereiland and Borneo Eiland	Cycling	-35	32.4	0	0	0		-2.6	0.93
Guisweg bike tunnel	PT	-40	9.2	5.1	8.9	0		-16.7	0.58
New bus connection IJburg - Bijlmer Arena	PT	-50	24	0	0	0		-26	0.48

A main result that follows from Table 3 is that 13 projects have a positive CBA score (positive net present value). The projects which focus on improvement for car users have high net present values and this also holds for the projects concerning safety. Public transport projects perform relatively poorly. The results can also be used to define the optimal portfolio when the investment budget of a government is limited. For instance, within a budget constraint of 100 million the optimal portfolio consists of three car projects: (Fly-over A10, extending MacGilavrylaan, extra lane Bovenkerkerweg), two safety projects (five extra policy officers, pedestrian tunnel Ilpendam), two cycling projects (cycling highway Amstelveenseweg, new bridge Purmerend) and a public transport project (acceleration bus connection Amsterdam CS – Zaandam). This portfolio has a net present value of 419.3 million euro.

5.2. Results of the PVE

This section presents the results of the PVE and also provides a comparison with the results of the CBA presented in Section 5.1.

5.2.1 Descriptive results

In the PVE, citizens were asked to select transport projects within a budget constraint of 100 million euros. The total costs of the 16 projects was 386.5 million euros, so it was not possible for participants to include all projects in their portfolio. Table 4 presents the number of projects selected by the respondents and shows that most respondents selected 3 or 4 projects.[3]

Table 4: Number of projects selected by respondents

Number of projects selected	Number of respondents
0	35 (2%)
1	42 (2%)
2	181 (8%)
3	475 (21%)
4	479 (21%)
5	362 (16%)
6	285 (13%)
7	216 (10%)
8	127 (6%)
9	23 (1%)
10	1 (0%)
11	1 (0%)

Figure 2 presents the percentage of respondents that selected each project. For each project the average costs (in million euros) are displayed between brackets.

[3] Around 15% of the respondents delegated their choice to an expert.

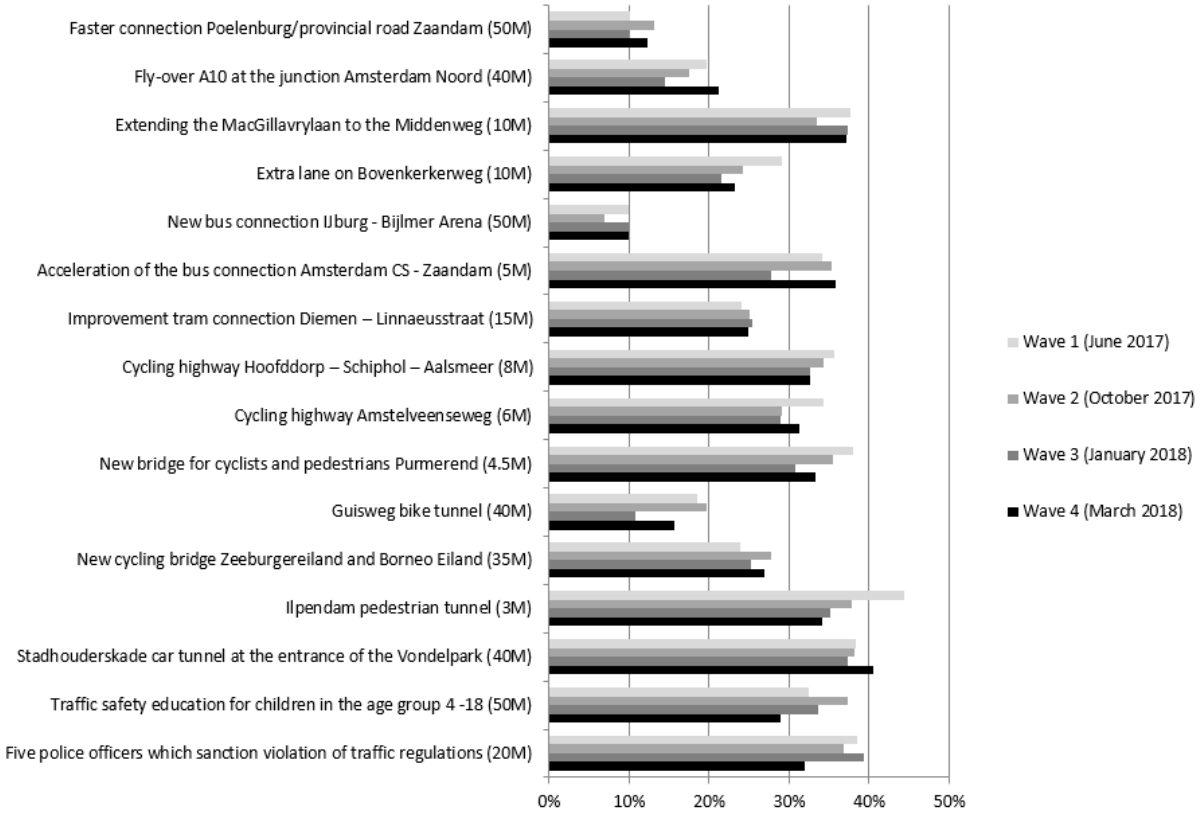

Figure 2: Percentage of respondents which selected the different transport projects

The first observation is that all projects are chosen by at least 5% of respondents. This is good news for the TAA as the choice set did not seem to include irrelevant projects (from the perspective of the citizens). Second, 12 out of 16 projects were selected by more than 20% of the participants in all experiments. Third, Figure 2 shows that the differences in the shares of respondents who select a project between the four waves of the PVE are not very large. As the waves took place at different time instances this is an indication that citizens' preferences for the 16 transport projects are fairly stable over time. For reasons of readability we will not further distinguish between the four waves in the remainder of this section. Participants were also asked to evaluate the PVE on four items. Table 5 presents the results.

Table 5: Answers of respondents to the four items rated at the end of the PVE

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
I was convinced of my choices	572 (32%)	991 (56%)	182 (10%)	22 (1%)	1 (0%)
I thought that the experiment was realistic	397 (22%)	907 (51%)	386 (22%)	87 (5%)	10 (1%)
I think it is good that the government aims to involve citizens in making choices between transport projects	1024 (56%)	687 (38%)	85 (5%)	20 (1%)	2 (0%)
This experiment provides the government with relevant information for making choices between transport projects	561 (31%)	908 (50%)	290 (16%)	56 (3%)	10 (1%)

Although the answers that respondents gave to these questions are not required to achieve the three research goals of our study, we think that it is relevant to report that citizens positively

evaluated their participation in the PVE. Table 5 shows that 88% of the respondents were convinced about their choice and 73% consider the experiment to be realistic. 94% of the respondents believed it is good that the TAA involves citizens in making choices between transport projects and 81% of the respondents agreed with the proposition: 'the experiment provides the government with relevant information in their decision-making process'. Only 10 respondents strongly disagreed with this proposition.

