“Correlations in Project Combinations:”

“An Empirical Investigation into Synergies in Participatory Value Evaluation”

Abstract. *Participatory Value Evaluation (PVE) has been developed in recent years as an ex-ante valuation method of government projects. PVE can be seen as an alternative for more traditional CBA, because it approaches it respondents as citizens rather than as consumers. Additional to the fact that the method is therefore arguably more appropriate to be used by governments in some public domains, there are other argued advantages to PVE. One of those is the idea that PVE can capture so-called synergies, implying that combinations of government projects can have a higher (or lower) effect on social welfare than merely the sum of the two projects separately. However, there is not a large body of empirical evidence to back up the theory behind PVE. In this paper, data from a PVE experiment on 16 government projects concerning urban mobility in and around Amsterdam is used to estimate these synergies. Standard logit is used as a benchmark, while mixed logit is used as a way to overcome the limitations of standard logit. Mixed logit is deemed to be somewhat more appropriate for analyzing PVE data, though not overwhelmingly so, since substantial evidence for the existence of synergies in the data cannot be found.*

Key concepts: *Participatory Value Evaluation, valuation methods, mixed logit, synergies, urban mobility.*

Geert van Bemmelen – 2579936

Master Thesis 2021

MSc STREEM

Supervised by Paul Koster

Table of Contents

1. Introduction

Whenever decisions on the spending of public budget on government projects or policies are to be made, whether they relate to infrastructure, public health or national security, governments will act in a similar manner to the one that all individuals follow according to economic theory: attempt to maximize the value subject to a budget constraint. Government bodies will often base their decisions on experts' analyses to achieve this. How these analyses are performed is thus an important topic of debate in and of itself. The public nature of the government's role means that simply looking at financial costs and benefits is not sufficient. To be able to make an adequately weighted decision, it is important that governments also look at benefits for citizens that are not merely financial (although these benefits are often translated to a monetary value). In other words, a process of valuation of a potential government project will take place. A traditional and very common way of doing so is the *cost-benefit analysis* (CBA). Although the method is so widespread that it is even mandatory in some countries in some sectors to perform an ex-ante CBA, it is not an undisputed method (e.g., Choy, 2018; Mouter et al., 2019; Nyborg, 2014; Spash, 1997). Measurements of welfare effects in CBA are based on market interactions or consumer surveys on private willingness-to-pay (WTP). An alternative and relatively recent method for the process of valuation is *Participatory Value Evaluation* (PVE). In PVE, welfare changes are not measured by market interactions or WTP, but rather by asking individuals how they would like government budget to be spent. The method is an extension of *willingness to allocate public budget* (WTAPB), with possibilities to allow respondents to lower of heighten the government budget through one-time tax increases or decreases. One of the argued advantages of PVE is its capability to capture respondents' preferences for combinations of projects, or so-called *synergies* (Bahamonde-Birke & Mouter, 2019). However, since PVE is a relatively young valuation method, there is no large body of literature and evidence yet on the effectiveness and reliability of the method. Inspired by the framework developed by Bahamonde-Birke & Mouter (2019), this paper will attempt to add to that body of literature by researching whether such synergies can be discovered – and possibly estimated – in data from a PVE experiment conducted by Mouter, Koster & Dekker (2021) in cooperation with the Transport Authority Amsterdam (TAA). A standard logit approach and a mixed logit specification will be used to achieve those results.

The next section will provide an overview of the relevant literature on the topic. It will go into the generals of PVE, compare it to traditional methods and explain its relevance to the field of urban mobility. The third section explains how the data was gathered by Mouter et al. in their 2020 paper, as well as the data preparations and extensions performed for this paper. Section 4 will introduce the methodology on which the quantitative part of this paper is based. Section 5 will show the results. Finally, section 6 contains conclusions, discussion and suggestions for further research.

2. Participatory Value Evaluation (PVE)

2.1 *PVE explained*

This paper revolves around a relatively new valuation method that was developed as a counterpart to – or an expansion of – traditional CBA. Currently, CBA is one of the most widely used methods for ex ante valuation of (government) projects. According to documentation on how to use this method, provided by the Dutch government, CBA can be used to compare economic benefits and positive welfare effects with their negative counterparts and factor in uncertainty to maximize the precision of ex-ante evaluations of government projects (Romijn & Renes, 2013). The government document also mentions the reliance on willingness to pay (WTP) for the positive consequences of a government decision. This means that individuals are asked how much they are willing to pay to experience the positive effects of a government policy, be it in the form of a good or in the form of a service. Willingness to pay is measured using stated preference data, by asking individuals how much money they are willing to spend on something or how much money they would have to receive to be indifferent between the money and the alternative situation in question. All benefits, including the collective WTP, are monetized. If they exceed the total costs, which also include monetized indirect costs, it suggests net positive effects and thus an argument for funding a project or implementing a policy. This is what is called the benefit-cost ratio (BCR). However, the most determinative factor often is the net present value (NPV). The government guideline on CBA in the Netherlands prefers NPV over BCR because the latter is more vulnerable to manipulation (Romijn & Renes, 2013).

Where CBA relies on WTP, PVE relies on a different form of stated preference data. In a PVE experiment, respondents are asked to choose a portfolio of government projects within a specified public budget. Information on economic costs and benefits and their societal costs and benefits, as well as potential consequences for the respondent's private budget is provided for every government project (Mouter et al., 2019). Although motivation, honesty and thoughtfulness are important for all stated preference experiments, an especially critical aspect of PVE is that respondents are convinced of the possibility that their choices have consequences. That is why they are told that the results of the evaluation process could be used to advise policy makers in their decisions. In that sense, respondents are asked to really approach the spending of government budget as if they are the policy makers who actually get to distribute the funds. This is an important informational element of PVE, in which there is an active effort to put the respondents in the mindset of a citizen rather than a consumer. The relevance of this lies in the fact that individuals possibly make different choices or can have different preferences in their role as citizen than they have in their role as consumer. The existence of this *consumer-citizen duality* has been shown to exist in the case of food safety standards in the US (Alphonce, Alfnes & Sharma, 2014) and decisions on road safety versus travel time in the Netherlands (Mouter, Van Cranenburgh & Van Wee, 2017), amongst others. Section 2.2 will go into this phenomenon in more detail.

Furthermore, Mouter et al. (2019) note that PVE is an extension of the *willingness to allocate public budget* model in the sense that, in a *flexible budget* PVE, it is also possible for respondents to choose to not exhaust the given budget. In that case, any remaining budget can either be saved for next year or it can flow back to citizens in the form of tax relief. A distinction can thus be made between a *fixed budget* PVE and a *flexible budget* PVE. In the former, respondents are not able to increase or decrease the budget (through a one-off increase or relief in household taxes), whereas this is possible in the latter (Mouter et al, 2020). Lastly, in PVE experiments, respondents can be given the opportunity to delegate their decision. Although it is often preferred for respondents to create their own portfolio, they can delegate the selection of projects to an expert in the field if they think that is more appropriate.

2.2 *The case for PVE*

As mentioned previously, the main aim of PVE experiments is to imitate more closely the setting for which respondents are asked to state their preference. In other words, PVE is useful for government decision making because it was designed for it (Mouter et al., 2020). PVE puts respondents in the decision maker's seat in order to estimate how respondents value the (potential) implementation of government projects. Why this is important can best be explained by the *consumer-citizen duality*. The application of this concept to the field mobility has been explored thoroughly in two papers by Mouter, Cranenburgh & van Wee (2017; 2018). In their 2017 paper, these authors conduct a series of stated choice experiments in which respondents are asked to make choices on hypothetical routes. The respondents are asked to make their choices either in the context of their role as a consumer or as a citizen, with the latter meaning that they are asked to give a recommendation to the government. The authors identify a discrepancy between how individuals value the tradeoff between reductions in travel time and additional traffic casualties. Their results suggest that individuals value safety relatively higher than reduction in travel time in their role as citizen (Mouter et al., 2017). Considering that CBAs (or the *value of a statistical life* and *value of travel time*, estimated through CBA to put it more correctly) are currently used in the decision-making process for similar projects, it is relevant that changing to a citizens' perspective might mean that different (sets of) projects would generate a larger amount of welfare. This is not to say that CBAs have no merit in ex ante valuation whatsoever, however it would make sense to lean more towards the citizen perspective when considering *government projects* that influence citizens.

In a 2018 paper by the same authors, possible explanations for the duality between consumer and citizen are given. A distinction is made between *normative* and *cognitive* explanations. The former category entails individuals' opinions on how the government should behave. In this case, the conviction is often that one of the core tasks of the government is to pursue and accommodate driver safety. Additionally, a normative explanation lies in the fact that individuals are often inclined to be more risk averse when another individual's health or safety is involved in their decision, due to the existence of social norms. The cognitive explanations for the consumer-citizen duality in the context of mobility include the inability to fully comprehend extremely small (changes in) probabilities, acceptance of higher risk levels due to overestimation of one's own capabilities and the influence of 'controllability' on the accepted level of risk (Mouter et al, 2018). PVE explicitly acknowledges these dynamics and actively encourages respondents to think normatively, since they are asked to think on behalf of the government. In that sense, PVE is a more political process than traditional CBA. However, one could argue that this is not problematic when performing ex ante evaluations of government projects, which is political by definition. The cognitive aspects of the consumer-citizen duality are somewhat more difficult to deal with. The inability of individuals to comprehend small changes in probabilities (of traffic mortality, for example) is not miraculously solved by doing a PVE experiment but translating these small probability changes to the larger scale will make them easier to comprehend. In a PVE experiment, respondents evaluate additional traffic injuries or deaths rather than small changes in probability of experiencing a traffic accident themselves. This also implies that respondents are likely to be less risk-averse, since they are considering other individuals' safety and thus overestimation of their own skills and controllability will also play a smaller role.

