The effect of the construction of cycling highways on cycling counts

An impact assessment of cycling highways and their usage in the province of Gelderland in the Netherlands.

Figure 1 RijnWaalpad Nijmegen - Arnhem. Source: Wikimedia

Master's thesis Spatial Planning – Urban and Regional Mobility
Paul van der Horst
February 2021

The effect of the construction of cycling highways on cycling counts

An impact assessment of cycling highways and their usage in the province of Gelderland in the Netherlands.

Author:
Paul van der Horst
Student number: s4628993
E-mail: pfa.van.derhorst@student.ru.nl

Master program:
Master spatial planning
Specialization in Urban and Regional Mobility

University:
Radboud University Nijmegen
Nijmegen School of Management
Nijmegen, the Netherlands

Supervisor:
Supervisor: dr. H. Ploegmakers

Preface

This thesis about the effects of cycling highways is the culmination of my period studying at the Radboud University. The road to completion was not always easy, however now that it is complete, I am proud to present the result. Writing this thesis, I learned a lot about infrastructural interventions, evaluation design, Stata and many other things. I want to thank Huub Ploegmakers for his guidance during this thesis. Without his help I would have stranded in, among other things, the bog of data preparation and the jungle that is difference-in-difference method. Besides that, I want to thank Sophie, my roommates, friends and family for keeping me sane during this weird period that is a pandemic. I want to close this preface with a quote from Blaise Pascal that kept running through my mind during the writing of this thesis.

"All of humanity's problems stem from man's inability to sit quietly in a room alone."

Summary

This thesis starts with the establishment of a problem. Namely that there is not enough ex-post evaluation research on the effect of cycling highways. It is interesting to fill this knowledge gap. To fill this gap this thesis tries to answer the question whether the completion of cycling highways change the bicycling counts on these routes. It tries answer this question using an approach that has not been used extensively when it comes to analyzing cycling highways. This thesis namely uses an ex-post evaluation design to try to estimate the effect of cycling highways. The evaluation design is an impact assessment with a natural experiment. The natural experiment is chosen because there is no control over the intervention. The analysis method that was chosen with this design is a difference-in-difference method.

To be able to use the methods and design these designs need to be understood well. To do this this thesis first dives into the evaluation literature. After this the relevant literature surrounding cycling interventions and cycling highways is reviewed and used as a basis for this research. Next the concepts that have been discussed in the second chapter are used in the third chapter to build the research design. In the fourth chapter the results of this thesis are presented using various models and analysis types. Using the most complex model the effects of cycling highways are estimated to increase the bicycle count per hour on these routes by 39,8%. However, in the fifth and sixth chapter this number is nuanced. The shortcomings and lessons from this thesis are also discussed in these chapters.

Inhoud

Appendices

List of Figures

List of tables

1. Introduction

1.1 Research problem statement

1.1.1 The demand for more cycling

Cycling as a transportation method has various benefits over other modes of transport (de Hartog et al., 2010; Handy et al., 2014). When compared to the use of a car cycling can potentially provide health benefits for individuals and society as well as environmental and economic benefits. Increased levels of cycling are for example linked to a lower mortality (Handy et al., 2014). More cycling also leads to less sick leave and therefore economic benefits (Fishman, Schepers & Kamphuis, 2015). It is therefore no wonder that various governments have aims to increase the modal share of cycling. Here in the Netherlands this can be seen on various governmental levels. In 2018 undersecretary for infrastructure and water management Van Veldhoven announced that she announced the goal of 200.000 more commuters on bikes and out of cars citing that this contributes to the national goals of more accessibility, living standards, sustainability and health. There also is the goal of 3 billion more kilometers on bikes compared to 2017 (Rijksoverheid, 2018). The total distance cycled in the Netherlands in 2017 was 14,5 billion kilometers (CBS, 2018). Taking this into account means that a major increase in the use of bikes and cycling is needed. The national government is investing €250 million euros in different measures and policies to achieve the intended increase in bicycle use (Rijksoverheid, 2018). Not only on the national level, but also on the regional and local level government agencies are trying to increase the use of bicycles. The Fietsersbond (cyclist' union) published a document where the different provinces and their programs related to cycling are summarized (Fietsersbond, 2019). The different provinces have various ambitions when it comes to cycling policy. The province of Utrecht has one of the more ambitious plans. Utrecht wants to make the bike the most attractive mode of transport for trips under 15 kilometers. They will invest close to €100 million euros until 2023 to try to achieve this goal *(Uitvoeringsplan fiets, 2019).* Less ambitious are for example the plans of the province Noord-Brabant. It has the ambition to increase the number of trips by bike by 75.000 this year compared to 2016. One of the measures for this is to increase the amount of cycling highways (Fiets in de versnelling, 2009) and Gelderland wants 35% of all trips being completed on (electric) bikes in 2030. This in comparison to 27% now (Koersdocument Duurzame Mobiliteit, 2018). Most provinces argue that stimulating the use of bicycles is both cost effective and has many benefits such as accessibility, climate and economy (Fietsersbond, 2019). On the local level various municipalities also have their plans to increase the usage of bikes. Here there is also a difference in ambition. For example, the city of Utrecht wants the bike to be the primary mode of transport in 2030 (Actieplan Utrecht fietst, 2015). Enschede has a less ambitious program and wants a 4-percentage point increase in the share of bike use this year compared to 2012 (Enschede Fietsstad, 2020). Lastly the municipality of Nijmegen wants to increase the use of bikes by 20% in 2027 compared to 2017 *(Ambitiedocument Mobiliteit, 2018).*

What is clear is that there are multiple governmental agencies that all want to increase the use of the bicycle as a mode of transportation. There is also a willingness to invest in measures and policy to achieve this goal.

1.1.2 Cycling highways as an answer

Interventions in cycling infrastructure has been one of ways in which cycling has been promoted in the past and is seen as one of the main ways to get more people to cycle (Mölenberg, Panter, Burdorf & van Lenthe, 2019). Interventions are changes in the cycling infrastructure are physical changes such as cycling paths or cycling bridges. One of the newer developments in infrastructural interventions is the development of cycling highways. Cycling highways are high quality bicycling paths were only cyclists are allowed and are meant for fast commuting over long distances typically up to 15 kilometers (European cyclist federation, 2014; Thiemann-Linden & Boeckhout, 2012). These paths have several characteristics that make them different from normal cycling paths. The European Cyclist Federation (2014) for example defines cyclist highways as being at least 5 kilometers long, separated from motorized traffic and pedestrians and avoid frequents stops. In the Netherlands they are a way to achieve several policy goals such as reducing traffic jams and reduce the pressure on public transport. They are mostly aimed at commuting traffic up to 15 kilometers (van Esch et al., 2017).Cycling highways are being developed in several countries such as Denmark, Belgium, and the Netherlands (European Cyclist Federation, 2014). Figure 1 shows the cycling highways that already have been completed and the plans. Some projects such as the cycling highway between Nijmegen and Arnhem are already completed. These completed projects are in green. However, even more projects, in yellow and grey, are being developed, explored and planned. What is clear that cycling highways are being planned all around the country.

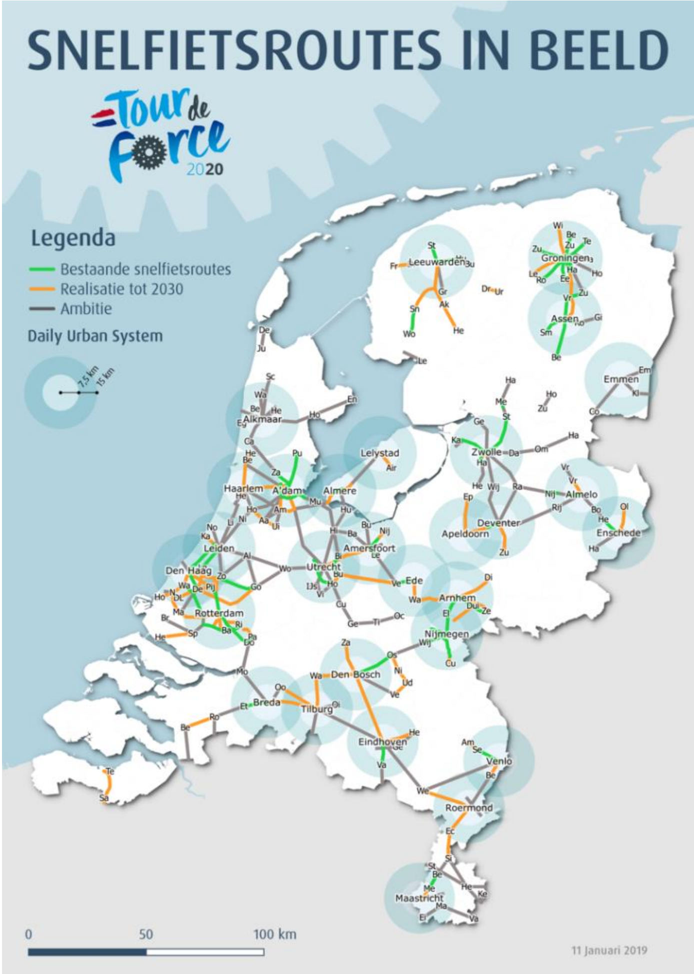


Figure 1 Already completed cycling highways are in red. The paths in green are being realised. The paths in yellow are currently in the planning fase and on the grey paths the possibilities are being explored. Souce: Rapport Tour de Force 2020.

1.1.3 State of research and practice

To answer the question why cycling highways are being built a look at research literature and policy documents is needed.