5.2.2 Computing the societal value of individual projects

To infer the societal value of the transport projects from the choices of the participants in the PVE, we estimated behavioural choice models which assume that participants aimed to select a portfolio of transport projects that in their view represents the portfolio which maximizes their 'utility' (i.e. they select the best portfolio). The technical details of the behavioural choice models can be found in the supplementary material and Dekker et al. (2019). Below we only discuss the most important assumptions that are made in the analysis. We assume that part of the value of an individual project is defined by the impacts that were explicitly presented to participants for each of the transport projects (henceforth: explicit impacts): reducing travel time, the number of traffic deaths, the number of severe traffic injuries, the number of households affected by noise pollution or the number of trees that need to be chopped[4]. We estimated so-called taste parameters in order to determine the importance of these explicit impacts on the individual's decision (this is comparable to stated choice surveys which estimate taste parameters for attributes of a public good). Because the (un)attractiveness of an individual project can also be defined by other considerations than the level of the five 'explicit impacts' so-called project specific parameters are estimated. These parameters capture the benefits 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). First, these project specific parameters capture the value resulting from other considerations than the five explicit impacts (such as the normative views concerning the future urban mobility system discussed in Section 3). Second, when citizens derive value from the fact that a project influences one of the explicit impacts irrespective of *the extent to which* the impact is affected, this is also captured in the project specific parameter. To illustrate this: when a participant selects a project because (s)he thinks that reducing traffic deaths in general is important, this is captured in the project specific parameters. When the participant values the *number of* traffic deaths that are prevented as a result of the project as well, this is captured in the taste parameter for traffic deaths. Both the taste parameters and the project specific parameters are presented in Table 5 in the column 'Estimate'. The column 't-value' depicts whether the parameter is significant at the 0.05 level (t-value higher than 1.96). All the project specific parameters and the taste parameters for traffic deaths and severe traffic injuries are significantly different from

[4] Note that the value of project(impacts) can potentially include the private benefits of the project, but also other-regarding considerations related to the impacts of the project on other citizens or future generations. Because we cannot identify the extent to which participants themselves experience benefits/costs from the transport projects, we cannot draw conclusions regarding the extent to which the value of the projects can be attributed to private impacts or other-regarding impacts.

zero. The arguments that respondents gave to underpin their selection of the projects (which will be discussed in Section 6) provide qualitative insights in the interpretation of the project specific parameters. For instance, if respondents argue that they selected a project because a project will lead to a certain impact and they did not receive explicit information about this impact in the PVE, then it is likely that this perceived impact contributes to the positive project specific parameter.

Log-likelihood:	-20,170.52	
Estimation results	**Estimate**	**T-value**
Taste parameters		
Reduction of travel time (per 1,000,000 minutes)	0.4806	1.13
Additional traffic deaths	-1.5814	**-2.76**
Additonal traffic injuries	-0.1896	**-2.31**
Additional households affected by noise pollution (per 100)	-0.0619	-0.85
Additional trees cut (per 100)	-0.0882	-1.09
Project specific parameters		
Faster connection Poelenburg/provincial road Zaandam	6.5555	**65.28**
Fly-over A10 at the junction Amsterdam Noord	6.6974	**38.09**
Extending the MacGillavrylaan to the Middenweg	5.5604	**53.77**
Extra lane on Bovenkerkerweg	5.3741	**71.39**
New bus connection IJburg - Bijlmer Arena	6.3883	**139.15**
Acceleration of the bus connection Amsterdam CS - Zaandam	4.9451	**118.33**
Improvement tram connection Diemen – Linnaeusstraat	5.7723	**134.40**
Cycling highway Hoofddorp – Schiphol – Aalsmeer	5.3959	**128.12**
Cycling highway Amstelveenseweg	5.0542	**74.96**
New bridge for cyclists and pedestrians Purmerend (Hoornselaan)	4.8378	**110.12**
Guisweg bike tunnel	6.5271	**149.09**
New cycling bridge Zeeburgereiland and Borneo Eiland	6.6641	**146.95**
Ilpendam pedestrian tunnel	4.5549	**101.35**
Stadhouderskade car tunnel at the entrance of the Vondelpark	7.0658	**108.88**
Traffic safety education for children in the age group 4 -18	7.1350	**77.60**
Five police officers which sanction violation of traffic regulations	6.1875	**65.71**

Table 5: Estimation results behavioural choice model

Table 5 shows that the level of these safety impacts is considered to be relevant when citizens choose their portfolio of projects. The taste parameter for reduction of travel time is not significantly different from zero. This means that *the level of* travel time savings does not significantly affect the (un)attractiveness of a project. However, *the fact that* a project reduces travel times can still impact a project's (un)attractiveness. In this case, this is captured in the project specific parameters. Using the taste parameters, it is also possible to establish the relative importance of the different impacts. For instance, we can infer from the results that citizens of the TAA think that the reduction of 1 traffic death provides the same societal value as the reduction of 8.34 severe traffic injuries (1.5814 / 0.1896). This is very close to the relative importance of these impacts in a CBA which can be computed by dividing the value of a statistical life (2.6 million) and the value of statistical severe traffic injury (0.3 million) = 8.67. In the context of a PVE, the societal value that individuals obtain from travel time savings is substantially lower than the value of traffic safety when compared with the standard numbers enumerated in CBA Guidelines. In a CBA, 1 million minutes of travel time per day will result

in a yearly benefit of 36 million euros (1 million / 60 * 9 euros value of time * 240 working days = 36 million) which is way higher than saving one statistical life (2.6 million euro). However, Table 5 reveals that respondents participating in the PVE obtain a higher societal value from preventing a traffic death (1.5184) than from saving 1 million minutes of travel time on an average working day (0.4806[5]). The result that individuals attach more importance to safety in a PVE context than in a CBA context aligns with previous research which compares these two impacts in a willingness to pay and a willingness to allocate public budget context (Mouter et al., 2017).

The behavioral choice models allow, after aggregation of impacts across the population, for the derivation of the probability that a project improves societal value compared to shifting the money to the next period, i.e. whether societal benefits are higher than the costs (see the supplementary material and Dekker et al., 2019 for a detailed discussion of conducting such a welfare analysis). These probabilities are the final indicators of the evaluation of the individual projects which reflect the extent to which a project improves societal value (probability higher than 50%). The desirability probabilities can be equated with the net present value in a CBA, although it should be noted that net present values computed in a CBA reflect the extent to which a project passes the Kaldor-Hicks efficiency test which differs from determining the societal value of a project (see Section 2). Computing the probabilities is a key step in the policy evaluation of a PVE since participants always have the fallback option of not spending any money in case they think that all the projects are undesirable. More specifically, in case all the participants in the PVE would have selected the null portfolio (a portfolio without any projects) thereby recommending to shift the entire public budget to the next year, the probability that one of the projects improves societal value compared to shifting the money to the next period would be (very close to) 0%.

Table 6 ranks the projects in terms of their desirability probability. The final column shows the ranking of the projects in the CBA analysis. Table 6 shows that seven projects have a probability higher than 50% to improve societal value. The Stadhouderskade car tunnel has a 56% probability to improve societal value compared to shifting budget to the next year. There are also some projects with a negative societal value. For instance, the new bus connection IJburg – Bijlmer Arena has a 31% probability to improve societal value compared to shifting budget to the next year.

[5] Note that Table 5 shows that the taste parameters for travel time savings is not significant at the 0.05 level.