Furthermore, Bergstrom et al. (2004) report that there can be differences in how individuals value policy (implications) or funding of public goods when the funds are taking from current government budget versus when funds are raised through an additional tax scheme. The results of their study on ground water in two states in the USA indicate that individuals' marginal value for the protection of ground water quality is higher when it is funded using a tax reallocation. Additionally, the authors find that such effects are likely to be independent of individuals' nominal income. This last finding is relevant in the context of the consumer-citizen duality. The opposing concepts that illustrate this comprehensively are those of one-person-one-vote (OPOV) and one-euro-one-vote (OEOV) (Mouter et al, 2019; Nyborg, 2014; van Wee, 2012). The former is arguably more appropriate in the context of government decisions, whereas the latter is more fit for markets and consumer-based contexts. The findings from Bergstrom et al. (2004) imply that methods like WTAPB – or PVE, though it had not been invented yet in the early 2000's – are a closer approximation of OPOV. CBA, on the other hand, can be seen as a method that puts OEOV into practice (Mouter et al, 2019; Nyborg, 2014).

Finally, one of the strong points of PVE that is mentioned regularly is the fact that it can capture how individuals value combinations of projects. When one attempts to specify the welfare changes of government projects using CBA and building on WTP, there is often a focus on one single project. Even when multiple projects are considered, their net costs and benefits are usually compiled linearly. In theory, implementing two projects with a positive net present value would result in a welfare change amounting to the sum of the two net present values. However, it is likely that this linear assumption does not always hold. It is not hard to imagine an example where two projects involving the same area might have net present values independently, but the cumulative effect of both projects might lead to an even larger positive welfare change if the two projects complement each other, or a smaller positive welfare change when the projects contradict or hinder each other. The way in which PVE experiments are conducted forces respondents not only to value projects separately, but also to think about the value of projects in relation to each other. They aim to maximize their value within the given budget, which is likely to influence their project choices within their portfolio. Not only the specifics of projects can play a role in this, but also how the respondents perceive and value fairness. The latter could be expressed by spreading the chosen projects over a larger geographical area or by choosing projects that improve mobility for different types of road users. Although this dynamic seems intuitive, there is not a lot of empirical evidence to confirm this hypothesis yet.

2.3 *PVE in urban mobility*

In their 2020 paper, Mouter, Koster & Dekker provide arguments for using PVE in the field of urban mobility and planning. The authors do so by comparing it to the 'status quo' method of CBA, mainly by evaluating how PVE improves on CBA's shortcomings. Firstly, they argue that CBA is relatively effective for the evaluation of short-term outcomes, both in terms of costs and benefits, whereas planning of urban mobility is an inherently long-term process with long-term goals. Both the transportation system as a whole, as well as additions and alterations to it are part of a larger process that works towards the future mobility system (Mouter et al., 2020). The authors argue that part of this process transcends issues of congestion, safety and travel time. There are also non-traditional effects that are difficult to quantify within a traditional CBA – and in general – as their value cannot simply be estimated by respondents or deduced from market settings. Examples of these effects, provided by Mouter et al. (2020), include quality of urban and public spaces, sustainability and social inclusion.

Secondly, urban mobility planning is often at least partly based on some standardized values. Two examples are the *value of (travel) time* and the *value of a statistical life* (both were mentioned previously). These values are often estimated centrally by government institutions through CBA or CBA-like processes. This is part of what Pesch, Correljé, Cuppen & Taebi (2017) call *formal assessment*, which includes all steps, standards, legislation and policy goals that have been formalized institutionally for such processes. The counterpart, *informal assessment*, involves all other values that are not captured in the formal assessment but that are important to other actors within the process. In the case of mobility projects, these actors are often local citizens or local companies that experience the consequences of the (lack of) implementation of a project. Such informal effects are often substantial in mobility projects, but they are not captured in CBA.

In theory, PVE improves upon the shortcomings identified in the previous paragraphs. Although PVE does not necessarily make it easier or even possible to quantify some of the abovementioned factors, it does provide respondents with an opportunity to express their opinions on them and grant them some value. By selecting a portfolio of projects, assuming that it was done carefully and with attention, respondents are in fact ranking projects in a broader context. This means that respondents can take into account not only financial consequences of (combinations of) projects, but also the more personal and informal aspects of projects. PVE even leaves room for respondents to express their long-term, idealistic views on what the future should look like. Again, these views are not necessarily expressed explicitly, but the setup of a PVE experiment helps respondents to take those considerations into account. Translating these values into monetary terms, as is necessary for CBA, can be quite a difficult task for respondents.

On the other hand, PVE is not unlike other valuation methods in the sense that it has some shortcomings. One of the most significant is the reliance on stated preference data. More generally speaking, one of the perceived issues with stated preference data is the fact that what individuals *state* as their preference may differ from how they would behave in real life (Kroes & Sheldon, 1988). In cases where respondents are asked to participate in a hypothetical market, this discrepancy can lead to real life outcomes that differ significantly from the estimates based on stated preference data. Although PVE overcomes this specific issue due to the fact that respondents are not approached as consumers. One could even argue that PVE has turned this disadvantage into a positive aspect of PVE: the method actively asks respondents to make their decision based on their ideals, rather than their own day-to-day behavior (Mouter et al, 2021). However, other disadvantages of stated preference data are more difficult to bypass. A main problem is the tradeoff between complexity and accessibility. Policy makers' questions for which PVE attempts to formulate a response are complex by definition, especially when one considers PVE leaves room for long term considerations and large portfolios. Weighing off the advantages and disadvantages, be it in monetary terms or not, is often difficult enough for policy makers to do so, let alone respondents who are asked to fill out a relatively short survey. Although this complexity might make the conclusions of PVE experiments more useful because they are (theoretically) more realistic, it is also important to make sure that PVE is an accessible method of valuation. For a citizen-based approach, it is important that the group of respondents is diverse and representative, both in terms of ethnicity and age, as well as education level and income (to name but a few aspects). Mouter et al (2021) argue that it is relatively simple to control for such factors, but it is still only possible when a sufficient number of respondents from each category fill out the survey. Even when this criterium is fulfilled, there can still be issues with the content of their response to the survey. The complexity might not only make it so that respondents do not wish take part in the survey to begin with, but it could also cause them to make arbitrary decisions in order to finish the survey quickly. Filtering these kinds of responses from the data can be very difficult, while leaving them in can cause troubling biases. To combat this, respondents are often given the option to delegate their decision to an expert in the field. Although this can be useful, it is often not encouraged (by means of a lower compensation when choosing this option) in PVE experiments. Table 1 summarizes all previously mentioned advantages and disadvantages of the PVE method. Note that the contents of table 1 do not apply exclusively to PVE, but their relevance in this context has been argued in the previous paragraphs.

Table 1. *Advantages and disadvantages of Participatory Value Evaluation.*

Advantages	- *Citizen-based approach in citizen-based situations* - *Allows for idealistic and long-term perspective* - *Possibility to capture implicit impacts*
Disadvantages	- *Relatively complex* - *Reliance on stated preference data, implying:* ▪ *Risk of discrepancy between hypothetical and real-life* ▪ *Risk of incorrect responses due to complexity* ▪ *Risk of arbitrary responses*

3. Data

3.1 *PVE and CBA in the TAA*

The data used in this paper was gathered in a large PVE experiment conducted together with the Transport Authority Amsterdam (TAA) on 16 projects in the area. This section will mostly focus on the data that is relevant for the analysis and on how that data was gathered. For specifics, it is advised to consult the 2021 paper by Mouter, Koster & Dekker. In their paper, the authors attempt to quantify to what extent PVE experiments yield different results compared to traditional CBA analyses. If so, these different results would subsequently imply differing policy recommendations. In order to compare the two methods, the authors first conduct a CBA analysis for all 16 projects. The impacts considered in their analyses include (estimates of) execution costs, the number of travelers affected by the project, minutes of traveling time saved due to implementation of the projects, changes in the numbers of traffic deaths and severe injuries, households affected by noise pollution and the number of trees cut (Mouter et al., 2021). Since this paper focuses mainly on the existence of synergies, the qualitative assessment and details will not be discussed in too much detail. Instead, for the table with descriptive information on all 16 projects I will refer the reader to Mouter, Koster & Dekker (2021), where the information on project details can be found in their table 1. Where possible, the information on project attributes is translated to monetary impacts following guidelines from the Dutch government. This enables the authors to compute a benefit-cost ratio (BCR) for all projects, after which 13 out of the 16 projects are found to have a positive BCR.