Starting with the research literature. As said before interventions in the cycling infrastructure is one of the main ways to increase the level of cycling. These interventions can vary from painting bike lanes next to existing roadways to creating extensive bike networks or cycling highways. Various studies have been done about the effectiveness of these interventions (Buehler, Pucher, 2012; Rayaprolu et al., 2018; Skov-Petersen et al., 2017; Stappers, van Kann, Ettema, de Vries & Kremers, 2018). Systematic reviews on infrastructural intervention studies, such as the one published by Molenberg et al. (2019) and Stappers et al. (2018), show that most studies find a positive effect of these interventions on the amount of cycling. These interventions are varied. From painted bike lanes on shared roadways to bike bridges. For example, Dill and Pucher's (2011) study in 90 American cities for example found that investments in cycling paths and lanes correlate with a higher percentage of cycling. This positive correlation has also been found for cycling highways. Research shows upgrading to cycling highways increased the use of these paths. It also increased the satisfaction of the users of these paths (Rayaprolu et al., 2018; Skov-Petersen et al., 2017). While various studies have been done there are still interesting opportunities for further research. Buehler and Dill (2016) note several of these opportunities. For example, there have been relatively few

studies that track changes in the use of cycling because of interventions over longer periods of time. There have also been few studies about specific types of cycling infrastructure. Mentioned are for example specific designs or the quality of paths. Cycling highway, although there has been limited research, could also be included in this list.
As said various governmental organizations want to increase the amount of cycling. The provinces of Groningen, Utrecht, Gelderland, Noord-Holland and Overijssel see cycling infrastructure as one of the measures that can be taken (Fietsersbond, 2019). For example in the Koersdocument Duurzame Mobiliteit (Provincie Gelderland, 2018) the province of Gelderland states that because of the increased growth of (quick) e-bikes longer distances can be reached on bikes and to utilize and stimulate this potential a high quality cycling network is needed. This network also includes cycling highways. The national government has reserved €250 million for measures to increase cycling where €26 million has been specifically allocated to cycling highways. The rest of the budget goes toward improving cycling routes and parking facilities.
Interventions in bicycling infrastructure and cycling highways are therefore interesting aeras for research. One the one hand there are still unknows in the knowledge and on the other hand there are governmental organizations that want to invest in these cycling highways.

1.1.4 Knowledge gap

What then is the knowledge gap that exists when it comes to cycling highways? And how can this thesis address this knowledge gap? Firstly, there is a growing demand for cycling related policy and measures. This is seen in the increasing attention and investment in ways to increase cycling in the Netherlands. Specifically, infrastructural investments are a way to increase cycling volumes. As said one of these measures are cycling highways. While this attention for cycling highways is growing the knowledge around these highways is not yet conclusive. Secondly there has been limited ex-post designed studies in planning in general and on cycling highways specifically (Guyadeen & Seasons, 2018).

1.2 Research aim and research questions

1.2.2 Research aims

This research has the following research aim:

- Contribute to the knowledge about cycling infrastructure interventions and their effects on bicycling volumes, specifically about cycling highways and their effects on cycling volumes.

This research statement has several components. The first is to contribute to the knowledge of cycling highways as infrastructural interventions to increase cycling. As is clear from the knowledge gap (§1.1.3) the current scientific research on this topic is not exhaustive. This research will try to fill that gap through examining cycling highways in Gelderland and other provinces in the Netherlands.

The second component is to improve the knowledge of policy makers. Through filling a knowledge gap policy maker can get a better understanding of the implications of their decisions and therefore make better informed decisions.

The study will be an evaluation study. This means that the aims of this research will be achieved through this research design. This research is chosen because it fits best with the available data and aims. This will be further explained in chapter 2 and 3.

1.2.3 Research questions

The research aim leads to the following main question and sub questions.

The main research question is as follows:

''How does the cycling infrastructure intervention of cycling highways impact the count of cyclist?''

This research question fits the research aim. Through the examination of cycling highways and their effect on cycling volume the research aim is targeted. To know how cycling highways, influence the amount of cyclist there needs to be data on the amount of cyclist on routes before and after the implementation of cycling highways. Secondly to effectively use this data the question how an evaluation study looks needs to be answered.

1.3 Societal and scientific relevance

In this paragraph the societal and scientific relevance of this research will be discussed.

1.3.1 Societal relevance

This research has societal relevance in several ways. Firstly, cycling has many benefits (De Hartog et al., 2010; Handy et al., 2014). More cycling leads to less risk of getting cancer and a lower mortality rate in general (Oja, Titze, Bauman, de Geus, Krenn, Reger-Nash & Kohlberger, 2011). This decrease in mortality outweighs the risk of increased inhalation of pollutants and increased risk of accidents (Handy et al., 2014). These health benefits also mean that increased levels of cycling have economic benefits through among other things less sick leave (Fishman, Schepers & Kamphuis, 2015). Besides health and economic benefits cycling also contributes to a better environment. Increased levels of cycling, and with that decreased levels of car use, contribute to better air quality (Garrard, Rissel & Bauman, 2012).

With these benefits in mind it is not strange that many governmental organizations want to increase levels of cycling. One of the measures to achieve this are cycling highways. This increased interest in cycling highways, and with that increased funds, are also reasons why this research is relevant for society. To be able to make well-argued decisions about the distribution of these funds, knowledge about the effects of these decisions is needed.

Lastly there are not enough ex-post studies on cycling highways. Most studies evaluating the impact of cycling highways are ex-ante design. Most of these are MKBA's (social cost-benefit analysis) (Decisio, 2012). These try to model the impact of cycling intervention before they are made. With this method of evaluation there are various assumptions made (Hanemaayer, 2012). These assumptions are not always fully backed up by research. The rapport MKBA van de fiets (2012) explicitly mentions that there needs to be more research on these effects. It also states that filling in, among other things, this gap would lead to better usefulness of the MKBA.

1.3.2 Scientific relevance

Besides societal relevance this research also has scientific relevance. Firstly, it contributes to scientific knowledge because of the gap that currently exists in the literature regarding cycling highways and their effects. As discussed, the current research is not yet conclusive. This research aims to contribute to this knowledge and expand it. Secondly this research tries to contribute to the knowledge of good evaluation study design. By trying answer the question what a well-designed infrastructural intervention evaluation study looks like this thesis tries to contribute to this scientific knowledge. As said Buehler and Dill (2016) name several missing links in the current research literature when it comes to intervention in cycling infrastructure. Many studies rely on self-reporting of participants on their amount of cycling. This means that these studies are not always representative for the whole community. Secondly cross-sectional studies and not longitudinal studies are the most common among bicycle infrastructural intervention studies. This means that the length of the study is most of the time short. And lastly newer types of infrastructure such as cycling highways have not yet been studied to the extent that some other interventions which have been studied intensively. Stappers et al. (2018) also found that the effect of built environment infrastructural changes varies for the types of interventions. This means that the effect of cycling highways might be different from other interventions. It is therefore interesting to research. The knowledge gap is thus that the relatively new intervention of cycling highways has not yet been extensively researched and that a research that does this could contribute on the knowledge about cycling related infrastructural interventions.

2. Literature review and theoretical framework

2.1 Evaluation

Evaluations have been done in various ways over the years. Different forms of evaluations are possible. In this chapter different theoretical frameworks and approaches will be discussed. Also, different parts of an evaluation will be dissected. Evaluation theory is part of the theoretical framework because of the many differences in evaluation and to assess what best fits this research. the framework. As said different kind of evaluations can be done**.** In this chapter the focus will be on program evaluation. To be able to discuss program evaluation firstly a definition is needed.

2.1.1 Definition

Evaluations have a long history in society. Already in the 19th century there were evaluations being done in for example Great Britain and the USA (Stufflebeam et al., 2000). These evaluations where aimed at reforming the educational system and other social agencies. During this time official agencies, such as Royal commissions in Great Britain, were also set up (Stufflebeam et al., 2000). Stufflebeam et al. (2000) named this period until 1900 as the first period in the history of evaluation the age of reform. After this first period several other eras in evaluation can be identified. In the period after 1900 the practice of evaluation steadily gained more attention. New techniques and methods were developed. These developments accelerated in the 1960's and the beginning of the 1970's (Stufflebeam et al., 2000). In this period discussion evolved around how evaluation should be conceived. In the following decades evaluation further professionalized and institutionalized. In all these years many different evaluations have taken place: from evaluating events or people to processes or things (Rossi et al., 1998). The evaluation of policies is also one type of evaluation. The focus in this thesis will be on this kind of evaluation. Policy evaluations are often also called program evaluation. There are several definitions of program evaluation (Guyadeen & Seasons, 2018; Rossi et al., 1998). Guyadeen and Season (p.99, 2018) define program evaluation as: *"systematic assessment of the operations and/or outcomes of a program, compared to a set of explicit or implicit stands, as a means of contributing to the improvement of the program*". Another definition by Rossi, Freeman and Lipsey (1998) is as follows *''Program evaluation is the use of social research procedures to systematically investigate the effectiveness of social intervention programs that is adapted to their political and organizational environments and designed to inform social action in ways that improve social conditions.'' (*Rossi et. al., p.20, 1998). While these definitions are somewhat different there are several components that overlap and seem to be important in defining program evaluation.

Firstly, program evaluation investigates some form of program (Rossi et al., 1998). A program is a set of planned actions that try to have an effect in a specific audience. Secondly program evaluation has some systematic assessment or investigation of this program. This means that there is an organized way of inquiry into a topic. There are certain standards of quality when it comes to research. The evaluation also has some form of judgements. This means that the program is judged on certain criteria. There needs to be a valid way of making this judgement when looking at the program and it effects. Thirdly the goal of an evaluation is to improve the program. An evaluation is not just done to assess the program but also to improve it in the future.

A difference in these definitions is that Rossi et al. (1998) add that social programs are adapted to their political and organizational environments. This is important to keep in mind. Programs can be intended to operate in a certain way. However, there can be a difference between intention and implementation. There might be certain differences that might influence the effectiveness of the program. All in all, we can distill four components that are important in the definition of program evaluation. Firstly, there is a certain program that is being investigated. Secondly this investigation happens according to certain methods. Thirdly there is some form of judgement of the program and this is used to improve the program. Lastly it is important to keep in mind the context of the program and its implementation when doing a program evaluation.