	Project	Project type	Project desirability		CBA
1	Stadhouderskade car tunnel at the entrance of the Vondelpark (40M)	Safety	56%	✓	6
2	Ilpendam pedestrian tunnel (3M)	Safety	55%	✓	7
3	Traffic safety education for children in the age group 4 -18 (50M)	Safety	54%	✓	8
4	Five police officers which sanction violation of traffic regulations (20M)	Safety	54%	✓	5
5	New bridge for cyclists and pedestrians Purmerend (4.5M)	Cycling	52%	✓	9
6	Extending the MacGillavrylaan to the Middenweg (10M)	Car	52%	✓	2
7	Acceleration of the bus connection Amsterdam CS - Zaandam (5M)	PT	51%	✓	13
8	Cycling highway Hoofddorp – Schiphol – Aalsmeer (8M)	Cycling	50%	?	12
9	Cycling highway Amstelveenseweg (6M)	Cycling	48%	x	10
10	New cycling bridge Zeeburgereiland and Borneo Eiland (35M)	Cycling	46%	x	14
11	Improvement tram connection Diemen – Linnaeusstraat (15m)	PT	44%	x	11
12	Extra lane on Bovenkerkerweg (10M)	Car	44%	x	4
13	Fly-over A10 at the junction Amsterdam Noord (40M)	Car	41%	x	1
14	Guisweg bike tunnel (40M)	Cycling	40%	x	15
15	Faster connection Poelenburg/provincial road Zaandam (50M)	Car	35%	x	3
16	New bus connection IJburg - Bijlmer Arena (50M)	PT	31%	x	16

Table 6: Probability that a project improves societal value

Another noteworthy result is that the project desirability of the majority of the projects is very close to 50%. This reflects the high uncertainty associated with the policy recommendations. The first cause for the high uncertainty is the tightness of the budget constraint in the experiment. For instance, many respondents included two very expensive projects in their portfolio ('Stadhouderskade car tunnel' and 'Traffic education for children') which already takes up 90% of the budget. From this selection we can infer that these respondents think that these two projects are highly attractive, but at the same time it is hard to evaluate how they judge the attractiveness of the other 14 projects. A second explanation for the high uncertainty in the policy recommendations is that the large majority of projects is selected by a substantial part of the participants. More precisely, 12 out of 16 projects were selected by more than 20% of the respondents in all experiments (see Figure 2). In the presence of such conflicting preferences, it is relatively difficult to determine which projects have a negative societal value compared to a situation in which various projects were only selected by a few respondents.

When we compare the ranking of the projects presented in the PVE with the ranking of the projects in the CBA (last column of Table 6) we see that the projects that score negatively in the CBA (new cycling bridge Zeeburg, Guisweg bike tunnel and the new bus connection IJburg – Bijlmer) also score negatively in the PVE. Hence, in this case CBA and PVE provide rather similar recommendations. The most important difference between the rankings is that the 4 car projects represent the 'top 4' in the CBA analysis, but these projects score relatively poorly in the PVE as 3 of these projects are likely to generate a negative societal value. It is also noteworthy that the five cycling projects and the four safety projects score better in the PVE than in the CBA in terms of their ranking.

5.2.3 Computing the portfolio with the highest societal value

A second output of the welfare analysis conducted in a PVE is the ranking of portfolios of projects in terms of social welfare. When the public budget is unlimited policy makers should implement all projects with a desirability probability of higher than 50%. However, in reality policy makers are faced with limited budgets and PVE allows for determining the best selection of projects (i.e. the best portfolios) for a given budget. Table 7 shows the top 10 of portfolios within a budget constraint of 100 million euros. A '1' indicates that a project is included in the portfolio and a '0' indicates that a project is not included.

Top 10 portfolio's	1	2	3	4	5	6	7	8	9	10
Faster connection Poelenburg/provincial road Zaandam (50M)	0	0	0	0	0	0	0	0	0	0
Fly-over A10 at the junction Amsterdam Noord (40M)	0	0	0	0	0	0	0	0	0	0
Extending the MacGillavrylaan to the Middenweg (10M)	0	1	0	0	0	0	0	0	0	0
Extra lane on Bovenkerkerweg (10M)	0	0	0	0	0	0	0	0	0	0
New bus connection IJburg - Bijlmer Arena (50M)	0	0	0	0	0	0	0	0	0	0
Acceleration of the bus connection Amsterdam CS - Zaandam (5M)	0	0	1	0	1	0	1	0	0	0
Improvement tram connection Diemen – Linnaeusstraat (15m)	0	0	0	0	0	0	0	0	0	0
Cycling highway Hoofddorp – Schiphol – Aalsmeer (8M)	0	0	0	0	0	0	0	0	1	0
Cycling highway Amstelveenseweg (6M)	0	0	0	0	0	0	0	1	0	0
New bridge for cyclists and pedestrians Purmerend (Hoornselaan) (4.5M)	1	0	0	0	1	1	0	0	0	0
Guisweg bike tunnel (40M)	0	0	0	0	0	0	0	0	0	0
New cycling bridge Zeeburgereiland and Borneo Eiland (35M)	0	0	0	0	0	0	0	0	0	0
Ilpendam pedestrian tunnel (3M)	1	0	1	1	0	0	0	1	0	0
Stadhouderskade car tunnel at the entrance of the Vondelpark (40M)	1	1	1	1	1	1	1	1	1	1
Traffic safety education for children in the age group 4 -18 (50M)	1	1	1	1	1	1	1	1	1	1
Five police officers which sanction violation of traffic regulations (20M)	0	0	0	0	0	0	0	0	0	0
Total costs portfolio	**98**	**100**	**98**	**93**	**100**	**95**	**95**	**99**	**98**	**90**

Table 7: 10 portfolios computed in the PVE which result in the highest expected societal desirability within budget constraint of 100 million

The first conclusion that we can draw based on these results is that the portfolio with the highest societal value consists of the Ilpendam pedestrian tunnel, the new cycling bridge in Purmerend, the Stadhouderskade car tunnel and the traffic education program. These are all projects that focus on safety and improvements for cyclists and pedestrians. Car projects and public transport projects are not included in the best portfolio. Finally, the Stadhouderskade car tunnel and the Traffic education program received high support by citizens and are included in all the top 10 portfolios within a budget constraint of 100 million euros. Notwithstanding the high total cost of these projects (90 million) citizens seem unwilling to sacrifice these projects for alternative projects.

Section 3 revealed that the sample is not representative for the population of the TAA. One of the strengths of PVE is that it is possible to control for this in the evaluation step. In a parallel research project Volberda (2020) analysed the choices of the participants in the TAA using latent class cluster analysis and found that the living area of the participant was the most

important variable explaining heterogeneity among the clusters. That is, participants disproportionally selected projects close to where they live. Because some regions were underrepresented in our sample, we conducted a new welfare analysis applying corrective weights for their living area. It is beyond the scope of this paper to report the full analysis, but the most important insight was that the portfolio's 1 and 2 changed position. Other than that, the welfare analysis correcting for this location effect did not substantially affected the ranking of portfolios. For instance, the Stadhouderskade car tunnel and the Traffic education program are still included in all the top 10 portfolios.

To summarize, CBA and PVE produce different results when assessing the 16 projects of the TAA. The most important difference is that car projects score relatively good in the CBA analysis. On a project level (section 5.2.2.) the 4 projects suggested by the car department of the TAA represent the 'top 4' in the CBA analysis. However, car projects do not perform very well in the PVE as 3 out of these 4 projects have a negative societal value. On the other hand, the PVE recommends projects that focus on safety. The four projects recommended by the safety department even represent the 'top 4' in the PVE analysis and two expensive safety projects are included in all the top 10 portfolio within a budget constraint of 100 million euros. Finally, on the project level, projects resulting in improvements for cyclists and pedestrians are ranked higher in a PVE than in a CBA.