Their PVE experiment is carried out in four waves. Half of those waves are flexible budget PVEs. The distinction between flexible and fixed PVE results will not be mentioned explicitly in the coming sections, for simplicity's sake. All the impacts mentioned in the previous paragraph are also presented to respondents in the PVE experiment, which is why the authors call them *explicit* impacts. Additionally, the estimated project specific parameters signify the values respondents assign to projects outside of the explicitly mentioned impacts. Furthermore, respondents are asked to comment on their portfolio choice. These written motivations were used to perform a content analysis, which the authors want to utilize to gain more insight into why the results from the PVE (hypothetically) differ from traditional CBA analysis. Also, these motivations can help with interpreting and specifying the random taste variation.

3.2 *Data preparation and descriptive statistics*

In order to be able to examine synergies between projects in PVE, some additional variables need to be generated. More specifically, there must be one variable for every possible combination of two projects. Because there are 16 projects in the data and we are looking at combinations rather than permutations (i.e., the order of the two projects does not matter), this means that there are 120 combination variables. For every combination variable, it is checked whether both projects have been selected. This is done for every individual in the dataset. The code that has been used to generate these variables is made available upon request.

The dataset now contains an additional 120 project combination variables for all 2227 individual responses to the PVE experiment. Figure 1 contains a correlation matrix of all project choices. A first glance suggests that only a few projects show notable correlation. Figure 2 shows the distribution of the choice for project combinations, for which the boxplot suggests no real outliers. The two projects with the highest correlation as shown in figure 1 is that of project 2 and project 11, which was chosen most often – 514 times. This suggests that there could be a positive synergy between these two projects. Since there can theoretically also be negative synergies, it is really the absolute values of correlations we should be considering. The two projects with a correlation level that is furthest from zero in the other direction are projects 1 and 15. This combination was selected 17 times. The combination that was chosen least often is that of project 1 and project 6, which was chosen only twice. The mean count of all project combinations is equal to approximately 178, with a standard deviation of 117. All combinations were chosen at least once. In general, comparing popularity of project combinations to the data on project choices in general gives rather logical results: popular projects, meaning projects that were chosen relatively often regardless of combination effects, are more prevalent in the top project combinations. The same – or rather the opposite – holds for unpopular projects. Figures 3 and 4 show the 25 most and least frequently selected project combinations. Note that these are project combination choices, which can thus be part of a larger portfolio.

In total, respondents selected 1199 unique portfolios which includes the empty set. The distribution of this variable is shown in the boxplot in figure 5. The mean of portfolio choices is approximately 1.9, with a standard deviation of 2.4. About 85% of portfolios were selected only once or twice. The figure shows that there are some outliers, which are relatively more popular than other portfolios. The most popular portfolio, however, is the empty set. The other portfolios that are among the 10 most frequently selected can be found in table 2. The fact that the top portfolio was selected by 35 respondents (roughly

1.6%) implies that preference for portfolios has a large spread. The spread gives an indication of how complex this valuation method can get – and perhaps the corresponding policy recommendations even more so. Another interesting detail is the fact that projects 14 and 15 show negative correlation in figure 1, while the combination is found in four out of the ten most popular portfolios.

Figure 1. Correlation matrix of project choices.

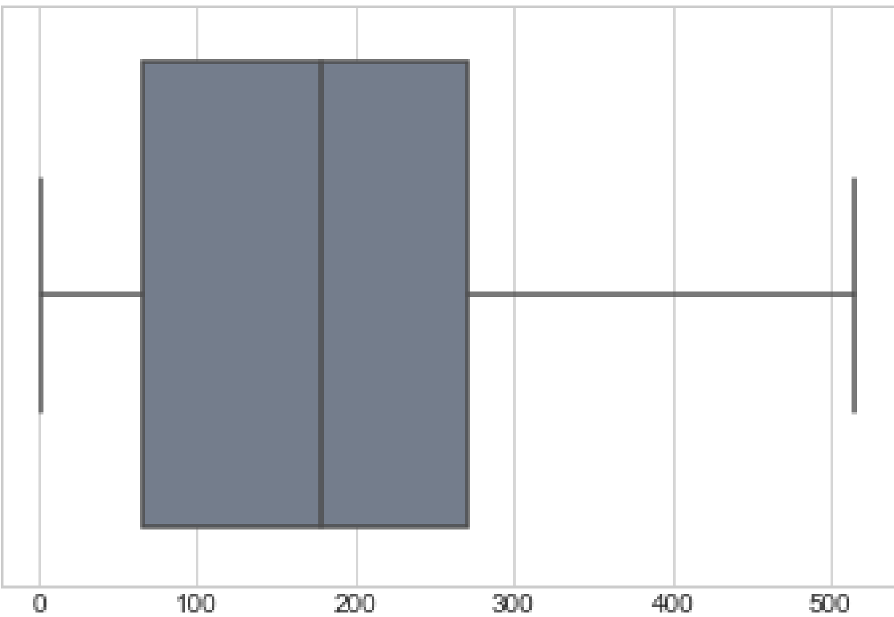

Figure 2. Distribution of project combination choice frequency.

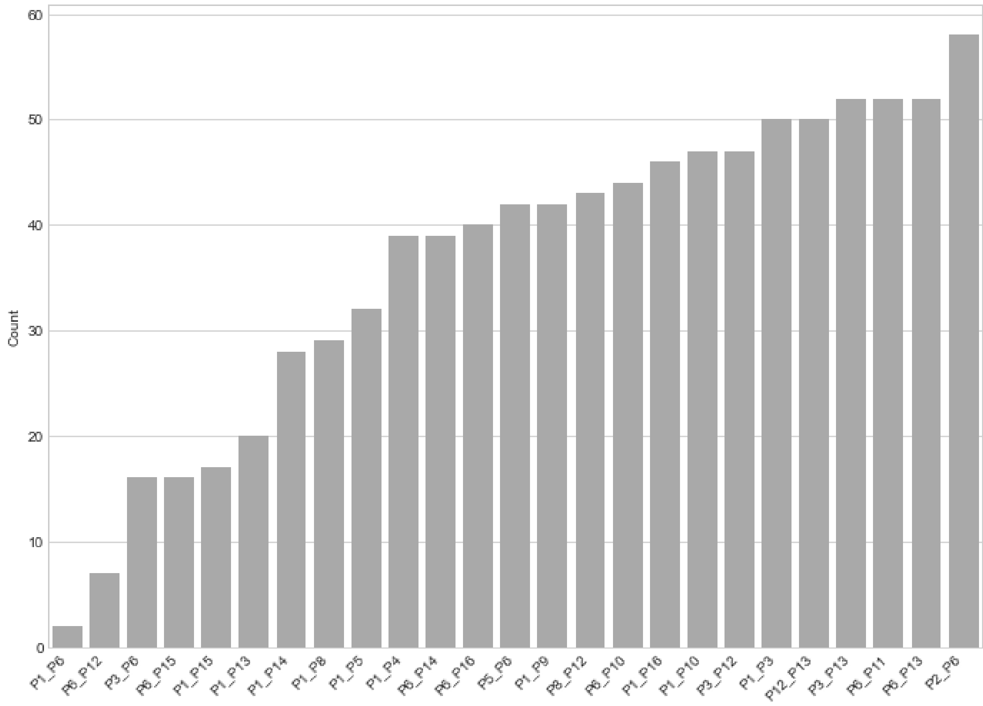

Figure 3. Low 25 of project combination choices.

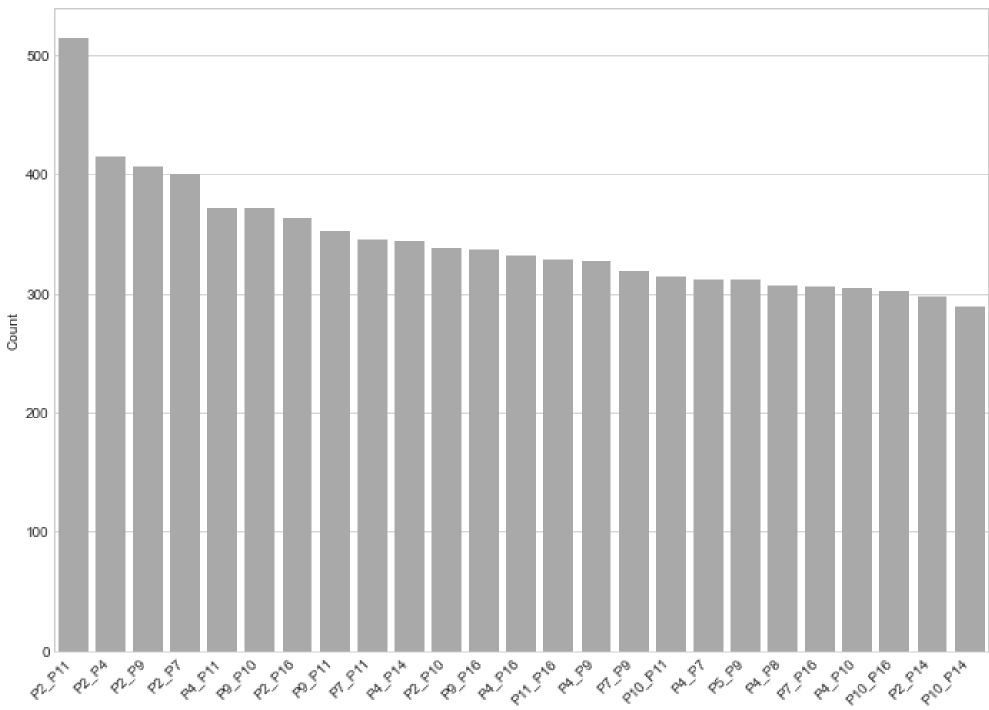

Figure 4. Top 25 of project combination choices.