2.1.2 Types of evaluations

Stage of program development	Question to be asked	Evaluation function	Explanation evaluation function	Formative or accountability
Assessment of social problems and needs	To what extent are social needs met?	Needs assessment	An evaluative study that answers questions about the social conditions a program is intended to address and the need for the program.	Formative
Determination of goals	What must be done to meet the needs of the society?	Needs assessment		Formative
Design of program alternative	What services could be used to produce the changes needed?	Assessment of program logic or theory	An evaluative study that answers questions about the conceptualization and design of a program.	Formative
Selection of alternative	Which of the possible programs' suites best?	Feasibility study	An evaluation where the different programs are best possible.	Formative
Program implementation	How should the program be put into operation?	Implementation assessment	An evaluative study that answers questions about program operations, implementation, and service delivery.	Formative
Program operation	Is the program operating as planned?	Process evaluation	An evaluative study where the program operation is evaluated.	Accountability
Program outcomes	Is the program having the desired effects?	Outcome evaluation or impact assessment	An evaluation study that answers questions about program outcomes and impacts	Accountability

Program efficiency	Are program effects attained at a reasonable cost?	Cost-benefit analysis	An evaluative study that answers questions about program costs in comparison to either the monetary value of its benefits or its effectiveness in terms of the changes brought about in the social conditions it addresses.	Accountability

Table 1 Types of program evaluation. Adopted from: Rossi et al., 1998

An important question when it comes to evaluations is what is the purpose of the evaluation? The types of evaluation that can be done depends on the type of questions that are asked (Rossi, Freeman, & Lipsey, 1998). Rossi et. al. (1998) distinguish four main reasons for doing an evaluation. Firstly, they mention formative evaluations. The goal of the formative evaluations is to improve the performance of the program. This means that they influence the program directly (Guyadeen & Seasons, 2018). To achieve this goal a formative evaluation means that it is often focused on program design, implementation, impact or efficiency. The purpose of the evaluation means that often the evaluator and the stakeholders work closely together during the program (Rossi et al., 1998).

The second form of program evaluation is aimed at accountability. These evaluations are also known as summative evaluations. Summative evaluations are aimed at the results of programs. In comparison to formative evaluations this means that they are done after the program is (nearly) complete. The goal is then to inform the decision makers if the program was successful in achieving the goals that were set out. This is useful because programs have a certain cost and decision makers, or critics of the program want to be informed of the effectiveness of the program. Rossi et. Al. (1998) note that these evaluations therefore must adhere to a sufficient scientific standard to be credible. This also extends the participation of stakeholders. There can be input from the stakeholders, but there should be no collusion.

The third reason, knowledge generation, fits with the post positivistic paradigm. These evaluations that are done because of knowledge generation are mostly contributing on how interventions work. These evaluations are therefore mostly intended to expand scientific knowledge. This implies that evaluations are not preformed to inform a decision makers or critics. This does not mean that these evaluations do not contribute to decision making. However, the findings of these evaluation might indirectly be useful for the development of new programs (Rossi et al., 1998). In these types of evaluation it is important to adhere to rigorous scientific framework.

The last kind of program evaluation is public relations. This goal is different from other types of evaluation. In a public relations evaluation study the purpose of the evaluation is not to gain knowledge or to improve a program. The purpose is political (Rossi et al., 1998).

Not only the reason for doing an evaluation matters when it comes to doing a program evaluation. The design of the evaluation is also dependent on the part of the program that is going to be evaluated. In figure 3 the stage of the program and the function of the evaluation is listed. The function of the evaluation is dependent on the stage of the program development and the question that needs to be answered.

In the first stage of the program the needs of the program often need to be assessed. These kinds of evaluation are called a needs assessment (Rossi et al., 1998). For the creation of a social program recognition of social problems is required. Something must be a problem to plan an intervention. A needs assessment can be made through surveying informants or analyzing data and statistics (Rossi et al., 1998). The results of a needs assessment are often recommendations on the how a program best fits on the needs that exist in society. A needs assessment can be used to design a new program or adjust an existing program to the needs that arise.

If the problem is recognized and the need for an intervention is clear the next evaluation type may be more useful. The social program needs to fit the problem it is trying to solve. An assessment of program theory asks the question what can be used to achieve the desired effect? (Rossi et al., 1998). This type of evaluation is often needed in the early stages of the program and seek to fit the program design with the intended purpose (refer to work on program theory). A part of this can be the feasibility study. In this study the alternatives that are possible are weighed (Rossi et al., 1998). This design of program alternatives answers the question which of the alternatives is most likely to have the desired effects and is most cost effective while achieving this?

If in theory the program is assumed to have the intended effect and the most suited alternative has been chosen, the next step is to implement the program. The implementation of the program can differ in its success. Between theory and successful implementation there are several hurdles. For example, the personnel that needs to implement the program can be poorly trained or the target demographic does not want to participate in the program. If the implementation is not working properly this is also known as an implementation failure. If the implementation is successful but does not have the desired effect this is known as a theory failure (Rossi et al., 1998). These possible problems in the organization or delivery of the program are investigated in this kind of evaluation. An implementation assessment thus assesses the effectiveness of the implementation of the program (Rossi et al., 1998). An implementation assessment is also known as a process evaluation. During this type of evaluation it is important to identify the crucial functions of the program and the corresponding criteria for success.

If the implementation of the program is completed the next evaluation step is an impact assessment. An impact assessment tries to measure to what extent the program or intervention has the intended effects on the problem it addresses and if there are other effects: it measures the outcome of programs (Rossi et al., 1998). In these kinds of evaluations it is also important that the objectives and successes are well defined. Clear outcome variables that can be measured are needed. Based on these objectives and associated success measures an impact assessment tries to estimate the effects of the program. To measure the effects data is needed. This data needs to be collected. The data must show the effects of the intervention. The effect also needs to be confidently attributed to attributed to the intervention and not to other causes. As Rossi et al. (1998) explain this is the hard part of an impact assessment. The counterfactual, or how the target group of the intervention would have been without the intervention, needs to estimated. Ideally an experimental design with control and experimental groups that are randomly assigned. However, this is often not possible due to practical constraints (Rossi et al., 1998). When this is the case different designs such as quasi-experiments might be needed. Rossi et al. (1998) note that an impact assessment is most useful when it is important to learn about the program effects because the program is for example innovative or it is the basis for further action. The conditions for undertaking an impact assessment also need to be suitable. A well-defined program with data of the results are most suited to an impact assessment.

An impact assessment determines the effect of a program however this does not say anything about the cost of the program. An efficiency assessment weighs the results of the program against the costs of the program (Rossi et al., 1998). A program for example can produce results but if the costs of this program are too high the program might still be cancelled: budget for programs is often limited and the effects need to be worth it. Two types of efficiency assessment can be undertaken. Firstly a cost-effectiveness analysis. This analysis looks at the cost per unit of outcome (Rossi et al., 1998). Secondly a cost benefit analysis. This looks at the cost and benefits in monetary terms (Alkin & Rossi, 2012). The difficulty in this type of evaluations is that it can be hard to translate benefits into monetary terms. As is the case in the previous types of evaluations during an efficiency assessment it is also important that the program and its results are clear. If for example not all the benefits are documented the cost for unit of result might be higher.

All these approaches and forms of evaluations have different roles and uses. It is therefore important to weigh these different approaches and see what fits best with this research. Choosing the type of evaluation to do depends on different factors. The context and the progress of the program that is to be evaluated are important (Rossi et al., 1998). If the intended purpose for example does not suit the needs to be addressed well a program theory evaluation might be best suited. However, if the program is working well established but the effects are unknown an impact assessment might be more suited. In this research the question is what the effects of cycling highways are. More specifically the program of constructing cycling highways and the effects of this program is to be assessed. An impact assessment will be used to assess these effects. This fits well for several reasons. Firstly, the program is well-defined and mature enough to do an impact assessment. A significant amount of cycle highways has already been constructed and are in use for several years. Secondly there is data available to use for an impact assessment. Various counting points are used on cycling highways and other cycling paths. This data is required to assess the impact of these cycling highways. Because there is data available on both cycling highways and other cycling paths it is possible to create an evaluation design that permits the estimation of the counterfactual. Lastly the impact assessment can be used well for decision making in the future. Multiple cycling highways are still being planned and constructed. This means that the assessment of existing cycling highways can contribute to these planned highways.

How can the assessment of cycling highways contribute to these plans? As discussed in the societal relevance often in the first stages of an infrastructure project in the Netherlands a (social) cost-benefits analysis is made (Hanemaayer, 2012). This is an ex ante analysis of the possible costs and benefits of a project. The problem with this analysis is that several assumptions are made about the effect of an intervention. As noted by the rapport Waarderingskengetallen MKBA Fiets (2017) there is still a lack of traffic models that predict how much extra traffic new infrastructure generates and where this extra traffic comes from. This means that this evaluation can help fill the gap in the knowledge that exists. It can help make better traffic models.

2.2 Impact assessment design

An impact assessment will be used in this research. However, there are still multiple ways in which an impact assessment can be done. In the following paragraphs the nuances and differences in impact assessments will be discussed.

An impact assessment is carried out after the effects of the program are supposed to be visible. The goal is to try to estimate the net effects of the program. Because the goal is to understand the effects of cycling highway programs an impact assessment suits the research. Impact assessments are ideally an comparison between two or more groups: those who received the intervention and those who did not (Rossi et al., 1998). However sometimes due to practical or other constraints it is not possible to compare between groups. As said ideally this would mean that randomized lab experiments would be used. The ideal experimental setup is also known as a randomized controlled trail (RCT). In a RCT subjects are randomly assigned to two or more groups and receive different treatments or no treatment. However, because of practical and time concerns this is often not possible. A researcher often does not have any control on the intervention. For example, the subjects sometimes cannot be assigned randomly. The relation between a program and its impact can therefore not be assessed in a straightforward manner. There are several caveats.

2.2.1 Confounding factors and bias

Firstly, there might be several confounding factors. These are unknown variables that influences the variables the impacts assessment tries to understand (Rossi et al., 1998). The confounding factors distort these variables. Selection bias is one of these confounding factors. This is when the selection of participants is not random. For example, when a social program is voluntary there might be self-selection in the participants that choose to enter the program. While this self-selection is one of the forms of uncontrolled selection this bias can also be found in whole communities. Some municipalities for example might be more inclined to invest in interventions than other municipalities. This is also called uncontrolled selection (Rossi et al., 1998). Another confounding factor is that the social program that is being researched is not the only one that is active during this time. Other programs might influence the results of the program for which the impact is being assessed. Besides these known confounders there might also be unknown cofounders. These are the effects that are not known but are there. It is still important to try and control for these factors. This can be done through the design or in the analysis.

Secondly there is also endogenous change. This is the change that happens naturally over time (Rossi et al., 1998). These endogenous changes can have different forms. Secular drift are long-term trends that might mask the net effects of the program. This is when a long-term trend is opposite to the effect of the intervention. In a cycling highway intervention this might be a long-term trend where cycling levels are declining while the cycling highway might have a positive effect. This long-term effect would then mask the effect of the cycling highway. Contrary to long-term effects there can also be short term events that mask the effects of the intervention. These are known as interfering events (Rossi et al., 1998). These might be natural disasters for example. The last type of endogenous change are maturational trends. This is the natural aging of a population. When for example children are being studied on math skills for multiple years their natural growth might influence the results.