6. Potential reasons for differences in results CBA and PVE

The goal of this section is to generate empirical insights into potential reasons why PVE and CBA might provide different results. To achieve this goal, we analyzed the written motivations respondents gave after they submitted their preferred portfolio. As discussed in Section 4 we grouped the 9,920 written motivations provided by respondents in 85 categories. Again, we wish to emphasize that we discuss potential reasons for differences in recommendations derived from PVE and private WTP-based CBA. We believe that many of these differences will disappear when one decides to use collective WTP-based valuation studies in CBA as in both collective WTP-based studies and PVEs participants consider the overall positive and negative impacts of a proposed government project together in the context of a government decision instead of a private decision. There is one exception being the category 'spatial equality'. This category includes written motivations indicating that respondents had chosen a project because infrastructure should be improved across the region and not only in Amsterdam itself. These respondents therefore assign a value to a fair distribution amongst communities. Respondents can only express this preference in valuation studies such as PVE in which they make a portfolio choice, but not in collective WTP experiments for a single policy. Below two illustrative statements of respondents are presented.

> *"As a resident of Amsterdam, I wanted to do something for the regions outside Amsterdam with the funds I had left over."*

> *"Spread the investments across the region and across the different modes."*

Figure 3 provides for each project an overview of the motivations put forward by citizens. For reasons of readability, various small categories are excluded from Figure 3.[6] Based on this analysis, we believe that the written motivations might help explain why safety projects and cycling projects perform better in a PVE analysis than in a CBA analysis and why car projects perform better in a CBA analysis. Section 6.1 discusses potential explanations for why safety projects perform relatively good in a PVE analysis. Section 6.2 presents potential reasons for why cycling projects perform relatively good in a PVE analysis, whereas car projects perform relatively good in a CBA analysis. Note that it was relatively difficult to identify potential reasons why respondents ranked car projects lower in a PVE because participants in the PVE were only asked to provide arguments for the projects they selected and not for projects they didn't select.

[6] Examples of such categories are: 'trees should not be chopped' (mentioned 30 times) and 'this project perfectly aligns with other policies' (mentioned 5 times).

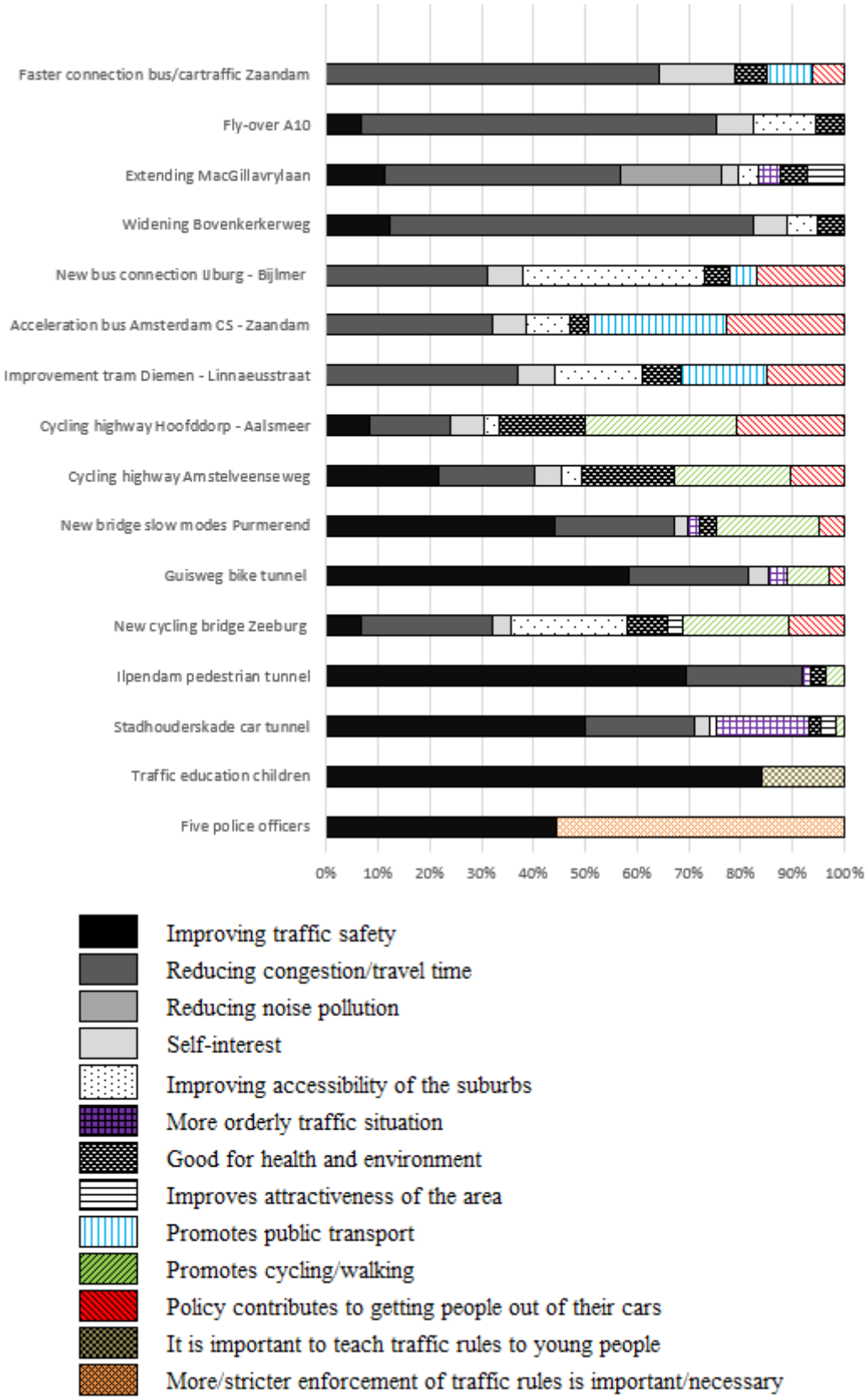

Figure 3: overview of qualitative motivations for each project

6.1 Potential reasons why safety projects perform better in a PVE than in a CBA

Improvement of traffic safety was mentioned 2,502 times to underpin the selection of a project which makes this the most frequently cited motivation for choosing projects. In the analysis of the written motivations of all the four projects which aimed to improve traffic safety (Ilpendam pedestrian tunnel, Stadhouderskade car tunnel, Traffic education, Additional policy officers) we found potential explanations for why safety projects perform better in a PVE analysis than in a CBA analysis.

6.1.1 Stadhouderskade car tunnel: individuals do not only value objective safety

Traditional CBAs generally operationalize traffic safety as the reduction in traffic deaths, serious injuries and slight injuries (Mackie et al., 2014). However, based on the statements of respondents who selected the Stadhouderskade car tunnel project we established that citizens' conceptualisations of traffic safety turn out to be broader than the reduction of injuries and deaths and this might explain why this project is ranked higher in the PVE than in the CBA. Respondents argued that they also valued the reduction of small accidents, 'near misses' and the subjective experience of safety. The fact that travellers value both 'objective safety' and 'perceived safety' is endorsed in the literature (Adam et al., in press). Moreover, many respondents who selected this project (18%) indicated that they did so because they believed it would make the traffic situation more orderly. Respondents characterized the present traffic situation as chaotic or stressful and hoped that the project would alleviate these issues thereby creating a calm, relaxed or pleasant travel experience. The importance of including these impacts into the appraisal of transport projects is emphasized in Gössling et al. (2019). Below, we present some illustrative statements provided by respondents who selected the Stadhouderskade car tunnel.