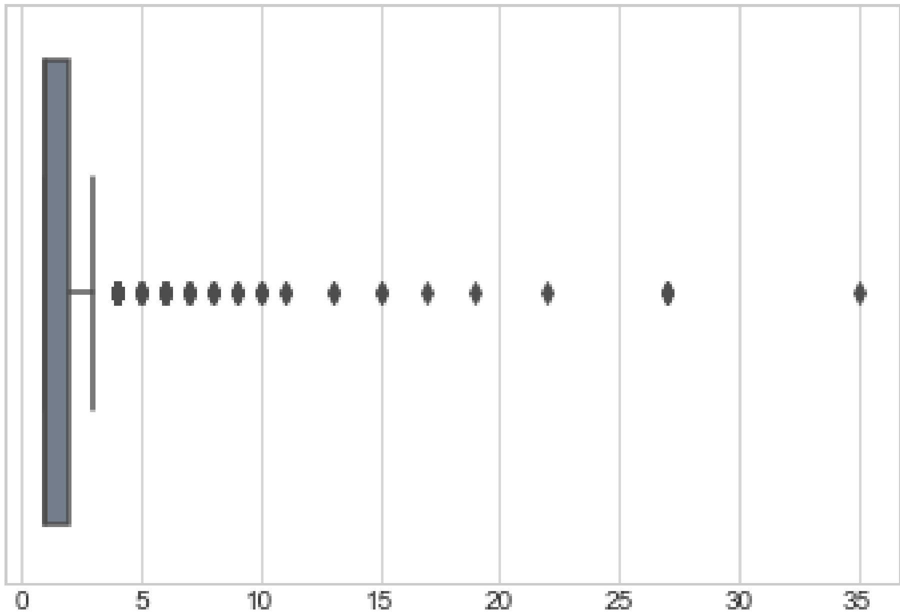

Figure 5. Distribution of portfolio choice frequency.

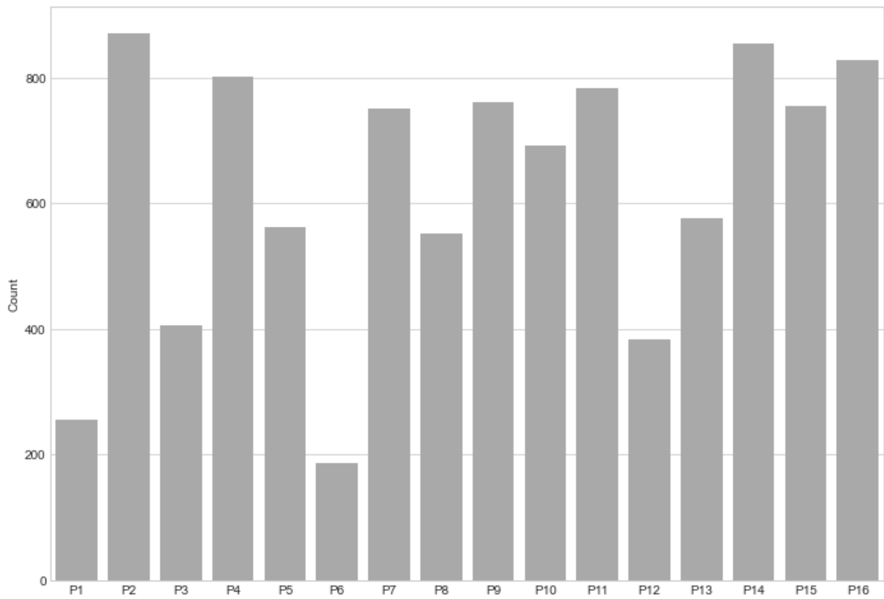

Figure 6. Frequency of single project choices.

Table 2. *Top 10 of most frequently selected portfolios.*

Portfolio	*Count*	*Percentage*
1. No projects	35	1.572 %
2. Project 1, Project 7, Project 12	27	1.212 %
3. Project, 14, Project 15	27	1.212 %
4. Project 4, Project 14, Project 15	27	1.212 %
5. Project 15, Project 16	22	0.988 %
6. Project 14, Project 15, Project 16	19	0.853 %
7. Project 13, Project 14, Project 16	17	0.763 %
8. Project 4, Project 8, Project 13, Project 14	15	0.674 %
9. Project 2, Project 3, Project 11, Project 15	15	0.674 %
10. Project 10, Project 14, Project 15	13	0.584 %

4. Methodology

4.1 *Synergies in PVE*

In their 2019 paper, Bahamonde-Birke and Mouter construct – and attempt to verify – a framework that measures the positive and negative interaction effects that PVE is theoretically able to capture. Their hypothesis is that there can be complementarities or substitution effects between projects. This implies that a portfolio of projects can yield substantially higher or lower social utility that merely the sum of its parts. It is possible to study these dynamics by generating all possible portfolios and examining choice probabilities, using a multinomial logit model at the portfolio level. The reasoning by Bahamonde-Birke and Mouter (2019) is as follows. The decision maker is assumed to make a selection out of the available projects such that the social utility of the portfolio is maximized, given the budget constraint. A portfolio's social utility (SU) is considered to be a function of all projects k included in portfolio p. In other words, the social utility of a given portfolio p chosen by individual i consists of the sum of social utilities for all projects k in the portfolio, plus the sum of all parameters α_{km} capturing expected utility changes due to *both* project k and project m being included in the portfolio, plus the total budget minus the sum of all project costs times parameter α_B. That final parameter thus measures the marginal utility of expenditures not being allocated to implement projects.

Although Bahamonde-Birke & Mouter (2019) verify their model with a synthetic dataset and it is therefore a contender for detecting synergies, their approach requires substantial computational power. Their synthetic data contains only four possible projects and a budget that leads to a set of 15 unique and feasible portfolios. The set of possible portfolios in the TAA contains all combinations of projects that have a total combined cost that is lower than or equal to the available budget, which enlarges the set of unique and feasible portfolios substantially. In the case of a flexible budget PVE, it is even possible to expand the budget which increases the set of possible portfolios even more. The number of unique and chosen portfolios in this dataset already equals 1199 options. The number of possible portfolios is even larger, since it is thinkable that there are feasible portfolio alternatives that have not been selected in the TAA experiment. To get to an estimate of these interaction parameters more easily, it is thus interesting to investigate whether other methods would also be able to detect synergies in the data. The two models that will be used to do so will be discussed in section 4.3, but first it is important to elaborate on the underlying utility theory.

4.2 Random utility framework for discrete choice models

Similar to the methodology of Mouter et al. (2021), this paper will also study the PVE data using behavioral choice models. Firstly, to establish some sort of baseline, the data will be utilized to run a set of standard logit regressions. The results from the application of binary logit will help in understanding the effect of the observed characteristics on choice probabilities, or what is sometimes called *taste variation* (Train, 2003). The random utility framework dictates that utility can be split up into two elements, of which the terminology can be somewhat ambiguous. In his 2009 book, Train distinguishes taste variation from *random taste variation*. In other instances, a distinction is made between random taste variation and *deterministic utility*. For ease of telling the two apart, I will make use of the latter distinction in the remainder of this paper. The deterministic element of utility contains, as mentioned previously, the effects that the observed characteristics have on the total utility. If there would be only deterministic utility, knowing all observed characteristics would enable one to determine individuals' choices (for projects, for example) flawlessly (Train, 2003). However, random taste variation hampers this process to some extent. Random taste variation can be made up of many different aspects, but what they all have in common is the fact that they are not caused by the known characteristics of the model specification. Mathematically and paraphrasing the specification provided by Train (2003), utility as specified by the random utility framework can be expressed as in equation 1, which states that the utility U is made up of deterministic utility V (depending on a vector of project characteristics, x) and random taste variation ε of individual i for project k. In line with McFadden (1973), the distribution of ε_{ik} is assumed to be Extreme Value Type I (EV-1).

$$U_{k_i} = V(x_k) + \varepsilon_{ik} \tag{1}$$

Since both U_{k_i} and ε_{ik} cannot be directly observed, estimations have to be performed resulting in choice probabilities. For this research, the left-hand side of the equation that forms the basis for these estimations consists of the probability of individual i selecting project k into their portfolio. The right-hand side contains the project's characteristics. In this model specification, it is assumed that an individual only chooses to include a project in their portfolio if the utility that results from such a choice exceeds the utility that one experiences when the project is not included. This condition is expressed in equation 2.

$$\Pr(\mathrm{k_i} = 1) = \Pr\left(\mathrm{U}_{\mathrm{k_i}=1} > U_{k_i=0}\right) \tag{2}$$

More specifically for PVE, the utility as a result from including a certain project should exceed the utility of the additional government budget equal to the project's cost (Bahamonde-Birke & Mouter, 2019). When the PVE is of the flexible type, it is not the additional government budget but the additional private consumption of which the utility should be exceeded.