Thirdly there are also design effects. These do not come from outside but are the results from the research themselves. The first of these effects are stochastic effects. These are effect that happen by chance (Rossi et al., 1998). To combat these effects larger sample sizes can be used. What also can be used is statistical power. With increasing sample size effects sampling variance will be lower. This is how likely is an impact evaluation will detect a net effect when taking account, the study design. This statistical power is needed to make a judgement about two types of error. A type I error is a false positive: concluding that an intervention has effect while it does not. A type II error is the opposite.

Another important caveat that is important is the measurement reliability and validity of the impact assessment. The reliability of a measure is that it produces the same result every time. Unreliability might obscure the real effects of the intervention. Validity is the question whether a measure measures what it is intended to measure. This has various factors. Firstly, a measurement should be consistent with other studies on the concept. Secondly if the measure is consistent with other measures it is more valid. Thirdly the measure must be internally consistent. If multiple measures are used it should produce similar results. And lastly there should be consequential predictability.

Lastly the way in which the outcome variable is measured can also impact the results of the impact assessment. To be able to measure the effect of the intervention reliably the outcome variable should reflect the effect that is being studied. This requires that the outcome variable measures the effect of the intervention and what would have happened if they had not been exposed to the intervention. This is important because the casual interference between the intervention and the effect needs to be made plausible. The outcome variable needs to be attributed to the intervention and not to other factors such as spill-over effects (Mölenberg et al., 2019).

All these effects can influence the result of the impact assessment. There ways to minimize these effects. These ways involve establishing control conditions. For example, through statistical controls or time-series controls. In time-series controls multiple measurements are taken before and after the intervention. As said ideally one would have a randomized lab experiment to control for these factors. However, when this is not possible there are other research designs, such as natural and quasi-experiments, to control for various factors. The next paragraph will go further into the effect of the study design.

2.2.2 Study design and biases

To deal with confounding factors the study design is important. Study design can influence the way in which confounding factors play a role and can deal with biases in different ways. As said before the ideal experimental setup would be a randomized controlled trail (Rossi, 1998). A randomized controlled trail (RCT) is an experimental design in which the participants are randomly assigned into two or more groups. The second characteristic of an RCT is the difference in treatment that the randomized groups receive (Matthews, 2006). Because of the design a RCT has the least amount of assumptions and has a high statistical power. This means that it is useful when it comes to trying to prove a causal relationship between variables. However, there might several ways in which it is not possible to do a randomized controlled trail. Some requirements need to be met in order to be able to execute an RCT (Matthew, 2006). Firstly, there needs to be an eligible population from which groups can be made. Secondly, as said, there needs to be a random allocation into the groups, and they need to receive different treatments. The groups also need to be comparable. Lastly the differences between these groups needs to be compared. When this is not the case other experimental design need to be considered. Natural experiments and quasi experiments are two of the designs that can be used when a randomized controlled trail is not possible or not ethical. The designs however do have their own caveats and design effects. Besides two there are also other designs that can be used as a study design.

Starting with the quasi-experimental research design. In this research design the comparison groups are different from those in a RCT by the fact that they are not randomly assigned (Rossi, 1998). This means that groups might not be comparable. To solve potential bias in the groups matching or statistical methods can be used. This design is used when there is no control over the assignment of participants due to a variety of reasons. These can be political, ethical or other. For example, when it comes to life-saving treatments of diseases. The goal of the statistical methods or matching is to be able to make groups that are comparable. While these methods might make the groups more comparable there might always some uncontrolled difference between the groups. When it comes to designing quasi-experiments, it is therefore important to consider for an uncontrolled difference between the groups. Quasi-experimental groups can be created ex-ante or post-ante. This means that groups can be created after the intervention took place or before. All in all, quasi-experiments try to mimic randomized experiments as closely as possible. When it comes to the assignment to groups statistical methods or matching is used to make pseudo-random groups.

Secondly the natural experimental design can be used. This. A natural experiment is a certain type of experiment where control over the intervention is not in the hands of the researcher (Mölenberg, 2019). This means that the researcher cannot control the exposure of the intervention to the population. The difference with a quasi-experiment is that the assignment to the groups in a natural experiment is not chosen by the participants. The natural experiment has some advantages and disadvantages. Natural experiments make it possible to research interventions that are not able to be done as an RCT. For example, when an RCT would be unethical or cannot be performed. This is for example the case when it comes to large infrastructural changes. Because of the nature of the intervention it is near impossible to create a randomized controlled trail. These interventions are not planned randomly and therefore populations are difficult to place into random comparable groups. This means that in these situations a natural experiment is the better option. However, the fact that populations are not placed in random groups and are not comparable means that there are some drawbacks to natural experiments. There might be selective exposure to the intervention. With natural experiments there is also more risk of biases and inaccuracies. To combat these drawbacks the design of a natural experiment is important. Firstly, it is important to make a difference in exposure to the intervention. Even though the comparison groups might not be random there needs a way to differentiate between exposure to the intervention. On top of that some elements can reduce bias and unobserved cofounders in the natural experiment. Multiple measurements before and after the intervention and accurately measuring cofounders also needs to be done. Considering these elements several natural experimental study designs can be used to create well designed studies. Firstly, the difference-in-difference method can be used. In this method changes in groups that are exposed and not exposed to the intervention are compared. It assumes that possible confounders are the same in the groups. The confounders that do vary across the groups are assumed to be time-invariant. This means that differences between groups are assumed to stay the same over time. For example, pre and post intervention. The time varying cofounders are assumed to be equal for the groups. This means that if changes, such as secular drift, occur the are assumed to be the same over time. All in all, this should isolate the effect of the intervention. Any change that would occur should then be able to be attributed to the intervention. Another design that can be used is a regression discontinuity design. In this design a level of a variable is chosen to divide groups. Above this level is considered exposure to the intervention while under this level is considered not exposed. These groups are then used to analyze the difference between these groups (Craig et al., 2017).

These two methods are useful for accounting for unobserved variables. Observed differences can be best tackled with other methods. Firstly, matching can be used. This is when individuals with similar characteristics in the treatment and non-treatment are matched. Alternatively, statistical adjustments can be used. If known differences exist this can be compensated for in the analysis (Craig et al., 2017). All in all, is a natural experiment a study design that can be used when there is no control over the distribution of the intervention. This design does however have some caveats when it comes to biases and cofounders. These can be addressed with different methods.

Other methods are also available when a randomized controlled trail is not possible. Firstly, a meta-analysis can be done. This is when an aggregate study is done, and relevant studies done on the subject are reviewed and combined. The idea is that biases that exist in individual studies even out over multiple studies. Observational studies might also be used. In these study design constructed groups are made that are afterwards analyzed through statistical methods. A strategy that can be used with this design are propensity scores. Propensity scores are used to estimate the likelihood that a participant entered a certain group and thus receive treatment (Thoemmes & West, 2011).

For this thesis the choice is made to use a natural experimental design. This will be used in combination with a difference-in-difference method. There are several reasons why this strategy is chosen. Firstly because of the nature of the intervention. The intervention, cycling highways, cannot be controlled by the researcher. Some cycling highways are already built, and some plans are underway. On top of that the exposure of the intervention to the population is not random. There is purpose to the placement of the cycling highways. What this means is that a randomized controlled trail is not possible, and an alternative method is needed. Secondly the data on cycling highways is of good quality and the population that is exposed to them is large. The intervention is also relatively large. These characteristics are named as being useful for natural experiments (Craig et al., 2017). This is because this makes it easier to compensate for biases and cofounders. The choice is not made for a quasi-experiment because exposure to the interventions is not chosen by the participants. While people might choose to ride on cycling highways the do not chose themselves to create a cycling highway in their neighborhood.

The difference-in-difference method is chosen because it fits well with the intervention and the data. There is data available of multiple years before and after the interventions. Also, because there is data on cycling highways and other routes these can be compared with each other.

2.3 Bicycle infrastructure interventions

Cycling highways are infrastructural changes. It is therefore important to discuss the effects of infrastructural changes. There have been many studies that researched the effect of infrastructural changes of cycling infrastructure (Buehler & Dill, 2016; Mölenberg et al., 2019; Pucher et al., 2010; Stappers et al., 2018). Buehler and Pucher (2012) for example looked at the influence of bike paths and lanes on the level of cycling in various American cities found that an increased supply of bike paths increases the amount of cycling. This positive effect of infrastructural intervention is seen in more studies. A study done in Brisbane in Australia following the construction of a new bikeway increased the amount of cyclist in the city (Heesch et al., 2016). A review done by Buehler and Dill (2016) found that most studies find a positive relation between infrastructure and levels of cycling. However, this is not the complete story. As Mölenberg et al. (2019) note there is a difference in studies regarding bicycle infrastructure interventions. For example, there is a difference in results when different measurement methods were used. Different infrastructural interventions also resulted in different outcomes. In this section the types of built environment infrastructural changes and their effects will be discussed.

2.3.1 Categorizing infrastructural interventions

Starting with the different objects of study. The aforementioned study by Buehler and Dill (2016) categories three main domains of research that have been done on cycling infrastructure. These are links in the bicycle network, nodes of the bicycle network and the third domain combines the links and node. Under the first domain fall all sorts of infrastructure: from cycling paths painted on roads to separated biking paths (Buehler & Dill, 2016). There is a difference in research results between the type of infrastructure. Buehler and Dill (2016) differentiate between bike lanes, cycle tracks, bike paths and cycle track. Bike lanes are separated from motorized traffic by paint or another barrier but share the same road while cycle tracks are physically separated from motorized traffic but follow the road network. Bike paths are also physically separated but do not follow the road network. Other facilities are for example sidewalks were biking is also allowed.