> *"I use this intersection on a frequent basis, and I think it is very dangerous. The fact that various traffic flows cross each other results in near misses on a frequent basis."*
>
> *"Many small accidents occur at this intersection which are not registered."*
>
> *"I bike across this intersection quite often and pretty much always end up ringing my bell at someone. The bike path running along the Vondelpark is narrow, and there is an enormous number of tourists in this section who don't watch where they're going. If this ceases to be an intersection, and instead becomes two roads above/below each other, then cycling will become a lot more relaxed."*
>
> *"Because this is such a chaotic scene the speed is very low and there are not a lot of traffic deaths and injuries. But still a tunnel is crucial to enhance travellers' sense of safety."*

The final statement is particularly interesting because the respondent seems to believe that an improvement of subjective safety (preventing near misses) is important even though objective safety (reduction number of injuries and traffic deaths) is not affected that much.

6.1.2 Ilpendam pedestrian tunnel: solving a specific safety issue for vulnerable travellers

The analysis of the written motivations uncovered two potential reasons why participants in the PVE assigned a particularly high value to this project. First, respondents referred to a specific characteristic of the safety issue being that pedestrians need to cross a busy road when they wish to access the bus stop. Currently, some pedestrians decide to disrespect the red traffic lights when they see that the only way to catch their bus is running through the red light. According to some respondents this results in very dangerous situations. Policy makers of the TAA were unaware of the details of this traffic safety issue prior to the completion of the PVE. Figure 4 shows the traffic situation in more detail.

Figure 4: Image retrieved from Google maps to illustrate the safety situation at Ilpendam.

Particularly the following statements clearly illustrate the specific safety issue in Ilpendam.

> *"The situation in Ilpendam is quite dangerous at present (certainly in the morning- and evening rush hours). Many pedestrians rush across even when the light is red – to catch the bus, for instance. A pedestrian tunnel would improve this dangerous situation."*

> *"Pedestrians just cross the street there right now, not at the crosswalk, not at a green light. They could get killed."*

A second reason why respondents particularly value this project is that it will improve safety for 'vulnerable travellers' such as pedestrians and children. The fact that citizens assign a higher value to traffic safety of pedestrians compared to car drivers was also found in the study of Johansson-Stenman and Martinssen (2008). Conventional CBA does not account for the two motivations that are discussed above as the standard numbers which are used for valuing a reduction of traffic deaths and traffic injuries do not differentiate between different groups (e.g. safety improvement for children or adults; pedestrians and car drivers) and between different causes of a safety issue (e.g. pedestrians negating a red light or drink and drive). Instead, participants in the PVE experiment were able to consider the local conditions of this specific

traffic safety issue when valuing the safety impacts of the pedestrian tunnel which might explain why this safety project is ranked higher in a PVE analysis than in a CBA analysis.

6.1.3 Traffic safety education for children: individuals value education as a matter of principle

Policy makers of the TAA expected a priori that respondents would only choose this project to improve traffic safety, but 71 respondents indicated that they supported such an investment because they saw it as important to impart the rules of the road to young road users. This normative belief was not included in the CBA and therefore potentially explains why it was ranked higher in the PVE.

> *"Education is a matter of principle. You should always be investing in it."*

> *"The education, guidance, and shaping of our youth contributes to a more intelligent, engaged, and respectful society in the years to come."*

6.1.4 Five additional police officers: enforcement of traffic laws is valuable in its own right

Although policy makers of the TAA proposed the project to add five additional police officers which sanction violation of traffic regulations solely to improve traffic safety, 281 respondents indicated that they had chosen it because stricter enforcement of traffic laws is desirable in and of itself. In fact, the number of respondents choosing this project for reasons of safety is lower than the number who chose it because they thought enforcement of traffic laws was valuable in its own right. This normative goal was not included in the CBA.

> *"Always good to show the traffic jackasses that they can't get away with everything. Most importantly: fines on the spot!!! No sneaky photos with payment after the fact…"*

> *"When you make rules you are obliged to enforce them."*

> *"It is about time to enforce the rules we made. Sometimes it feels that no one is obeying the rules. I know a lot of elderly people who do not cycle anymore because they are too afraid. This is madness of course."*

> *"I would not be surprized if better enforcement in traffic also results in positive impacts outside mobility because the essence is that people's inappropriate behaviour is reprimanded."*

The two final statements also suggest that participants in the PVE considered impacts which were not included in the CBA (e.g. increase in number of elderly people who are now afraid to cycle).

6.2 Potential reasons why cycling projects perform better in a PVE and car projects perform better in a CBA

This section presents potential reasons for why, on a project level, cycling projects perform relatively good in a PVE analysis, whereas car projects perform relatively good in a CBA analysis.

6.2.1 Respondents think that the TAA should promote cycling and discourage car use

Figure 3 reveals that participants in the PVE clearly include a broader set of reasons than the traditional goals of transport planning which were included in the CBAs presented in Section 4 (improving safety, reducing travel time, preventing noise pollution) in their selection of cycling projects. For instance, for the cycling highway Hoofddorp – Schiphol – Aalsmeer only 25% of the motivations can be clustered in traditional transport goals and 71% of the motivations can be attributed to broader goals of transport (18% positive impact on health and the environment; 31% promoting cycling; 22% trying to get people out of their car). The fact that more than 50% of the respondents selected the cycling highway Hoofddorp – Schiphol – Aalsmeer because they think that the TAA should promote cycling and reduce car is quite surprising as these goals were not discussed in the project description that was included in the PVE. These motivations exemplify that a group of inhabitants of the TAA has a negative attitude towards car use and these individuals positively evaluate government projects which result in a reduction of car use. Below we provide two illustrative quotes:

> *"Amsterdam is a cycling city. I believe it should always be made easier for cyclists to move throughout the city so that people are less likely to drive to where they need to be."*

> *"Perhaps if you make a fast cycling route, you'll be able to get a few people out of their cars. A few is already enough for me. If you don't do anything, absolutely nothing will happen."*

The first quote expresses that respondents clearly have a normative idea regarding the urban mobility system of Amsterdam: cycling friendly, less place for the car. The fact that such preferences are not reflected in private WTP-based valuation studies that are currently used in (Dutch) CBAs potentially explains why car projects perform better in a CBA than in a PVE. The second quote suggests that simply getting a few drivers to opt for alternate means of transportation is enough to make the project worthwhile for this respondent.

Respondents also mention other normative reasons for why they think that the TAA should invest in projects which promote cycling. First, there are respondents who feel that cycling is a part of the identity of Amsterdam itself.

> *"My choices are based on the idea that Amsterdam is a cyclists city par excellence. This idea should be further developed and therefore we should encourage cycling by expanding cycling infrastructure."*

Moreover, there are respondents who seem to believe that cycling is a desirable behaviour that should be stimulated by the government:

> *"This is an additional incentive to take the bicycle and it is also a reward for cyclists."*

> *"Those who bike deserve a comfortable route."*

One unique comment comes from a respondent who feels that cycling infrastructure should be improved in order to ensure that children learn to bike independently (to the sports club):

> *"Since more and more families today see both parents working, it is important that their children can get to the sports club by themselves during their "free" time. It is often the case that parents have to avoid making any commitments so they can get their kids to their sporting activities. If these children can safely bike along their "protected" routes, that reduces pressure on the parents and makes it less likely that someone has to "rush" home."*

Moreover, there are respondents who feel that there should now be more investment in cycling infrastructure for reasons of fairness. For instance, this could be because there has historically been much more money spent on the road network, or because proportionally less is done for cyclists and pedestrians as compared to drivers.

> *"There has been a lot of investment for drivers on the road network around Amsterdam. It's now time to consider the interests of cyclists and pedestrians."*

> *"It's nice that they're thinking about cyclists for once. This is why one would value this project more highly."*

Finally, a number of respondents emphasize the importance of high-quality cycling infrastructure because it is an inexpensive mode of transportation that is important to those who cannot afford to buy/use a car or use public transport.