4.3 Standard logit

Now, to return to the data at hand. One of the simpler discrete choice models is standard binary logit, partly because its results are relatively easy to interpret when translated to marginal effects. Additionally, standard logit performs quite well in estimating the deterministic portion of utility (Train, 2003). Performing some mathematical manipulations that combine equation 1, equation 2 and the EV-1 assumption leads to the logit specification that can be found in equation 3.

$$Pr(k_i = 1) = \left(1 + e^{-\beta x_k}\right)^{-1} \tag{3}$$

This specification is now prepared to be used in statistical software to estimate the deterministic utility and thus the effects of the projects' characteristics on the choice probabilities in the PVE experiment conducted in the TAA. This model will serve as a control for the somewhat more suitable yet more complex mixed logit that will be discussed below.

The need for mixed logit is due to the fact that the usage of standard logit is not without limitations. Most notably, Train (2003) mentions three main limitations. Firstly, though it works well for estimating the deterministic utility, standard logit does not excel in estimating the random taste variation found in the random utility framework. Secondly, logit is only able to capture substitution effects between alternatives when these effects behave according to the model's specification. This specification is decided by the researcher, based on expectations grounded in theory. If substitution occurs different to the specification's prediction, logit is not able to represent those accurately. The *independence of irrelevant alternatives* (IIA) assumption implies that the odds ratio of two options in the choice set is independent of whether or not a third option is included. Standard logit is built around IIA, while it is not always a reasonable assumption to make. In the context of PVE, it is likely that this assumption is unrealistic. Mixed logit, on the other hand, does not rely on this assumption. The third limitation formulated by Train (2003), concerns issues that arise when applying logit to data containing individuals taking the same decision at different points in time. Since the data in this paper is transformed in such a way that the selection of projects in portfolios simulates decisions over time, this limitation forms an issue. Standard logit is not able to deal with unobserved characteristics that are correlated over time. The unobserved characteristics of individuals' choices over the 16 projects (or 120 project combinations) are assumed to be highly correlated.

The first limitation can be accounted for in standard logit to some extent by making use of so-called *alternative-specific constants* (Train, 2003). These constants can be estimated by including dummy variables for every project in the regression. The coefficients for these project dummies are then interpreted as the alternative-specific constants, or *project-specific parameters* as they will be called in this context. Their function is similar to that of a 'regular' constant in a linear model, in the sense that they correct the model in such a way that the mean of the predicted values is equal to the mean of the observed values in the data. It the context of random utility, these constants estimate the random taste variation as they contain everything that is not part of deterministic utility as captured by the coefficients for project attributes. Because of the other limitations mentioned in the previous paragraph, however, logit is not able to estimate these constants without bias.

4.4 Mixed logit

The model specification that is able to overcome the issues explained in the previous paragraph is mixed logit (Train, 2003). Mixed logit has been utilized as a valuation method in several domains (e.g., Brownstone & Train, 1998; Campbell, 2007; Revelt & Train, 1998; Small, Winston & Yan, 2005). Mixed logit builds upon standard logit by taking the estimated probabilities, as defined in equation 3, at different levels of coefficients. Additionally, mixed logit lets go of the IIA assumption. This allows for capturing various substitution patterns as has been described previously. Whereas the basic logit model had general coefficients for the entire population, mixed logit allows the coefficients to vary over the population, according to a density function of those coefficients. The coefficient levels are weighted based on the density function of the parameters in order to portray a weighted average of the regular logit function. Abstractly speaking, the choice probability is then given by the integral of the logit probability L_{k_i} at values of β, accounting for the distribution of the $\beta's$, expressed by a density function f which is in turn defined by parameters θ (Train, 2003). This is what is expressed in equation 4.

$$Pr(k_i = 1) = \int L_{k_i}(\beta)\, f(\beta|\theta)\, d\beta \qquad (4)$$

In contrast with logit, this does not result in closed form choice probabilities. These thus have to be estimated by statistical software using simulation (Train, 2003). Based on the random utility framework we assume utility to be linear (like in equation 1), but since the specific preferences of individuals are not directly observed we specify those to be random variables with a density function equal to $f(\beta|\theta)$. The distribution of the β's has to be specified before fitting the model. A normal or log-normal distribution is most common, where the latter is mostly applicable to situations where coefficients are assumed to have the same sign for all individuals (Train, 2003). Technically speaking, the error term ε_{ik} in equation 1 can now be split up into two parts: one part containing the random taste variation and one part that captures all the remaining error.

To account for individual heterogeneity in the mixed logit specification, it is possible to run a mixed logit model for panel data. This is similar to the specification in Brownstone & Train (1998), where the utility is divided in three: one part capturing all non-random utility due to observed factors, one part capturing random utility that is correlated between alternatives but differs between individuals and one part capturing utility that can be heterogeneous over both individuals and alternatives. It is possible to do so because of the relaxed IIA assumption (Hensher & Greene, 2003).

To apply this panel data mixed logit, however, the PVE data first has to be transformed to resemble panel data. That means that for each individual, there are 16 rows of observations in the case of single project choices and 120 rows of observations in the case of project combination choices. These choices are thus treated in the same way as different time instances would be in regular panel data. Individual-specific preferences are assumed to be equal over the choice occasions. Since mixed logit is a choice model, the alternatives are specified for each of these rows of observations. The choice is binary, so the alternatives are to select (1) and to not select (0) a project to be included in the portfolio.

5. Results

5.1 *Standard logit*

The first empirical application of the data consists of a binary logit model. This model acts as a baseline in order to verify the model specification before applying it to the project combinations. Table 3 contains the results from this primary model. The table looks similar to table 6 in Mouter et al. (2021) though interpretation differs. The first six estimates reflect the influence of the projects' attributes on choice probabilities, also described as explicit impacts in previous paragraphs. For this specification, the attributes are estimated over all sixteen projects since it is assumed that the decrease in utility – and in turn the decrease in choice probability – for a project as a result of the implication of an additional traffic death (for example) is the same for all projects. Due to the logit specification of the model, the coefficients are not to be interpreted directly. However, it is possible to look at the signs and significance of the parameters. These results align with intuition and are also in line with the results of Mouter et al. (2021). Additional costs for a project are associated with a negative effect on choice probability. The same holds for additional traffic deaths and additional severe traffic injuries. The coefficients for additional households affected by noise pollution and additional trees cut show the same sign, but they do not have statistical significance. The only coefficient which shows a positive relationship with choice probability is the one capturing additional travel time saved (in millions of minutes), but this coefficient is again not statistically significant. The results of this regression on PVE data can be put in perspective of traditional CBA analyses by looking at the relative importance individuals place on traffic deaths prevented versus severe traffic injuries prevented. Traditional CBA claims that the prevention of 1 traffic death provides the same utility (or holds the same value) as the prevention of roughly 8.67 severe traffic injuries (Mouter et al., 2021). From these results, the relative value equals approximately 8.5 (0.2134 / 0.0251) which is in the same range. It is worth noting that the results from Mouter et al. (2021) show that PVE and CBA differ in the fact that individuals show a relatively higher preference for safety compared with additional travel time saved in PVE than in CBA. Such a statement cannot be made based upon the results in table 3.

In section 4.2, it was already specified that the project-specific parameters are not to be interpreted directly, since this logit specification is not able to estimate them without bias. In the case of table 3, the project-specific parameters are such that they correct the model to make perfect predictions on average. Furthermore, it is important to note that the project-specific parameters are merely saying something about choices on the project level. Since PVE respondents are creating a portfolio subject to a budget constraint (either public or private budget) and not just selecting projects based on their features alone, these parameters have not a lot of explanatory value quantitatively speaking. This also links into the previously mentioned inability to capture substitution effects due to the IIA assumption. The possibility remains to look at the signs of the parameters since, theoretically speaking, a positive value for a project-specific parameter implies an inherent preference, or random taste, for the corresponding project. The results in table 3 suggest that such a preference only exists for project 14 and project 15, though the coefficients show no statistical significance. These projects are among some of the most popular projects as shown in figure 6 and, as noted earlier, they also occur relatively often in the most frequently selected portfolios as shown in table 2. The fact that the coefficients for all other projects have a negative sign would imply that there is an inherent aversion to these projects on average. However, almost all individuals include at least one project in their portfolio, so this aversion seems implausible.

Applying the same model to the data containing project combinations results in the coefficients that can be found in table 4. The major change in these results, apart from the additional project-specific parameters, is the fact that the project attributes are added to the regression in the form of the sum of the impacts of the two projects separately. By doing so, they equal the total effect of the project combinations. These parameters show resemblance to the previously estimated parameters, though some are statistically more significant in this estimation. The relative value that individuals ascribe to the prevention of an additional traffic death in relation to an additional severe traffic injury is somewhat higher in this estimation. The prevention of one additional traffic death has the same value as the prevention of roughly 9.4 severe traffic injuries.