The second domain of research in cycling infrastructure are the nodes of the bicycle network (Buehler & Dill, 2016). These are intersections with other roads and paths. There are two types of studies in this domain. The first type looks ate the characteristics of intersections (Buehler & Dill, 2016). These can look at the preference of cyclist when it comes to using or avoiding intersections or the study the traffic volumes on intersections. Traffic lights at intersections have also been studied. In these studies intersections are seen as having a greater potential for conflict and delay and are the preference for cyclists is to avoid these points (Heinen et. al., 2010). Besides preference these studies also look at interventions on intersections with the goal of promoting cycling have found that the safety at these intersections increases, the effects on cycling volume has not been well studied (Buehler & Dill, 2016). The second type of research focuses on bicycle-specific intersection treatments (Buehler & Dill, 2016). Bicycle specific intersection treatments are treatments such as bicycle activated signal crossings and bike specific boxes at intersections. Buehler and Dill (2016) note that there are a limited number of studies on this subject.

The third and final domain is research that combines both the nodes and links in the bicycling network and researches the network as a whole. There are various ways of measuring networks as a whole however remains difficult (Buehler & Dill, 2016).

Many earlier studies on bicycling networks used stated preference to obtain information about the quality of the networks. However newer more complex measures of cycling networks have been developed (Buehler & Dill, 2016). These use objective measures such as GPS data. An example of this is an index developed on the basis of safety and distance (Klobucar & Fricker, 2007). In this index the presence of a bike lane, traffic speed and other factors are considered and a level of service is calculated. This level of service score and other indexes however have not yet been used in empirical studies (Buehler & Dill, 2016). A form of these networks are cycling highways. These are connections of several paths, intersections and other forms that together form one highway. All these different forms

2.3.2 Results

These three categories mentioned above have been studied in multiple researches (Buehler & Dill, 2016). Mölenberg et al. (2019) and Stappers, et al. (2018) have reviewed some of the studies done on cycling infrastructure intervention. In this paragraph the results from these researches will be discussed using the reviews from Mölenberg et al. (2019) and Stappers, et al. (2018).

To be able to discuss the results it is first useful to elaborate on the reviews of Mölenberg et al. (2019) and Stappers, et al. (2018). Starting with the last one Stappers, et al. (2018) investigated 19 different built environment infrastructural changes (BEICs) and their effects on physical activity, active transportation and sedentary behavior in adults. The review focused on natural experiments, where the researcher has no control over the intervention, and quasi-experiments, where there is some control of the researchers over the intervention. There were several results from this review. Firstly, there is a bias in most studies (Stappers et al., 2018). The bias was observed in seven categories. The categories in which the most problems occurred were risk of bias in selection of participants, selection of reported results and bias in the measurement outcome. The lowest risk of bias was in the bias due to departure from intended intervention and risk due to missing data.

Mölenberg et al. (2019) set out to, similarly to Stappers, et al., (2018), to summarize the effects of infrastructural interventions on cycling levels and physical activity in adults and to evaluate whether study design and methods influence the results of these studies. The review included 31 studies and included a variety of interventions and outcome measures. The difference, as the authors note themselves, is that this review tries to quantitatively lists the results of the interventions. The review found that studies that reported behavioral changes were smaller than studies that reported the usage of infrastructure. Smaller effects were also found when using objective studies that had tested for statistical significance than those that did not. Other methodological differences in studies were also named. Casual interference was named as one of these. There was sometimes no way to control for this. Besides these spill-over effects were a source of bias in the results. For example, cyclist coming from other routes. Another important design effect was controlled versus uncontrolled studies. This means that there is a control population in the study that can be compared to the population receiving the intervention. This control is noted to also help with the casual interference (Mölenberg, et al., 2019). Lastly, as Stappers, et al., (2018) also note, the longer the infrastructural intervention has been completed the more effect was observed. Lastly, they note that it is important to include equity effects in the study design. Not many studies included these. Equity effects are population characteristics.

If these problems with biases and other confounding factors exist it in the varying studies, it is interesting to learn how various studies tried to address these issues. These will be discussed according to the seven categories.

The first bias found in studies was the problem of cofounding factors. Confounding factors are effects that influence how participants receive the intervention or not. These factors can among other things be demographic variables or weather. Rainfall for example can influence the number of cyclists for example. When it rains it is plausible that there will be fewer cyclists. This means that adjustment for these possible cofounders is needed. Mölenberg et al. (2019) found that various studies do not adjust for these cofounders. For example, a study done in Finland on an improvement of cycling and walking paths did not adjust for any cofounders (Aittasalo, Tiilkainen, Tokola, Suni. Slevänen, Vähä-Ypyä, 2019). They compared the use of these paths using a survey before and after the intervention and did not take into account any factors such as weather and demographics. The result was that there were no significant results on the frequency or distance cycled. However due to not considering possible cofounders' casual interference between the intervention and the effects is at risk. Due to these results of this study are weaker and can possibly contribute to not finding significant results. Another study done in the USA that looked at the effect of eight new bicycling boulevards did however control for possible cofounders (Dill, McNeill, Broach, Ma, 2014). In this study demographic variables, weather, distance to downtown and attitudes towards cars and bikes were considered. The results of this study were significant and showed a mixed result. More participants cycled at least ten minutes a day but made less trips a day. Taken into account the confounders in this study might have helped to find significant results. What is thus important for this thesis is that confounding factors should be considered. These can be weather or demographic variables for example. It may help the plausibility of the casual interference.

The second risk of bias is in the selection of participants. Stappers et al. (2018) not that in their review they only found one study that had a sample size calculation. This is needed to assess the amount of data that is needed to be able to get statistically significant results. Besides the number of participants needed in the study the way in which these participants are selected is also important when it comes to biases. Restrictions on the participants that can be selected can increase the risk of bias. The aforementioned study in Finland for example selected only from participants working in the area (Aittasalo, 2019). Other people living near the intervention were not able to participate. The risk of bias increases due to this choice. When only a part of the population can participate due to a certain characteristic, in this case work status, this might be related to the outcome. It is possible that workers might use the intervention differently from others. In other studies, the effect of selection is minimized in various ways. When studying the effects of new infrastructure in Cambridge

Heinen et al. (2015) presented the study to potential participants as a commuting study and not explicitly as a study evaluating new infrastructure. With this method they avoided potential bias in self-selection of participants into the study because of particular attitudes towards this new infrastructure. What is clear is that when it comes to the study population it is important to keep in mind the ways in which the selection process can influence the results.

The third risk of bias is in the measurement of interventions. This can occur due to wrongly classifying the status of the intervention to participants. This wrong classification of participants status might have an influence on the outcomes, however this is not necessarily the case. However, because the risk exists it is important to try and control for it. This can be done for example by validating surveys that have been done (Panter et al. 2016). Not all studies take these kinds of measures, however. A study using census data in the twin city area in America for example did not have any controls for the measurement of interventions (Krizek, Barnes, Thompson, 2009). It is thus useful to see if it is necessary to control for this possible bias when looking at the counting data.

The fourth bias can be due to departure from the intended intervention. This is when the intervention turns out different than was planned. This can be due to several factors. For example, something occurred in the control group but not in the intervention group after the initial division of the participants in these groups. Two groups with similar characteristics but in different cities might be compared in a study when an event occurs in one of the two cities that alters the characteristics. This then effects the intervention and its effect. An example of a study in which the risk of departure from intended intervention is possible is a study done on the effects of a new bikeway in Australia (Heesch, James, Washingoton, Zuniga, Burke, 2016). The bikeway consisted of three parts which were built during four years. The goal was to measure the effect of the third and final part of the bikeway (Stage C). In this study a measurement was taken before building the complete bikeway and one after. The problem with this method is that the effects of the intended intervention, stage C, was not the only effect that is possibly measured. The other two stages might also have influenced cycling levels. This means that there was a departure from the intended intervention. What is thus needed is a clear understanding of the intervention and possible other interventions. The measurements should reflect the intervention.

Bias due to missing data is the fifth possible risk. In studies where lots of participants miss lots of follow up appointments data might be incomplete. Another reason is that lots of data has been deleted from a dataset. If this data is incomplete the analysis might show biases. For this thesis that means it is important the datasets are as complete as possible and can be analyzed in full.

Another risk of bias is in the measurement outcome. This means measurements outcomes are measured with some error. This error can occur due to different reasons. For example, the measurement devices might not be working well. Another reason is that measurements outcomes were subjective instead of objective. Subjective for example are surveys, while automatic counting stations are objective. As noted before the outcome measure must be able to measure the effect of the intervention or program. These subjective measures might not accurately measure the effects. This is for example the case in the aforementioned study in Australia. In this study surveys are used in combination with GPS data from Strava. This is an app were users can register their activities (Heesch et al. 2016). The problem with these measurements outcomes is that the surveys are subjective, and the GPS data might be incomplete because only a sport-minded portion of the population uses this app. A better way thus to measure outcome is to use objective measures and use multiple sources of data.

The last risk of bias is in the selection of the reported result. With this bias some analyses or results might not be fully reported. This means that possible outcomes are not available and a complete picture of the effects of the intervention is missing. A risk for example exists when the choice for analysis has not been explained. Stappers, et al. (2018) found that most studies did not make it clear on what basis the analysis of the results was done. For this thesis it is therefore important to explain the choice for the analysis and show other possible analyses.

These risks should be considered when designing a study. Stappers et al. (2018) also have some more conclusion from their review. Firstly, it seemed that older articles seemed to be of less quality but yielding more results. That is, they found more significant results than newer studies which had more non-significant results (Stappers et al., 2018). This was attributed to the more complex design of these studies and thus less bias. This more complex design is for example due to incorporating proximity to the intervention into the study design. Duration of measurement after the intervention, in particular more than one year after, also seemed to improve the study. Furthermore, objective measurements resulted in more validity and reliability in the results. The last note the review makes is that the context of the research is important. For example, cycling levels are already higher in some countries than in others making the need for better study designs more necessary to see results. The subsequent recommendation of the review was then to design high quality studies that consider the biases named in this paragraph. Mölenberg et al. (2019) conclude that, to control for all these factors, it might be useful to use existing data in a natural experimental design. However, this is not a recommendation to only use this design. It is here important that this data fits the implementation of the intervention.

What can be concluded is that study design is important to keep in mind when looking at designing a study that investigates an infrastructural intervention. However, the results of all these studies are also important. In this paragraph these results will be discussed along the three categories of interventions.

Links are the first category of cycling infrastructure interventions. In this category a separation can be made between shared cycle paths and separated ones, and other links. Preference studies show that cyclist prefer certain types of infrastructure. In stated preference studies shared cycle paths are less preferred than separated ones. However, in cities with limited bicycle networks, such as the United States, these cycling paths are often still used. On these shared cycling paths, the preference is for roads that have less and slower traffic (Dill & Buehler, 2015). This preference is however less pronounced for more experienced cyclists.