> *"This is important for the cyclists who do not want to use public transport and can therefore save a bit of money by using their own bike to get around."*

6.2.2 Respondents use personal judgment to assess safety impacts

Several respondents stated that they selected a cycling project for safety reasons even though we communicated to them in the experiment that the project would not have any safety impacts. This was particularly the case for the cycling highway Amstelveenseweg project. The civil servants of the TAA that were involved in the design of the PVE were of the view that this project would not have any safety impacts. However, 23% of the motivations provided by the respondents related to safety improvements. For instance, one respondent made the following statement:

> "*I know this situation and think it is unsafe. Hence, I think that this problem should be tackled immediately.*"

Hence, we can conclude that respondents base their choices both on information that is offered to them in the experiment (formal assessment) and personal experience (informal assessment). These perceived safety impacts were not included in the CBA because they were not known by

policy makers. This might explain why this project is ranked higher in a PVE analysis than in a CBA analysis.

7. Conclusions and discussion

Participatory Value Evaluation (PVE) is a new evaluation method which establishes the desirability of government projects. It is a survey-based method in which individuals select their preferred portfolio of government projects given a constrained public budget.

The primary goal of the present study is to investigate whether CBA and PVE lead to different policy recommendations. More specifically, we conducted CBAs and a PVE for 16 transport projects (car projects, public transport projects, cycling projects and safety projects). We find that projects that focus on improving traffic safety and improvements for cyclists and pedestrians perform relatively good in the PVE, whereas car projects perform relatively good in the CBA analysis.

The second goal of the present paper is to explain how the PVE method should be positioned next to the CBA. When PVE is simply considered as a valuation method, the results can be used as input for a standard CBA analysis. In this case the outcomes of a CBA are based on the results of a PVE-based valuation instead of valuation methods that are currently used, such as hypothetical route choices. However, we show in the paper that outcomes of a PVE can directly be used to evaluate the policies that are assessed by participants in the PVE using social welfare functions. In that case PVE can be positioned as an evaluation method that replaces CBA.

The third goal of the paper is generating empirical insights into potential reasons why safety projects and cycling project perform relatively good in a PVE and car projects relatively good in a CBA. We find two kinds of potential reasons for why safety projects and cycling projects perform relatively good in a PVE: 1) conventional CBAs value impacts of government projects through observing people's past consumer choices (e.g. hypothetical route choices and behavior in the real estate market). Instead, PVE allows individuals to include normative ideas regarding their preferred *future* urban mobility system. Many of these forward-looking normative statements referred to the importance of fostering cycling and traffic safety (e.g. individuals value traffic education as a matter of principle, the normative belief of citizens that the mobility system of Amsterdam should be cycling friendly with less place for the car); 2) a conventional CBA values impacts based on standardized price tags, whereas a PVE experiment allows participants to include specific (local) characteristics of a project that are not on the radar of policy makers when valuing the impacts of a project. Especially for the safety projects participants grounded their judgments in personal experiences that policy makers were unaware of prior to the completion of the PVE. This might explain why these projects performed relatively good in the PVE-analysis. For instance, the policy to add five additional police officers is a good example of a selected project by citizens based on other motivations than the policy makers of the TAA expected a priori. Although the policy makers expected that citizens would only choose these projects to improve traffic safety, the most mentioned argument by respondents was that they thought that a stricter enforcement of the traffic laws is desirable in and of itself. Because policy makers were a priori unaware of these motivations it is highly likely that these would be omitted into the valuation of these projects in a conventional CBA.

Hence, a PVE can provide focused, and otherwise unavailable, information about citizen values, preferences and perspectives on the details of urban policy (Fung, 2003).

The most important conclusion of this study is that PVE and CBA produce different policy recommendations and we believe that we provide convincing empirical evidence which explains this divergence. Both the CBAs and the PVE that we conducted in this study were subject to several limitations. At this point, we cannot guarantee to which extent the observed differences would disappear (or increase) if these limitations would have been addressed. We believe that the results of our study warrant further research which compares the results of better-quality CBAs with a better quality PVE. Such a study could, for instance, improve the quality of the CBAs to conduct dedicated valuation studies for valuing the impacts of transport projects (e.g. dedicated studies to value of travel time savings resulting from the projects under scrutiny). Moreover, it would be interesting to compare results of a PVE with results of CBAs that are based on collective willingness to pay studies in which individuals are asked, for each project, whether they agree with a one-time tax increase for the entire community to finance the construction costs of the project. As discussed in Sections 2 and 3, we hypothesize that any divergences between policy recommendations of CBAs and PVE are much smaller when CBAs are grounded in collective WTP experiments.

We think that the quality of PVE studies can be improved by augmenting the information provided to participants. A key result of our study is that respondents selected their preferred portfolio based on impacts and considerations for which they did not receive any explicit and systematic information in the PVE. For instance, respondents argued that they selected cycling projects to promote cycling and to try to get people out of their car. Because respondents were not provided with any information, they were forced to make arbitrary judgments which primarily ended up in the project specific constants (Carson, 2012; van Wee and Börjesson, 2015). One possible solution is to transform PVE into an iterative assessment method which starts with a first round in which a relatively small focus group conducts the PVE. After this initial PVE the new motivations to select projects put forward by participants are further explored. For instance, in the case of the PVE of this paper, literature on the impact of cycling highways on mode shift from car to bicycle (e.g. Goodman et al., 2014; Heinen et al., 2015) and resulting impacts on health (e.g. de Hartog et al., 2010) would be explored. In addition, it would be interesting to interview respondents that mention new motivations. For instance, in the case of the PVE of this paper respondents who argued that they selected cycling projects to reduce car use would be interviewed to identify their underlying motivation to pursue a reduction of car use. Subsequent to these investigations, the new information would be integrated into a final PVE in which a larger group of citizens participates.

A related avenue for further research may investigate the generalizability of our results to different contexts. First, it is questionable whether CBA and PVE provide alternative policy recommendations when these methods are deployed for the assessment of transport policy options in which normative considerations regarding the future mobility system and local characteristics play a less important role (possibly the evaluation of motorway extensions in non-urban areas). In addition, normative considerations regarding the importance of city cycling might be a specific empirical result for countries with an omnipresent cycling culture such as Denmark and the Netherlands, making these results less generalizable to car-oriented countries/cities.

Apart from providing an alternative method for valuing impacts of transport projects, PVE also facilitates the participation of large groups of citizens in the design of public policies. The importance of the active involvement of citizens in the decision-making process on transport schemes to secure high-quality implementation is also recognized in the transport literature (Banister, 2008; Handy, 2008). As said before, PVE mobilizes the local knowledge of a large group of citizens to express their values and concerns regarding the specific characteristics of transport policy options in an efficient way. In essence, PVE can be conceived as a method which combines formal assessment and informal assessment (Pesch et al., 2017). The standard impacts computed by a transport model comprise the formal part of the assessment. The informal part of the assessment refers to the fact that PVE leaves room – and can be adapted to – the values and concerns of citizens that are not on the radar of policy makers and experts. Apart from mobilizing local knowledge, Table 5 confirms that participants valued the fact that participating in the PVE made them feel involved in the decision-making process. PVE may also overcome issues that result from the fact that conventional approaches to citizen participation (e.g. public hearings) generally require a substantial time commitment which many citizens would prefer to avoid. This has the potential to lead to a poor representation of the general population, insofar as those with a high motivation to participate will be those that have the most to gain by influencing decisions, but also have the free time and economic resources to do so (e.g. Irvin and Stansbury, 2004; Day, 1997). Various studies find that in conventional participation approaches white, middle aged well-educated males are overrepresented (e.g. Huitema et al., 2007; Public Agenda, 2016; Wittmayer and Rach, 2016). A key benefit of PVE is that the entry barrier for participating is relatively low. Participants generally spend 20 minutes to submit their choice(s), and the respondents can choose themselves when and where they conduct the PVE. Hence, the probability is relatively high that a more representative part of the population participates. However, our study finds that the same group of people is overrepresented as in conventional participation approaches but such concerns can be addressed by setting representative sampling targets. Moreover, a key strength of PVE is that the welfare analysis can correct for underrepresentation (see section 5.2.3) when all segments of the population are to some represented in a PVE. Further research may provide a comprehensive comparison of PVE with other approaches to citizen participation (e.g. citizen juries and focus groups).