As mentioned in section 3.2 containing descriptive statistics, there are some project combinations of which the correlation suggests the existence of synergy. There is no real indication for these synergies in the project-specific parameters in table 4. For example, the combination of project 2 and 11 showed the highest correlation but it has no stand-out parameter, equaling -1.2058. The combination of project 1 and 15 with a low correlation has an even lower parameter of -1.8835, but it is also not stand-out. The signs of both parameters suggest that there is a preference for not including these projects. The only project combination that shows to have some inherent preference is that of project 14 and 15, just like in the previous estimation. However, this coefficient is again not statistically significant. It is very likely that the counterintuitive results from the logit specification for both single project and project combination choices are caused by the logit specification's shortcomings in estimating random taste variation.

5.2 *Mixed logit*

The mixed logit specification requires a few adjustments in comparison with the previous model. First of all, a distinction has to be made between random and fixed coefficients. The former type includes coefficients for which heterogeneity between individuals' preferences is assumed. The latter, on the other hand, contains variables for which all individuals are assumed to have the same preference. In order to keep the runtime of the statistical software somewhat reasonable, some coefficients are assumed to be non-random. That is to say, they are assumed to be equal over all individuals. The previous two specifications indicate that the value ascribed to the prevention of traffic deaths and severe traffic injuries is somewhat stable, both in relation to the model specifications and in relation to CBA results. Therefore, these preferences are assumed to be homogeneous in this dataset, together with the values for additional households affected by noise pollution and the number of trees chopped due to their low explanatory power in the previous specifications. The parameters for the remaining two project attributes are assumed to be heterogeneous. The distribution of the coefficient for project cost is assumed to be log-normal, which is appropriate for reasons discussed in section 4.2, since it is expected to have a negative correlation with choice probability for all individuals. The coefficient for travel time saved is assumed to have a standard normal distribution.

The results from the mixed logit specification for single project choices can be found in table 5. All coefficient signs for the project attributes are in line with expectations and with previous specifications. The only project attribute that still has both economically and statistically significant explanatory power is the project cost, which is log-normally distributed around -3.4449 with a standard deviation lower than 0.0001. The size of the coefficient for project cost has increased notably compared to standard logit, namely with a factor 100. This increase is to be expected – though not necessarily in these proportions – since the mixed logit specification allows the non-stochastic variables to capture more of what is left to the error terms in standard logit (Brownstone & Train, 1998; Revelt & Train, 1999). The other results for project attributes imply that the cost of a project is the main explanatory variable for project selection, apart from the random taste for projects. Even though the relative value for the prevention of additional traffic deaths versus the prevention of additional severe traffic injuries is still

Table 3. *Logit regression on single project choice.*

Log-likelihood: -20,276.173		
	Coefficients	**t statistic**
Attribute parameters		
Project cost (in 1,000,000 euros)	-0.0257***	(-5.43)
Time saved (in 1,000,000 minutes)	0.4083	(0.55)
Additional traffic deaths	-0.2134*	(-2.45)
Additional traffic injuries	-0.0251*	(-1.98)
Additional households affected by noise pollution	-0.0009	(-0.71)
Additional trees cut	-0.0020	(-1.31)
Project-specific parameters		
Project 1	-0.7759**	(-2.65)
Project 2	-0.4115***	(-7.88)
Project 3	-0.2048	(-0.57)
Project 4	-0.4747**	(-2.70)
Project 5	-0.7708***	(-5.96)
Project 6	-1.1294***	(-4.55)
Project 7	-0.5498***	(-10.74)
Project 8	-0.7309***	(-8.21)
Project 9	-0.4575***	(-7.74)
Project 10	-0.5349***	(-4.85)
Project 11	-0.5347***	(-9.97)
Project 12	-0.6016**	(-3.05)
Project 13	-0.1962	(-1.15)
Project 14	0.3381	(1.66)
Project 15	0.2976	(1.11)
Project 16	-0.3920*	(-2.40)

* p<0.05, ** p<0.01, *** p<0.001

Table 4. *Logit regression on project combination choice.*

Log-likelihood: -69,092.021	**Coefficients**	**t statistic**
Attribute parameters		
Project combination cost (in 1,000,000 euros)	-0.0321***	(-14.46)
Time saved (in 1,000,000 minutes)	0.3871	(1.16)
Additional traffic deaths	-0.1611***	(-4.32)
Additional traffic injuries	-0.0171**	(-2.92)
Additional households affected by noise pollution	-0.0012*	(-1.99)
Additional trees cut	-0.0012	(-1.72)
Project-specific parameters		
Project 1 & Project 2	-1.5801***	(-8.74)
Project 1 & Project 3	-0.6788*	(-2.16)
Project 1 & Project 4	-2.2788***	(-9.67)
Project 1 & Project 5	-2.2733***	(-9.06)
Project 1 & Project 6	-3.8224***	(-5.13)
Project 1 & Project 7	-1.0628***	(-6.26)
Project 1 & Project 8	-2.2439***	(-9.01)
Project 1 & Project 9	-2.0959***	(-9.73)
Project 1 & Project 10	-1.9781***	(-9.24)
Project 1 & Project 11	-1.7511***	(-9.21)
Project 1 & Project 12	-0.5248*	(-2.16)
Project 1 & Project 13	-2.0270***	(-6.69)
Project 1 & Project 14	-1.6429***	(-5.70)
Project 1 & Project 15	-1.8835***	(-5.55)
Project 1 & Project 16	-1.8895***	(-8.18)
Project 2 & Project 3	-0.8575***	(-4.54)
Project 2 & Project 4	-1.2683***	(-12.54)
Project 2 & Project 5	-1.5433***	(-16.91)
Project 2 & Project 6	-1.9731***	(-11.07)
Project 2 & Project 7	-1.2917***	(-21.71)
Project 2 & Project 8	-1.4574***	(-18.43)
Project 2 & Project 9	-1.1880***	(-19.21)
Project 2 & Project 10	-1.4066***	(-17.80)
Project 2 & Project 11	-1.0258***	(-18.02)
Project 2 & Project 12	-1.2708***	(-10.01)
Project 2 & Project 13	-1.0661***	(-9.64)
Project 2 & Project 14	-0.6942***	(-5.80)
Project 2 & Project 15	-0.6184***	(-4.20)

Table 4. *(continued)*

Project-specific parameters	Coefficients	t statistic
Project 2 & Project 16	-1.2083***	(-12.38)
Project 3 & Project 4	-1.2689***	(-6.09)
Project 3 & Project 5	-1.2065***	(-5.12)
Project 3 & Project 6	-1.8535***	(-5.29)
Project 3 & Project 7	-1.0501***	(-5.46)
Project 3 & Project 8	-1.6353***	(-7.12)
Project 3 & Project 9	-1.2274***	(-6.09)
Project 3 & Project 10	-1.3371***	(-5.89)
Project 3 & Project 11	-0.9394***	(-4.91)
Project 3 & Project 12	-1.1113***	(-4.14)
Project 3 & Project 13	-1.1764***	(-4.53)
Project 3 & Project 14	-0.4007	(-1.60)
Project 3 & Project 15	-0.6013*	(-2.19)
Project 3 & Project 16	-1.1611***	(-5.27)
Project 4 & Project 5	-1.4446***	(-11.07)
Project 4 & Project 6	-1.6446***	(-8.47)
Project 4 & Project 7	-1.5122***	(-14.76)
Project 4 & Project 8	-1.2149***	(-10.61)
Project 4 & Project 9	-1.3689***	(-13.07)
Project 4 & Project 10	-1.4524***	(-12.60)
Project 4 & Project 11	-1.3507***	(-13.19)
Project 4 & Project 12	-1.6692***	(-9.96)
Project 4 & Project 13	-0.8554***	(-5.95)
Project 4 & Project 14	-0.4437**	(-2.92)
Project 4 & Project 15	-0.8547***	(-4.85)
Project 4 & Project 16	-1.2359***	(-9.76)
Project 5 & Project 6	-2.0200***	(-9.50)
Project 5 & Project 7	-1.6233***	(-17.54)
Project 5 & Project 8	-1.6428***	(-14.60)
Project 5 & Project 9	-1.2169***	(-13.44)
Project 5 & Project 10	-1.3076***	(-11.62)
Project 5 & Project 11	-1.7081***	(-17.91)
Project 5 & Project 12	-1.9193***	(-10.97)
Project 5 & Project 13	-1.4829***	(-9.90)
Project 5 & Project 14	-0.8990***	(-6.05)
Project 5 & Project 15	-0.8620***	(-5.11)
Project 5 & Project 16	-1.3473***	(-11.70)
Project 6 & Project 7	-1.8077***	(-10.19)
Project 6 & Project 8	-1.5080***	(-7.75)