Besides preference for certain types of infrastructure the effect on cycling levels has also been studied. A positive relation has been found between the amount of bike lanes and the amount of cyclist (Dill & Buehler, 2016). However, some studies did not find this relationship. This is in line with what Stappers et. al. (2018) found. Results in this review ranged from an increase in physical activity to a decrease. More positive results were found with small interventions in comparison to total bike network overhauls. As said other factors influenced these results as well. However, these studies looked at the effect on physical activity not the use of infrastructure. Mölenberg et al., (2019) note that measures that are more related to the intervention are more likely to find effect. For example, they found that most studies that investigated infrastructure usage found an increase with a median increase of 62%. Again, as noted before there were various study design elements that influenced these results.

Secondly there are the nodes of the network. These are the intersections (Dill & Buehler, 2016). At these intersections there is potential for more conflict. Preference studies show that cyclist tend to avoid intersections if possible (Dill & Buehler, 2016). Several measures are possible to make intersections better for the (perceived) safety of cyclists, however, there have not been many studies that look effects of intersection treatments on the levels of bicycling.

Lastly there are whole network interventions. Studies have found a correlation between cities with extensive networks and cycling levels. As is with bike lanes networks seem to have a positive effect on cycling levels (Dill & Buehler, 2016). There is a preference for these extensive cycling networks over discontinued bike paths and lanes. Stappers et. al., (2018) note that with extensive interventions on the whole network it is difficult to detect the precise effects. As said cycling highways could also be seen as a network. In the next paragraph the results for this intervention will be discussed.

2.3.3 Results for cycle highways

The research mentioned above is all done on relatively well-known infrastructural interventions such as single bike paths or cycling lanes. Studies that investigate new kinds of cycling infrastructure have not yet been studied exhaustively (Buehler & Dill, 2016). This is also the case for cycling highways. The studies that have been done will be discussed in this paragraph.

Cycling highways started emerging in the modern form since 2004 in the Netherlands and other places (Liu, 2019). With the emergence of cycling highways studies involving cycling highways also started to be published. These studies took various perspectives on cycling highways. From the impact on health, the induced travel demand, physical design perspective or a practitioner's perspective (Liu, 2019). For this research studies that focus on the impact of cycling highways on cycling volumes are more important to focus on. This is because the focus in this thesis is on the effects of cycling highways on cycling volumes. In this category of research three studies are of note.

To begin with the study done by Skov-Petersen et. al (2017). In this study the effect of an upgrade of two cycling paths to cycling highways on bicycling volume and cyclist behavior and experience. To achieve this objective two methods were used. To assess the effects on bicycle volume data from automatic counting stations on the routes themselves was used. This data was collected over 35 months (Skov-Petersen, et. al, 2017). Data from three surveys from before, and one and two years after the intervention was used to analyze the experience of cyclist on these routes. A control survey on a nearby route which had not received an upgrade. The results from the automatic counting stations on the two routes show an increase in bicyclist on both routes. However, from this increase only 4-5% of the new cyclist were from new trips. Most trips were from cyclists coming from other routes. The results from the surveys show an increase in the satisfaction of cyclist on the upgraded routes. This satisfaction is significantly higher than on the control route. From these results the authors of the study conclude that the investments in cycling highways increased the volume of cyclist but mostly due to cyclist coming from other routes (Skov-Petersen, et. al, 2017). The satisfaction on the routes has also increased. This study shows that investments in cycling highways can have effects on the cycling volumes. However, there are more aspects of this study that are important for this research. Firstly, as the authors note themselves the longitudinal study design is important to control for time of day and week and the weather. It is thus important to consider these variables when collecting the data for this research. Besides to control for these effects it is also shows that the effects on cycling volumes takes some time to express themselves and this also increases the reliability of the measurement of bicycling volumes.

The other two studies done on cycling highways are modeling studies. These do not retroactively look at the effects but try to model them beforehand. The first of these two studies have been done in Flanders in Belgium (Buekers, et al, 2015). The objective of this study is to model the cost and benefit of two cycling highways in Flanders. To assess the effect a model was created using two indicators. These are external costs and disability adjusted life years (DALYs) (Buekers, et al, 2015). This last indicator is used as an indicator of healthy life years. Besides these indicators' variables such as the number of cyclists and the costs of the infrastructure are put into the model. Here assumptions are made about the effect of the cycling highways. There are several scenarios about the number of cyclists on the cycling highways. One of the scenarios increase in cyclist is taken from counting done on the highways. However, the problem is that these amounts might not be accurate. For example, there is no control for cyclist coming from other routes. This means that an impact evaluation would be useful to increase the accuracy of the model. The model is then applied to two cycling highways. One between Antwerp and Mechelen and the other between Brussels and Leuven. The results of this were mostly positive (Buekers, et al, 2015). The cycling highways were expected to have a positive effect on the health indicators and external costs. The study therefore concludes that investments in cycling highways have health benefits that are worth these investments.

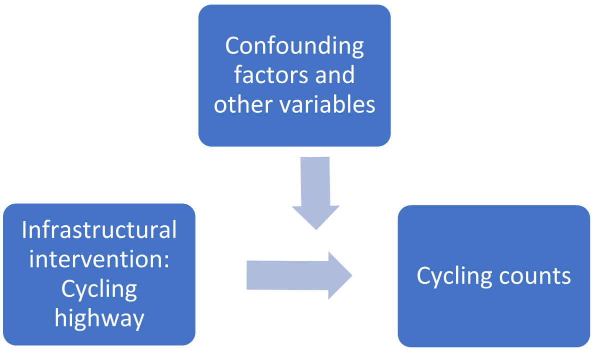
The third and final study done on cycling highways that will be discussed in this section was conducted by Rayaprolu, et al (2018). This study. The study also modeled the effect of cycling highways on commuter mode choice and travel time reduction. To achieve this a discrete choice model based on a German household travel survey was made. Similarly, to the previous study this model was applied to a proposed cycling highway in Munich. The results of this modeling were that the implementation of the bicycling highway would lead to an increase in cycling levels and a decrease in motor traffic. A modal shift was thus expected. The effect of the intervention was greater closer to the intervention than further away. The model however has some limitations. While it does factor in proximity it does not account for differences beyond travel time reductions such as safety or convenience. Secondly this study was a modeling study and not an impact evaluation. This means that the effects of a cycling highway were being modeled ex-ante and not ex-post.

These studies are not in the two reviews that have been discussed in this chapter, however the lessons from these reviews can be used with the literature on cycling highways. Firstly, the study design has influence on the results of the study. The two modeling studies for example do not use empirical data and are not impact studies and are therefore less useful for this study. This study will use counting data from automatic counting stations. The study done in Denmark did use counting stations and is therefore more aligned with this study. What Skov-Peterson et al., (2017) do is well is that they control for the weather and daylight by observing different hours. They also try to control for cyclist coming from other routes. They do this through the questionnaire. However, this is a subjective measure and is therefore less reliable than objective measures. This increases the risk of bias in measurement outcomes. The questionnaires and counting moments were also spread out through time. There were three years of data used. This makes for a more robust study. Lastly, they do not however control for equity effects. This means that there is more risk of bias due to confounding factors.

2.4 Conceptual model

The conceptual model visualizes the predicted effect between the dependent and independent variables. The dependent variable in this research are the cycling counts. The independ variable is the infrastructural intervention of cycling highways. The dependent variable are the cycling counts. This researchs tries to estimate the effect of this independent variable on the dependent variable. However as discussed in the previous paragraphs there are various confounders and variables that can influence the explanatory value of the infrastructural intervention. As mentioned these can be weather, distance to various locations and time. These need to be accounted for with the use of various methods. In chapter 3 the ways in which this can be done will be discussed.

Figure 2 The conceptual model



3. Methodology

In this chapter first the research strategy that will be used will be examined. Here the general way of conducting research and the research methods will be discussed. This includes the way of collecting and analyzing data. After this some paragraphs will be dedicated to the internal and external validity of the research. Finally, the ethics considerations of this research will get some attention.

3.1 Paradigms in evaluation research

Different paradigms exist in the evaluation literature. In this theses four main paradigms in program evaluation are presented (Guyadeen & Seasons, 2018). These paradigms are not exclusive to evaluation theory but rather are part of larger debates that are ongoing within the literature about, for example, how we can perceive reality (Stufflebeam et al., 2000). The first of these paradigms is the post-positivistic approach. In a post-positivist approach, there is thought to be a reality out there. This reality is only imperfectly approachable (Guba & Lincoln, 1994). This approach was the scientific dominant approach in evaluation studies: There is a focus on the scientific research methods with a preferences for quantitative methods and randomized research designs (Guyadeen & Seasons, 2018). This is the case because these methods are the best for trying to prove casual interference

The pragmatic paradigm came as an answer to the overly scientific methods of the post-positivistic approach. The pragmatic paradigm states that the evaluation method should be coupled to the object that is being evaluated (Guyadeen & Seasons, 2018). This discussion between these two discourses is exemplified by the two authors that were proponents of the two paradigms. In 1969 Donald Campbell published an influential article were the post positivist paradigm was applied to evaluation research. This scientific paradigm was challenged by Cronbach (1982). He proposed the pragmatic approach. Evaluations need to be tailored to that what is being evaluated. The evaluation should have maximum utility for decision makers. It might seem that these two paradigms might be compatible. One the hand has very high scientific standards and be totally useful for decision makers. However, there are often not enough time and resources available to do both. On top of that it might not be possible or ethically desirable to do an evaluation that adheres to the highest scientific standards (Rossi et al., 1998).

The third paradigm is the interpretative paradigm in evaluation. This paradigm has been championed by Guba and Lincoln (1989). This paradigm is about the stakeholders and their views on the evaluation. These views are central in this paradigm. The last paradigm is the normative paradigm. Here the stakeholders are encouraged to negotiate and recognizes the respective realities of the other stakeholders. The key value in this type of evaluation is the emancipation and empowerment of stakeholders. The difference between these two paradigms is that in a participatory evaluation the process of negotiation is central. In a sense this goes further than the interpretative paradigm where a diversity of opinions is considered but the discussion between these stakeholders is not as central.