It should be noted that the introduction of PVE does not disregard the role of experts in urban mobility planning. First of all, experts have an important role in the design of the PVE experiments. Second, citizens can delegate their choice to an expert. Moreover, in a graduation project (Darteé, 2018) citizens were asked which value should be assigned to the results of the PVE in the decision-making process. More than half of the respondents argued that the results of the PVE should not be decisive as the opinion of experts, civil servants and politicians should also count in the final decision. Hence, it can be argued that experts also have a third role on top of the design role and being an expert to whom participants can delegate.

This paper has illustrated that PVE presents a valuable evaluation methodology that is distinctive from CBA. The development of PVE fits well within the natural development of the CBA method itself where an increasing number of effects is being monetised and by now more than just time savings are quantified for transport projects. We believe that PVE is useful as an evaluation method on its own, but also as a method to challenge existing CBA practices. Indeed,

conducting a PVE requires more time and money than a standard CBA because the analyst needs to collect new data and, possibly provide a financial compensation to the participants. However, on the other hand conducting one PVE substitutes conducting multiple CBA studies. Finally, we believe that the present paper warrants further research into the advantages and disadvantages of CBA and PVE (for different types of transport policy options). Such a study may also scrutinize the (dis)advantages of PVE (and CBA) compared to other appraisal methods such as (participatory) multi-criteria analysis (Beria et al., 2012; Cornet et al., 2018; Dean et al., 2019; Hickman and Dean, 2018).

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 Transport Authority Amsterdam for financing the data collection and the design of the PVE.

8. 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 nothing. The New Press. New York.

Adam, L., Jones, T., Te Brömmelstroet, M.C.G., in press. Planning for cycling in the dispersed city: establishing a hierarchy of effectiveness of municipal cycling policies. Transportation.

Adler, M., 2012. Well-being and fair distribution: beyond Cost-Benefit Analysis, Oxford University Press, 2012.

Allen, M.T., Austin, G.W., Swaleheen, M., 2015. Measuring highway impacts on house prices using spatial regression. J. Sustain. Real Estate, 7 (1), 83-98.

Anderson, E., 1993. Value in Ethics and economics. Harvard University Press.

Ardıç, Ö., Annema, J. A., van Wee, B., 2013. Has the Dutch news media acted as a policy actor in the road pricing policy debate?. Transportation Research Part A: Policy and Practice 57, 47-63.

Asplund, D., Eliasson, J. 2016. Does uncertainty make cost-benefit analyses pointless? Transportation Research Part A 92, 195-205.

Banister, D. 2008. The sustainable mobility paradigm, Transport Policy, 15(2), 73-80.

Batley, R., Bates, J., Bliemer, M., Börjesson, M., Bourdon, J., Cabral, M.O., Chintakayala, P.K., Choudhury, C., Daly, A., Dekker, T., Drivyla, E, Fowker, T., Hess, S., Heywood, C., Johnson, D., Laird, J., Mackie, P, Parkin, J., Sanders, S., Sheldon, R., Wardman, M, Worsley, T., 2019. New appraisal values of travel time savings and reliability in Great Britain. Transportation 46 (3), pp. 583-621.

Bergson, Burk, A. 1938. A reformulation of certain aspects of welfare economics. Quarterly Journal of Economics, 52 (2), 310-334.

Beria, P., Maltese, I., Mariotti, I. 2012. Multicriteria versus Cost Benefit Analysis: a comparative perspective in the assessment of sustainable mobility, Eur. Transp. Res. Rev., 4, pp. 137 - 152

Beukers, E., 2015. Shaking up the Cost Benefit Analysis process: Issues and directions for improvement when assessing integrated spatial transport plans through a cost benefit analysis. PhD thesis.

Beukers, E., Bertolini, L., Te Brömmelstroet, M.C.G., 2012. Why cost-benefit analysis is perceived as a problematic tool for assessment of transport plans: a process perspective. Transportation Research Part A 46 (1), 68-78.

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

Börjesson, M., Eliasson, J., 2014. Experiences from the Swedish value of time study. Transportation Research Part A: Policy and Practice 59, 144-158.

Carson, R.T., 2012. Contingent valuation: a practical alternative when prices aren't available. Journal of Economic Perspectives 26(4), 27-42.

Cornet, Y., Merrill Jones, B., Barfod, M., Hickman, R. 2018. Giving current and future generations a real voice: a practical method for constructing sustainability viewpoints in transport appraisal. European Journal of Transport and Infrastructure Research, 18 (3).

Cuppen, E., Pesch, U., Taanman, M., Remmerswaal, S., 2016. Normative diversity, conflict and transitions: shale gas in the Netherlands. Technol. Forecast. Soc. Change.

Daniels, R.F., Hensher, D.A., 2000. Valuation of Environmental Impacts of Transport Projects. The challenge of self-interest proximity. Journal of Transport Economics and Policy 34 (2), 189-214.

Darteé, K., 2018. Assessing the application of the Participatory Value Evaluation method for urban storm water management in a The Hague case study

Day, D. 1997. Citizen participation in the planning process: an essentially contested concept?. Journal of Planning Literature, 11 (3), 421-434.

Dean, M., Hickman, R., Chen, C.L., 2019. Testing the application of participatory MCA: the case of the South Fylde Line. Transport Policy (73), pp. 62-70.

De Bruijn, H. & Veeneman, W., 2009. Decision-making for light rail, Transportation Research Part A: Policy and Practice, 43(4), 349-359.

De Hartog, J. J., Boogaard, H., Nijland, H., Hoek, G., 2010. Do the health benefits of cycling outweigh the risks? Environ. Health Perspect.118(8),1109–1116.

Decisio, 2014. MKBA Ring Utrecht. Commissioned by Ministry of Infrastructure and the Environment, The Hague.

Dekker, T., Koster, P.R., Mouter, N., 2019. The economics of participatory value evaluation experiments. Working paper Tinbergen Institute.

Ferreira, A., Beukers, E., Te Brömmelstroet, M., 2012. Accessibility is gold, mobility is not: a proposal for the improvement of Dutch transport-related cost-benefit analysis. Environment and Planning B 39, 683-697.

Fung, A., 2003. Survey article: recipes for public spheres: eight institutional design choices and their consequences. The Journal of Political Philosophy 11 (3), 338-367

Goodman A, Sahlqvist S, Ogilvie D. 2014. New walking and cycling routes and increased physical activity: One- and 2-year findings from the UK iConnect study. Am J Public Health.104, 38–46.