Table 4. *(continued)*

Project-specific parameters	**Coefficients**	**t statistic**
Project 6 & Project 9	-1.7538***	(-9.57)
Project 6 & Project 10	-2.0688***	(-10.26)
Project 6 & Project 11	-2.0347***	(-10.97)
Project 6 & Project 12	-2.9386***	(-6.88)
Project 6 & Project 13	-1.0786***	(-4.62)
Project 6 & Project 14	-1.3287***	(-5.17)
Project 6 & Project 15	-1.9679***	(-5.83)
Project 6 & Project 16	-2.0572***	(-9.05)
Project 7 & Project 8	-1.3775***	(-17.13)
Project 7 & Project 9	-1.3792***	(-20.48)
Project 7 & Project 10	-1.5782***	(-18.98)
Project 7 & Project 11	-1.4217***	(-22.44)
Project 7 & Project 12	-1.1078***	(-8.66)
Project 7 & Project 13	-1.2627***	(-10.65)
Project 7 & Project 14	-0.9165***	(-7.34)
Project 7 & Project 15	-0.8762***	(-5.72)
Project 7 & Project 16	-1.3128***	(-13.26)
Project 8 & Project 9	-1.3878***	(-16.07)
Project 8 & Project 10	-1.4442***	(-14.95)
Project 8 & Project 11	-1.6098***	(-18.98)
Project 8 & Project 12	-2.2139***	(-11.26)
Project 8 & Project 13	-0.8188***	(-6.10)
Project 8 & Project 14	-0.6362***	(-4.41)
Project 8 & Project 15	-1.2154***	(-6.60)
Project 8 & Project 16	-1.3067***	(-11.17)
Project 9 & Project 10	-1.1073***	(-13.95)
Project 9 & Project 11	-1.3095***	(-19.96)
Project 9 & Project 12	-1.3883***	(-9.79)
Project 9 & Project 13	-0.9697***	(-8.13)
Project 9 & Project 14	-0.6791***	(-5.31)
Project 9 & Project 15	-0.5549***	(-3.58)
Project 9 & Project 16	-1.1116***	(-11.00)
Project 10 & Project 11	-1.4422***	(-17.85)
Project 10 & Project 12	-1.6233***	(-10.64)
Project 10 & Project 13	-1.0514***	(-8.32)
Project 10 & Project 14	-0.5385***	(-4.11)
Project 10 & Project 15	-0.6939***	(-4.36)
Project 10 & Project 16	-1.2379***	(-11.33)
Project 11 & Project 12	-1.1101***	(-8.76)

Table 4. *(continued)*

Project-specific parameters	Coefficients	t statistic
Project 11 & Project 13	-1.1091***	(-9.70)
Project 11 & Project 14	-0.8586***	(-6.90)
Project 11 & Project 15	-0.6801***	(-4.51)
Project 11 & Project 16	-1.2778***	(-12.74)
Project 12 & Project 13	-1.4689***	(-6.77)
Project 12 & Project 14	-1.1547***	(-5.29)
Project 12 & Project 15	-0.8775***	(-3.64)
Project 12 & Project 16	-1.4491***	(-8.18)
Project 13 & Project 14	-0.0156	(-0.09)
Project 13 & Project 15	-0.5899**	(-2.72)
Project 13 & Project 16	-1.1304***	(-7.34)
Project 14 & Project 15	0.1929	(0.86)
Project 14 & Project 16	-0.4775**	(-2.87)
Project 15 & Project 16	-0.1875	(-0.95)

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

in the same region – the ratio is now roughly equal to 9.6 – both coefficients have lost their statistical significance. Considering the project-specific parameters, mixed logit requires a baseline project which changes interpretation somewhat. However, taking that fact into account, mixed logit results for project-specific parameters show no substantial differences to those in standard logit. Most notably, the project-specific parameters for project 14 and 15 have gained statistical significance while remaining to have a positive sign. Other project-specific parameters, and especially for 'unpopular' projects (projects 1, 3, 6, 12 and 13 according to figure 4) are now not significantly different from zero. The exception from this is project 6, which has the lowest coefficient in this specification, implying it is the project with the least random taste for it. This makes sense, due to the fact that it is the least selected project in this PVE experiment.

The mixed logit specification for the project combinations needs some working towards, since the full specification can require substantial computational time. Firstly, a random sample can be taken from the original dataset. To keep the sampled data representative of the original dataset, the sample is proportionally taken from the four waves of the original PVE experiment. This could even be done in several iterations, in order to decrease the standard deviation and biases that occur due to sampling. Secondly, it is possible to make a selection of project-specific parameters for project combinations before running the regression. This gets somewhat more tricky, as qualitative considerations can bias

***Table 5**. Mixed logit regression on single project choice.*

Log simulated likelihood: *-20,279.073*			
	Coefficients	**Std. error**	**t statistic**
Random parameters			
Project cost (in 1,000,000 euros)	-3.4449***		(-27.72)
Standard deviation (log-normal distr.)	0.0000	0.00	
Travel time saved (in 1,000,000 minutes)	0.4504		(0.61)
Standard deviation (normal distr.)	0.0000	0.00	
Non-random parameters			
Additional traffic deaths	-0.0835		(-1.23)
Additional traffic injuries	-0.0087		(-0.82)
Additional households affected by noise pollution	-0.0010		(-0.73)
Additional trees cut	-0.0019		(-1.29)
Project-specific parameters			
Project 1	-0.4716		(-1.79)
Project 2	-0.3716***		(-7.50)
Project 3	0.0333		(0.10)
Project 4	-0.4221*		(-2.42)
Project 5	-0.7421***		(-5.76)
Project 6	-0.8223***		(-3.87)
Project 7	-0.5185***		(-10.48)
Project 8	-0.6378***		(-7.96)
Project 9	-0.4089***		(-7.36)
Project 10	-0.5015***		(-4.58)
Project 11	-0.4875***		(-9.77)
Project 12	-0.3199*		(-2.02)
Project 13	0.0125		(0.09)
Project 14	0.6959***		(5.01)
Project 15	0.8172***		(5.18)
Project 16 (baseline)	N/A		(.)

* p<0.05, ** p<0.01, *** p<0.001

Table 6. *Mixed logit regression without project-specific parameters for project combinations.*

	Full dataset		Sample data	
	Coefficient	t statistic	Coefficient	t statistic
Project cost (in 1,000,000 euros)	-3.0622***	(-298.84)	-3.1107***	(-95.07)
Standard deviation (log-normal distr.)	0.2252***	(23.16)	0.2041***	(6.81)
Travel time saved (in 1,000,000 minutes)	-9.6612***	(-31.21)	-10.4637***	(-9.95)
Standard deviation (normal distr.)	11.2864***	(38.35)	12.7182***	(13.62)
Additional traffic deaths	0.1733***	(7.60)	-0.1253	(-1.82)
Additional traffic injuries	-0.0643***	(-19.21)	-0.0312	(-2.94)
Additional households affected by noise pollution	0.0010***	(4.84)	-0.0000	(-0.05)
Additional trees cut	-0.0043***	(-15.98)	-0.0048***	(-5.68)

* p<0.05, ** p<0.01, *** p<0.001

Table 7*. Mixed logit regression on sample data with selection of project combinations.*

Log simulated likelihood: *-9,236.006*			
	Coefficients	**Std. error**	**t statistic**
Random parameters			
Project cost (in 1,000,000 euros)	-2.9569***		(-92.90)
Standard deviation (log-normal distr.)	0.3046***	13.54	
Non-random parameters			
Travel time saved (in 1,000,000 minutes)	-3.2310***		(-8.65)
Additional traffic deaths	0.2904*		(2.02)
Additional traffic injuries	-0.0272		(-1.38)
Additional households affected by noise pollution	0.0027***		(4.51)
Additional trees cut	-0.0051***		(-6.07)
Project-specific parameters			
Expected positive			
Project 3 & Project 14	2.2272***		(8.62)
Project 3 & Project 15	1.6539***		(4.78)
Project 4 & Project 14	1.9740***		(11.89)
Project 9 & Project 15	0.4361		(1.95)
Project 10 & Project 14	1.1049***		(6.27)
Project 13 & Project 14	1.8567***		(9.30)
Project 13 & Project 15	-0.3089		(-0.60)
Project 14 & Project 15	2.0278***		(8.69)
Project 14 & Project 16	1.1924***		(5.97)
Expected negative			
Project 1 & Project 5	-11.689		(-1.63)
Project 1 & Project 6	-178.437		(-0.00)
Project 1 & Project 8	-1.4058*		(-1.96)
Project 1 & Project 9	-1.7517*		(-2.45)
Project 1 & Project 12	1.5264***		(5.04)
Project 5 & Project 6	-0.7739		(-1.52)
Project 6 & Project 10	-0.3628		(-0.98)
Project 6 & Project 11	-1.2650**		(-2.78)
Project 6 & Project 12	-168.470		(-0.01)
Project 8 & Project 12	-1.5242**		(-3.00)

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

the results to some extent. For now, I will make a selection based on figures 3 and 4 and the results from table 4. The 10 project combinations with the highest coefficient in table 4 were selected as well as the 10 with the lowest coefficient, given that the coefficient was highly significant. As a baseline, table 6 shows what the results from the mixed logit specification would look like without project-specific parameters for both the random sample and the entire dataset. Excluding all project-specific parameters for project combinations leads to heavily biased results. The coefficient for minutes of travel time saved in millions is negative and statistically significant, which is counterintuitive. However, the two models do show resemblance. Other descriptive statistics of the sample data, like mean income, age and area of residence, are also similar to those of the dataset as a whole. Table 7 includes the results for the mixed logit model on sampled data with a selection of project-specific parameters for project combinations. For simplicity, only the attribute parameter for project cost has been assumed to be random with a log-normal distribution. The results are in line with intuition and hypothesis to some extent. The cost attribute has both economically and statistically significance, though its estimated standard deviation has increased somewhat. The coefficient for travel time saved, however, has become negative implying that marginal travel time saved is associated with a lower choice probability. This does not seem logical. It is possible that this is due to the fact that the majority of project-specific parameters for project combinations have not been included. The coefficient is thus likely to contain bias. Considering the project-specific parameters that were included, most of the project combinations show to have the expected sign. A number of project combinations containing project 14 are substantial in size and statistically significant, which is similar to the result in the standard logit regression. Project 14 is also one of the most popular projects as described in figure 6. It is likely that the selection process for coefficients affects the results, although it also implicates that project 14 is chosen more often in the same combination than project 2 is. The latter being the most popular project according to figure 6. That means that project 14 has a higher potential of synergy, whereas project 2 is selected more often but into a more diverse set of portfolios.