What is clear from these different paradigms is that they have a significant impact on the design of the evaluation research. Choosing a certain paradigm also means that this paradigm needs to be translated into practice. The choice in this thesis is for a post-positivistic paradigm. To further explain this choice three considerations as described by Alkin (2012) for translating evaluation paradigms from theory into practice will be used. Firstly, the issues related to the methodology should be addressed. The choice for impact evaluation implies that an experimental approach is the best suited. This is because an impact assessment tries to prove the effect, or causal relationship, of an intervention and an experimental design has the best chance of proving this relationship (Alkin, 2012). This approach fits well in a post-positivist paradigm. Secondly the way in which data is valued is important. This has to do with the way in which the data is analyzed. A post-positivist paradigm also fits well in this criterium. The data in this thesis is quantitative and a scientific approach to analyzing this data is well suited. Thirdly the audience of the evaluation effort differs for different paradigms. This has to do with the audience for which the evaluation is done. In the case of this thesis the evaluation has two different audiences. Firstly, and most importantly it is a research that has a scientific goal. This research wants to determine the effects of cycling highways in a scientific manner and thus contribute to scientific knowledge. Secondly it has a societal purpose and audience. The results and recommendations that follow from this research can be the basis for new policy surrounding cycling highways. It also contributes to the accountability of existing policy. As stated in the relevance of this thesis there is currently not enough existing ex-post research done on the topic of cycling highways. This thesis can contribute to the knowledge about the effect of the intervention, because a natural experiment is always ex-post research. This means that a post-positivistic paradigm fits well with the scientific goals of the thesis. On top of that can the results be translated for policy uses. When considering the three considerations a postpositivist paradigm fits best.

3.2 Research strategy

The research strategy is the construction in which the data will be collected. This can be done through various ways. To find the best way to collect the data it is useful to know what the ideal type of data is for this research. For an impact evaluation the type of data is needed to make assumptions about the casual interference. For this to happen data relevant to the outcome is needed and relevant variables that might explain differences is needed (Ravallion, 2001). In this case an outcome indicator is needed that is related to the amount of cyclist on the cycling highways and other routes. Secondly variables are needed that explain possible heterogeneity in the data. Besides that, the data needs to be relevant to the outcome the data also needs to made able to control for biases and cofounders. This can be for example due to the data being longitudinal (Rossi et al., 1998).

The data that will be used in this study is count data from the province of Gelderland. This data fits the above description well. The count data is from multiple years before and after the intervention. This means that it is possible to control for biases. Other data that will be used is weather data and population data. These can help indicate possible cofounders.

The nature of this data helps shape the design of the natural experiment. A natural or quasi experiment can take different forms (Rossi et al., 1998). Because the count data contains information from multiple years before, during and after the intervention this means that is possible to design a panel study. Besides this, the assignment of the intervention is not random and not uniform. A panel study can help control for this non-randomness and non-uniformity. A cycling highway is planned deliberately and is therefore not random. There is non-uniformity because not all participants live equally close to the intervention and are therefore not uniformly exposed to the cycling highway. Lastly as Stappers et al. (2018) note that results differ longer after the infrastructural intervention. Therefore, a longitudinal or panel study is interesting. The choice for this type of design does have some disadvantages. Firstly, a randomized controlled trail has more possibility to prove causality, however this is not possible in this case. Secondly a quantitative design might yield less depth in the analysis. This thesis for example will not give answer to the question why cycling highways might be used more or less.

A panel study can be analyzed in various forms. As is clear from chapter 2 the choice in this thesis is to use a difference-in-difference method to analyze the data. In the next paragraph this method will be further explained.

3.3 Data collection

The data collection of this thesis consists of various datasets. The first part of the data in this research will be count data of (fast) cycling routes in the Province of Gelderland. The data was provided by the province of Gelderland through the supervisor of this thesis Dr. Ploegmakers. This count data contains counts from various points along cycling highways from 2011 until 2020. It also contains data about other cycling routes and for some point the type of cycling path. The data is however not uniform and in a poor state. The count data is spread among 14 different cycling highways and 166 different counting locations. The cycling highways were completed in different years. In the table 2 below the years of completion are shown. There are still some cycling highways that will be completed in the future. On top of this many cycling highways were completed in 2015. This is important to keep in mind when analyzing the data. This data is the basis of the analysis.

Table 2 Completion of cycling highways over the years. Source: own analysis

Completion of cycling highways			
Year	Frequency	Percent	Cumulative
2007	21,022	4.61	4.61
2015	107,893	23.64	28.24
2016	36,468	7.99	36.23
2017	86,744	19.00	55.24
2018	7,798	1.71	56.95
2019	5	0.00	56.95
2020	33,329	7.30	64.25
2021	88,285	19.34	83.59
2022	59,432	13.02	96.61
2024	9,423	2.06	98.67
2025	6,057	1.33	100.00
Total	456,47	100.00	

The first step of the analysis was preparing the received datasets. This was necessary because the datasets were varying in format and not useable for analysis. The data for counting stations was available in separate Excel files and this excel files were also separate from the location data (x and y coordinates) for these counting points. The format of the excel files also differed per year and per counting point. There was also no overview of what counting stations contained data for what year. On top of this the names, locations and sometimes dates were not correct. This first step was to index the counting locations with their names in the separate excel files and the years that were available. Secondly the names of the counting stations needed to be matched to the names in the location dataset. After this was done the location data could be matched to the counting data. This process of matching took several iterations as problems and missing locations kept appearing. After the datasets were prepared and formatted the next step of the analysis was the joining of the different datasets. This was also done through importing the datasets into Stata and merging them. After the datasets were joined the analysis could begin. This was done through Stata. The process of the data preparation was not linear. Several iterations of formatting, indexing and merging was needed. Backtracking were errors occurred was also needed.

This data has advantages and disadvantages. The first advantage of this data is the time the data encompasses. The multiple years can give a good insight on the count of bicyclists on cycling highways longer after an intervention. The second advantage of this data is that in contrast to surveys there is no self-selection. Every cyclist is counted and therefore there is a large sample size. The third advantage is that objective data such as count data gives better results. The last advantage is that the count data on other cycling routes can give an indication of what percentage of cyclist on cycling highways can be attributed to induced cycling. The disadvantage of the dataset is that there is no data on 2013 and 2014. Secondly the data is also not complete for the other years. Th is means that for this might skew some of the results.

The counting data is also not the only dataset that was used. To be able to control for various cofounders and biases other data needs to be added to this dataset. The data that is added consist of two parts. The first part is weather data coming from the Royal institute for metrology (KNMI). This data contains the temperature, precipitation and amount of sunlight per hour for the years 2010-2020. To be able to use this dataset some preparation was needed. Firstly, the data needed to be converted from a CSV-file into a Stata dta-file. Secondly the counting points had to be coupled to the closest weather station. The KNMI has numerous weather stations across the Netherlands and the closest station to the counting points was used. This was done through ArcGIS. After the counting points were matched to the closest points the weather data per hour was merged through Stata.

The second part of the data comes from the Netherlands bureau of statistics (CBS). This dataset contains information about demographics and amenities in a 100 by 100 grid. The demographics include income groups and population. The amenities include distance to supermarkets, restaurants, train station, highways and schools. This dataset was also coupled to the dataset with the counting data and the weather data through Stata. The final dataset was used to perform the analysis. In the next paragraph the analysis of this dataset is discussed.

3.4 Data analysis

The research method that was used is the difference-in-difference method. The difference-in-difference method is an ex-post design, as said so are other experimental designs. This means the analysis is done after the intervention has been completed. As said there has not been enough ex-post research in planning (Guyadeen & seasons, 2018). Difference-in-difference is a statistical technique for quasi and natural experiments where there is a need to control for background changes and cofounders (Dimick & Ryan, 2014). With this method two groups are compared. One group is exposed to the policy, while another is not. These groups need to experience the same trends and are compared pre and post exposure. There are two main assumptions with the difference-in-difference method. The first being parallel trends. This is the assumption that the two groups (control and intervention group) experience the same trends/growth before the intervention. If this is the case then the expectation is that without any intervention the two groups would continue to be parallel (Dimick & Ryan, 2014). The second assumption is common shocks: this means that unexpected events affect both groups equally. For example, that the current coronavirus changes travel behavior in both groups the same.

A difference-in-difference method needs to make it plausible that there is casual interference (Wing et al., 2018). With this method the net effect of the intervention needs to be determined. In this case the average treatment effect of the completion of a cycling highway. This is different from the gross effect of the intervention (Rossi et al., 1998). The gross effect is all the difference that is observed from before and after the intervention. The gross effect occurs both in the treatment group and the control group. Between the treatment and control group there is a difference before and a (possibly different) difference after the intervention. The difference between the differences before and after the group is compared in the difference-in-difference method. This is nearly the net effect. However the difference might still contain uncontrolled difference, design effects and stochastic effects (Wing et al., 2018). These need to be subtracted or added to the effect.

The difference-in-difference method can also be written in formula. To build the models that will be used in this research we start with the following formula:

$$\mathrm{Y}_{\mathrm{gt}} = \delta D_{gt} + T_g + P_t + \varepsilon_{gt}$$

In this research the aim is to estimate the treatment effect of the construction cycling highways on bicycle counts. Let Y_{gt} here term represents an outcome. In the case of this research outcome effect of the construction of a cycling highway in bicyclist per hour on a count point. The g and t terms underneath represent the groups and time respectively. In this research the t would be the year from 2010 to 2020. The g represents the intervention and control group. In this research the control group are the paths where cycling highways are planned, and the intervention group is the group were the cycling highways were already completed. An indicator to the formula is added for the treatment group dummy to try and control for selection effects. It could be plausible that cycling routes with more potential were chosen to construct cycling highways. The dummy variable T_g is whether a certain group has received and intervention (the construction of cycling highway) before 2020 and is time-invariant. The P_t term is the group invariant but time varying dummy variable. Next a treatment dummy is added for before and after the intervention. This is D_{gt} term. The δ term before the D is the treatment effect. This is the effect of the intervention. In this research these would be the difference in cyclist before and after the intervention This leaves the ε term. This is the residual term. This is the difference between the observed and the estimated mean. This formula is the basis for the first model of the analysis. However, two more models will be used.

$$\mathrm{Y}_{\mathrm{gt}} = \delta D_{gt} + T_g + x_{gt} + P_t + \varepsilon_{gt}$$

The second model will add variables for inhabitants, weather and distance to locations. This adds the term x_{gt} to the formula. This term encompasses all these variables.