Gössling, S., Choi, A., Dekker, K. and Metzler, D., 2019. The social cost of automobility, cycling and walking in the European Union. Ecological Economics (158), pp. 65-74.

Hajer, M.A., Pelzer, P. 2018. 2050—An Energetic Odyssey: Understanding 'Techniques of Futuring' in the transition towards renewable energy. Energy Research & Social Science 44, 222 – 231.

Handy, S.L., 2008. Regional transportation planning in the US: an examination of changes in technical aspects of the planning process in response to changing goals. Transport Policy, 15, 2, 113 - 126.

Hauer, E. 1994. Can one estimate the value of life or, is it better to be dead than stuck in traffic? Transportation Research Part A 28 (2), pp. 109–118.

Heinen, E., Panter, J., Mackett, R., Ogilvie, D. 2015. Changes in mode of travel to work: a natural experimental study of new transport infrastructure. International Journal of Behavioral Nutrition and Physical Activity 12, p. 81.

Hickman, R., Dean, M., 2018. Incomplete cost – incomplete benefit analysis in transport appraisal, Transport Reviews, 38:6, 689-709

Huitema, D., Van de Kerkhof, M., Pesch, U., 2007. The nature of the beast: are citizens' juries deliberative or pluralist? Policy Sci. 40, 287–311.

Ivehammar, P., 2008. Valuing in actual travel time environmental encroachment caused by transport infrastructure. Transportation Research Part D 13, 455-461.

Ivehammar, P., 2014. Valuing environmental quality in actual travel time savings: The Haningeleden road project in Stockholm. Research in Transportation Economics 48, 349-356.

Irvin, R., Stansbury, J., 2004. "Citizen participation in decision-making: is it worth the effort?". Public Administration Review, 64 (1), 55-65.

Jara-Díaz, S.R, 2007. Transport economic theory. Elsevier Science, Amsterdam.

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

Kouwenhoven, M., G.C. de Jong, P. Koster, V.A.C. van den Berg, E.T. Verhoef, J. Bates and P.M.J. Warffemius. 2014. New values of time and reliability in passenger transport in The Netherlands. Research in Transportation Economics 47, 37-49.

Lancaster, K.J., 1966. A new approach to consumer theory. Journal of Polical Economy 74 (2), 132-157.

Lloyd-Smith, P., 2018. A new approach to calculating welfare measures in Kuhn-Tucker demand models. Journal of Choice Modelling 26, pp. 19-27.

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

Manaugh, K., Badami, M., & El-Geneidy, A., 2015. Integrating social equity into urban transportation planning: A critical evaluation of equity objectives and measures in transportation plans in North America. Transport Policy(37), 167-176.

Mouter, N., 2017a. Dutch politicians' attitudes towards Cost-Benefit Analysis. Transport Policy 54, 1-10.

Mouter, N., 2017b. Attitudes of Dutch politicians towards Cost-Benefit Analysis. Transport Policy 54, pp. 1-10.

Mouter, N., 2018. A critical assessment of discounting policies for transport Cost-Benefit Analysis in five European practices. European Journal of Transport and Infrastructure Research 18 (4), pp. 1-7.

Mouter, N., Annema, J.A., Van Wee, B., 2013a. Attitudes towards the role of cost-benefit analysis in the decision-making process for spatial-infrastructure projects: a Dutch case study. Transportation Research Part A 58, 1-18.

Mouter, N., Annema, J.A., Van Wee, B., 2013b. Ranking the substantive problems in the Dutch Cost-Benefit Analysis practice. Transportation Research Part A 49, 241-255.

Mouter, N., Chorus, C.G., 2016. Value of Time: a citizen perspective. Transportation Research Part A 91, 317-329.

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, pp. 53 – 63.

Mouter, N., Koster, P.R., Dekker, T., 2019. An introduction to Participatory Value Evaluation. Working paper Tinbergen Institute 19-024/V

Mouter, N., Ojeda Cabral, M., Dekker, T., van Cranenburgh, S. in press. The value of travel time, noise pollution, recreation and biodiversity: a social choice valuation perspective. Research in Transportation Economics.

Nicolaisen, M. S., Olesen, M. & Olesen, K., 2017. Vision vs. Evaluation - Case Studies of Light Rail Planning in Denmark, European Journal of Spatial Development, 65.

Nyborg, K., 1998. Some Norwegian politicians' use of cost-benefit analysis. Public Choice. 95, 381-401.

Nyborg, K., 2014. Project Evaluation with democratic decision-making: What does cost–benefit analysis really measure? Ecological Economics 106, 124–131.

Nurmi, V., & Ahtiainen, H., 2018. Distributional Weights in Environmental Valuation and Cost-benefit Analysis: Theory and Practice. Ecological Economics 150, 217-228.

Pesch, U., Correljé, A., Cuppen, E., Taebi B., 2017. Energy justice and controversies: Formal and informal assessment in energy projects. Energy Policy 109, 825-834.

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

Public Agenda (2016) "Public Spending, By The People. Participatory Budgeting in the United States and Canada in 2014-15". Available online at July 2nd 2019: http://www.publicagenda.org/files/PublicSpendingByThePeople_PublicAgenda_2016.pdf

Rijkswaterstaat 2018, Steunpunt Economische Expertise, kengetallen. www.rwseconomie.nl.

Rip, A., 1986. Controversies as Informal Technology Assessment. Sci. Commun. 8, 349–371.

Romijn, G., Renes, G., 2013. General Guidance for Cost-Benefit Analysis.. CPB en PBL, Den Haag.

Samuelson, P.A. 1947. Foundations of economic analysis. Cambridge: Harvard University Press.

Sagoff, M., 1988. The economy of the earth. Cambridge University press. Cambridge.

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

Sen, A.K. 2000. The discipline of cost-benefit analysis. The journal of Legal Studies 29 (2), 931-952.

Seo, K., Golub, A., Kuby, M., 2014. Combined Impacts of Highways and Light Rail Transit on Residential Property Values: A Spatial Hedonic Price Model for Phoenix, Arizona. Journal of Transport Geography 41, 53–62.

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

Sunstein, C.R., 2005. Cost-Benefit Analysis and the Environment. Ethics 115 (2), pp. 351-385.

Te Brömmelstroet, Marco, Nikolaeva, Anna, Glaser, Meredith, Nicolaisen, Morten Skou, Chan, Carmen, 2017. Travelling together alone and alone together: mobility and potential exposure to diversity. Appl. Mob. 2 (1), 1−15.

Van Wee, B., Börjesson, M., 2015. How to make CBA more suitable for evaluating cycling policies. Transport Policy 44, 117 – 124.

Volberda (2020) How to cover for distribution of citizens' views within the Participatory Value Evaluation? A case study to identify preference profiles among citizens towards the allocation of public budget to spatial-infrastructure projects. Graduation thesis TU Delft.

Vonk Noordegraaf, D., Annema, J.A., van Wee, G.P. 2014. Policy implementation lessons from six road pricing cases. Transportation Research Part A 59, 172-191.

Vigren, A., Ljungberg, A. 2018. Public Transport Authorities' use of Cost-Benefit Analysis in practice. Research in Transportation Economics 69, 560 – 567.

Weber, R.P., 1990. Basic Content Analysis, second ed., Newbury Park, CA.

Weimer, D. L., 2017. Behavioral economics for cost-benefit analysis: Benefit validity when sovereign consumers seem to make mistakes. Cambridge University Press.

Wittmayer, J.M., S. Rach, 2016. "Participatory Budgeting in the Indische Buurt"; Chapter 5 of TRANSIT Case Study Report Participatory Budgeting.