6. Conclusion and discussion

6.1 *Interpretation and implications*

The novelty of the PVE method complicates the process of formulating a rich context around the research results in this paper. Nonetheless, it is possible to formulate some general remarks about the outcomes of the different model specifications on the data from the PVE experiment in the TAA. Firstly, the standard logit specification and its assumptions result in biased estimations of individuals' preferences for projects, both for the project's attributes and the random taste for projects. However, this is not surprising. On the other hand, the results do show sensible signs for project's attributes' coefficients though, implying that it is reasonable enough to be used as a baseline with which to compare mixed logit.

Secondly, mixed logit does not immediately seem to be a substantial improvement on standard logit, at least not in the way that it was applied in this research. The mixed logit specification for single project choice implies that a project's cost is a main explanatory attribute for the choice probability of a project. Though it is not surprising that a higher price leads to lower desirability, this is an effect that is to be more expected in a consumer context rather than in the citizen context of PVE. In the context of fixed PVE, this large negative cost effect would mean that there is in this sample on average a high preference for shifting the budget to the next period. This could be the case if participants consider all projects to be somewhat unattractive and hope that next year's choice set will be better. In the case of flexible budget, it would imply that individuals gain higher utility from a one-time tax relief than from the realization of (one of) the project(s). However, this interpretation is contradicted by the fact that almost all participants do in fact select projects in their portfolio and that two relatively expensive projects (project 14 and 15) are selected relatively often. It is possible that the spread of projects and project choices leads to somewhat fuzzy results. This effect is also party present in the findings of Mouter, Koster & Dekker (2021). They find that at least half of the projects have a higher than 50% probability to increase social welfare. Positive as it sounds, it actually complicates the interpreting and advising process of PVE experiments. The fact that the cost parameter is the only remaining significant one, both statistically and economically, could be due to the fact that it is the most consistent one. That is, all projects have a non-zero value for project cost whereas the other project attributes are equal to zero for a substantial number of projects, as is the case for the number of trees cut.

Thirdly, looking at the project-specific parameters for project combinations, standard logit results contradict the project combinations that are expected to be relatively desirable for participants. Though some project-specific parameters for combination are relatively high, no real uniform conclusion is to be drawn from the standard logit results. Mixed logit seems to improve on this slightly, being able to capture more accurately which projects are relatively more likely to be selected. However, the results do not give a uniform indication of the effectivity of the model application. Although mixed logit shows potential in this application, the limitations that were part of this research, which will be explained in the next section, have to be overcome to be able to formulate a more definitive conclusion.

6.2 *Limitations*

It is important to note the limitations to the research design as was implemented in this paper. One rather obvious limitation is the fact that this research only looks at combination effects between two projects. In the context of a portfolio decision, this focus does not capture all possible synergies. The existence of synergies between more than two projects is rather likely, since the selected portfolios often contained more than two projects. A more general comment on this is the fact that approaches like the one in this research are always difficult due to the fact that a choice of portfolio is separated into choices for projects or project combinations. This simplification of the selection process as it occurs during the experiment is problematic. The degree to which this simplification occurs, and in turn the credibility of the remaining model (estimations), differs between different research design and the desired level is up for debate.

Finding a computationally efficient method for estimating positive and negative synergies in PVE data is useful. There are multiple possibilities to work towards this, which are also interesting alleys for further research. Firstly, it is possible to take the Bahamonde-Birke & Mouter (2019) framework into account when designing a PVE experiment. However, this would create a trade-off between computational capabilities and qualitative outcomes of the experiment. Keeping the number of possible portfolios low, by lowering the budget or limiting the number of different projects for example, can help with computations but it also puts a cap on the explanatory power of the results and allows for estimating a smaller number of possible synergies. Secondly, the addition of qualitative data to this approach could be helpful. As mentioned in section 3.1, Mouter et al. (2021) also collect statements by respondents on why and how they made their decision. Using methods for text analysis methods to comb out possible negative or positive synergies, which can then be affirmed by quantitative insights, could be a way to lower the amount of synergy variables. Additionally, it could be possible to simply ask respondents whether there were particular combinations of projects they liked or disliked. The resulting qualitative data can then be used to see whether such preferences can be found in the larger population as well.

6.3 Concluding remarks

Although the specifications used in this research do not definitively prove the existence of synergies for project combinations – or PVE's capabilities to capture those synergies – they also do not disprove it. The underlying theory of PVE suggests that there are possibilities to capture synergies and the research by Bahamonde-Birke & Mouter (2019) suggests their existence too. This research has made a small contribution to the empirical evidence, but not overwhelmingly so. The simplest conclusion that can be drawn upon the results in this research is that more research is needed to determine more securely that project synergies can be estimated from PVE data. The best available methods for analysis are not as set in stone yet as they are for CBA. This means that the use of CBA is probably not over its peak yet. More conclusive findings of the importance and size of synergies will have to be found to help PVE gain attention. Research papers like this one can and should aid in finding the most appropriate tool for ex ante valuation of government actions, since government budget can only be spent once.

7. Bibliography

Alphonce, R., Alfnes, F., & Sharma, A. (2014). Consumer vs. citizen willingness to pay for restaurant food safety. *Food Policy*, *49*, 160-166.

Bahamonde-Birke, F. J., en Planologie, S. G., & Mouter, N. (2019). About positive and negative synergies of social projects: treating correlation in participatory value evaluation.

Bergstrom, J. C., Boyle, K. J., & Yabe, M. (2004). Trading taxes vs. paying taxes to value and finance public environmental goods. *Environmental and Resource Economics*, *28*(4), 533-549.

Brownstone, D., & Train, K. (1998). Forecasting new product penetration with flexible substitution patterns. *Journal of econometrics*, *89*(1-2), 109-129.

Campbell, D. (2007). Willingness to pay for rural landscape improvements: Combining mixed logit and random-effects models. *Journal of agricultural economics*, *58*(3), 467-483.

Choy, Y. K. (2018). Cost-benefit analysis, values, wellbeing and ethics: an indigenous worldview analysis. *Ecological Economics*, *145*, 1-9.

Hensher, D. A., & Greene, W. H. (2003). The mixed logit model: the state of practice. *Transportation*, *30*(2), 133-176.

Kroes, E. P., & Sheldon, R. J. (1988). Stated preference methods: an introduction. *Journal of transport economics and policy*, 11-25.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

Mouter, N., Van Cranenburgh, S., & Van Wee, B. (2017). Do individuals have different preferences as consumer and citizen? The trade-off between travel time and safety. *Transportation research part A: policy and practice*, *106*, 333-349.

Mouter, N., Van Cranenburgh, S., & Van Wee, B. (2018). The consumer-citizen duality: Ten reasons why citizens prefer safety and drivers desire speed. *Accident Analysis & Prevention*, *121*, 53-63.

Mouter, N., Koster, P., & Dekker, T. (2019). An introduction to Participatory Value Evaluation.

Mouter, N., Koster, P., & Dekker, T. (2020). Participatory Value Evaluation for the evaluation of flood protection schemes. *Working Paper*.

Mouter, N., Koster, P., & Dekker, T. (2021). Contrasting the recommendations of participatory value evaluation and cost-benefit analysis in the context of urban mobility investments. *Transportation Research Part A: Policy and Practice*, *144*, 54-73.

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

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.

Romijn, G., & Renes, G. (2013). General guidance for cost-benefit analysis. *CPB Netherlands Bureau for Economic Policy Analysis and PBL Netherlands Environmental Assessment Agency, The Hague*.

Revelt, D., & Train, K. (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics*, *80*(4), 647-657.

Small, K. A., Winston, C., & Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, *73*(4), 1367-1382.

Spash, C. L. (1997). Ethics and environmental attitudes with implications for economic valuation. *Journal of Environmental Management*, *50*(4), 403-416.

Train, K. E. (2003). *Discrete choice methods with simulation*. Cambridge university press.

Van Wee, B. (2012). How suitable is CBA for the ex-ante evaluation of transport projects and policies? A discussion from the perspective of ethics. *Transport Policy*, *19*(1), 1-7.