$$\mathrm{Y}_{\mathrm{gt}} = \delta D_{gt} + n_p + P_t + x_{gt} + \varepsilon_{gt}$$

The last model of the analysis will add the fixed effects. The term n_p is added for the fixed effects. This is because it is plausible that the cycling highways are not randomly distributed. Adding cycling highway route based fixed effects means that the term T_g will be dropped because it is constant within the group. With the fixed effects the goal is to control for unobserved attributes that might influence the construction of a cycling highway. The last model will also try to model the effects after several years have passed since the construction of the cycling highway. It could be assumed that the effects of the intervention will rise after some years. To account for this in the formula a term needs to be added for the adjustment effects or lagged effects after the construction of the cycling highway. For this the term $\sum_m^M \delta D_{gt} + m\lambda$ is added. The Σ denotes the sum of the effects over the years. The m term denotes the time after the time in years after the intervention has happened in years. Lastly the M term denotes is the year in which the intervention has happened, if no intervention happened this is zero. There is also the assumption in the model that future interventions are not anticipated and thus do not have an influence on current outcomes. This is also called strict exogeneity null. The final formula looks like the following:

$$\mathrm{Y}_{\mathrm{gt}} = \sum_m^M \delta D_{gt} + m\lambda + n_p + P_t + x_{gt} + \varepsilon_{gt}$$

3.5 Population, intervention and control group

Every experiment contains an population along with an intervention an control group (Dixon, 2018). This is also true for this research. Starting with the population. The population is a collection of individuals or objects with a common trait (Rafeedalie, 2019). In this case these are all the cycling highways in the Netherlands. This thesis tries to answer the question what the effects of the completion of these routes is on cycling levels. Not the whole population will be researched in this thesis. A selection of the population will be used. This is due to the practical constrains such as time and availability of data. The subset that will be used is as said cycling highways in the province of Gelderland. With the use of this dataset control and intervention groups can be created. The intervention group are the counting locations were a cycling highway has been completed before 2021. The control group will be the locations were there will be cycling highways in the future. This is because it can be assumed that these locations have similar observed and unobserved characteristics (Busso et al., 2010). This is because the assumption is that the cycling highways that have been completed probably go through a similar development process as the cycling highways that are to be completed. There will be two different sets of control and intervention groups for comparison. One were the intervention group contains routes that were completed in 2020 and one were the routes were completed in 2019. The 2020 group will mainly be used, the other group is for comparison to see if there are significant differences. Now that the control and intervention group are defined it is interesting to know how many observations each group contains. In table 3 these numbers are presented. What can be seen is that 2013 and 2014 are low in observations. It can be seen that the amount of observation differs per year. Here the 2019 control group can show if the missing of these years makes a difference.

Table 3 Amount of observation per group per year

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Total
Control	802	2,818	802	130	131	14,746	19,617	15,754	19,625	25,381	25,331	125,137
Intervention	13,44	33,768	12,264	0	0	43,379	37,8	23,688	28,525	28,973	29,869	251,706
Near cycling highway	7,392	17,472	7,392	0	0	21,84	22,008	0	1,512	1,694	1,512	80,822
No cycling highway	0	4,704	0	0	0	4,704	5,096	504	504	504	504	16,52
Total	21,634	58,762	20,458	130	131	84,669	84,521	39,946	50,166	56,552	57,216	474

3.6 Validity

The internal validity causal relationship between the independent and depend variable exists. The study design and methods also try to make causal interference as plausible as it can be. With considering all the possible biases and cofounders the casual interference can be assumed to be plausible.

The external validity is the ability to generalize results of this research to the entire population.. However, as Stappers et al. (2018) note that interventions done in countries with high cycling counts might give different results compared to other countries. What this means is that for this research to be externally valid it is important to give a context of the research subject. Through the research design and methods, the external validity has been tried to secure.

3.7 Reliability

Reliability also consists of internal and external reliability. The internal reliability is how reliable the results are. By conducting statistical analysis that is grounded in an appropriate method and considering the significance of the results the internal reliability will try to be upheld. The external reliability is the replicability of a study (van Thiel, 2010). If the data source is open and an extensive description of the analysis can enhance the external replicability. This thesis contains a detailed description of the analysis and the choices for the analysis are explained. The data has been added in the appendix. This means that the external reliability of this thesis has been relatively well secured.

3.8 Ethics

Ethics in research are important to consider. There are several considerations for ethics in research (Barker, Pistrang, Elliott, 2016). Firstly, whether there is harm to participants. In this research no harm to the participants. This is because of the anonymous data collection of count data that does not interrupt the life of the participants. Informed consent is the second considerations. This is not a problem in this research because the data is collected anonymously. Because it is anonymous the third consideration is also not be a problem in this study. The third consideration is namely the privacy of the participants. All in all, the ethics of this research are well managed.

4 Results

4.1 Outcomes – descriptive statistics

The results of this thesis are divided into two parts. The first part is the descriptive statistics and the second part are the statistical analysis. The first part is to get an overview of the dataset and the variables that they contain. The second part is for the analysis. In this part the different models that will be used to analyze the data will be presented.

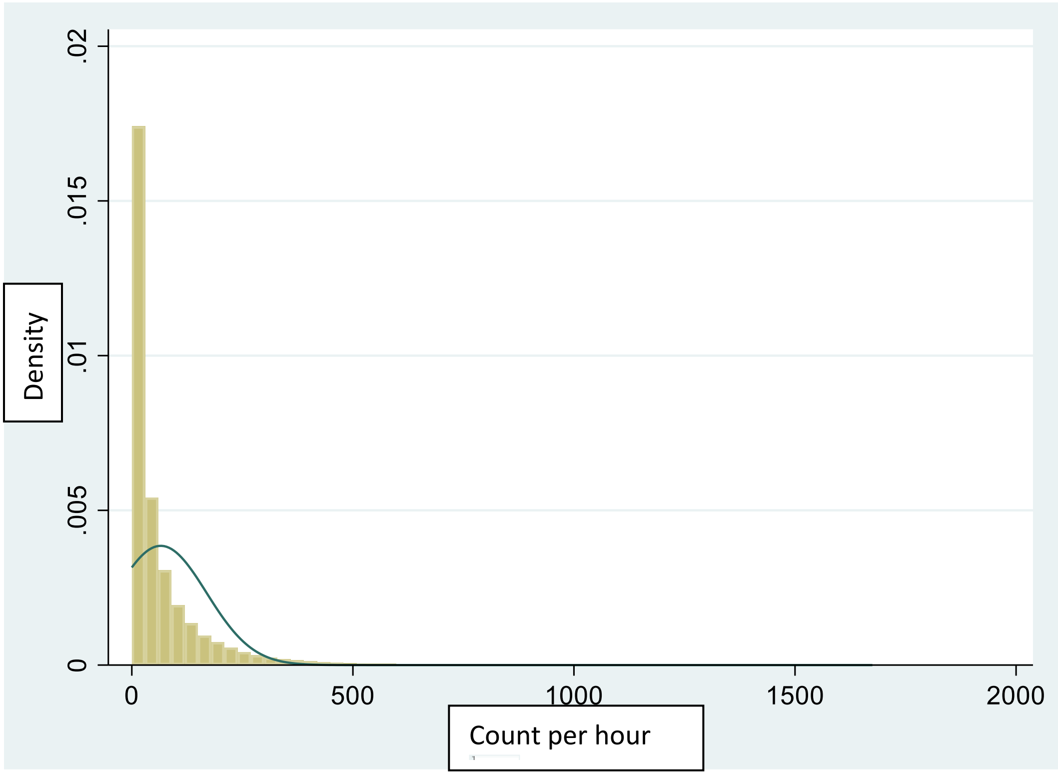
Starting with the descriptive statistics. The most important variable in the dataset is the 'telling' variable. This contains the count data of the counting points along the cycling highways in Gelderland per hour. In total there are 456.391 different observation across 10 years. From 2010 until 2020. The mean of the count per hour is 62,652. With a standard deviation of 97,492. This is important for the analysis. Because if the mean is lower than the standard deviation other statistical analyses are important. There are fourteen different cycling highway routes that have been used. Table 4 shows that the count on these is 456.391. The other counts come from the cycling routes that are not cycling highways. The average count on the cycling highways varies from 20.319 to 132.8 with a total mean of 62.652. As can be read from the table the count per cycling highway varies. From a minimum of 7,567 observations for the Zaltbommel – Den Bosch route to 107.884 for the RijnWaalpad. This means that the RijnWaalpad is a greater percentage of the total observations. This is important to take into account in the analysis. Figure 5 shows a histogram with the variation in the count dataset. All the counts in the dataset are positive and there is a skewness to zero.

Table 4 Descriptive statistics per route

Route	Mean	Standard Deviation	Minimum	Maximum	Amount of observations
Arnhem - Zevenaar (de Liemers)	79,888	117,907	0	1.030	36.463
Arnhem - Nijmegen (RijnWaalpad)	89,803	128,205	0	1.675	107.884
Nijmegen - Wijchen	62,775	105,436	0	1.097	21.021
Nijmegen - Beuningen	66,350	87,675	0	1.397	56.308
Nijmegen Zuid - Beuningen	72,490	95,174	0	1.032	30.436
Zaltbommel - Den Bosch	30,305	41,858	0	360	7.567
Apeldoorn - Deventer (F344)	80,657	103,489	0	673	7.791
Arnhem - Wageningen	34,364	47,167	0	583	49.252
Arnhem - Zevenaar (F12)	20,319	24,272	0	250	40.992
Nijmegen - Mook - Cuijk (MaasWaalpad Gelderland)	45,087	63,468	0	597	25.746
Arnhem - Dieren	57,734	83,106	0	908	47.278
Ede - Wageningen	132,800	14,701	0	1.198	6.055
Deventer - Zutphen (F348)	35,631	46,353	0	281	9.422

Apeldoorn - Epe (F50)	31,177	49,593	0	559	10.176
Total	62,652	97,429	0	1.675	456.391

Figure 3 Histogram of the count data



4.2 Analysis – statistical analysis

The statistical analysis consists of various parts. Different models have been tested and used to estimate the effects and to see which model fits best. The dependent variable in the models is 'Telling'. The independent variables differ per model. There are different models that will be used in this analysis. The first models can be considered a classical difference in difference model. In the second model various variables are added. The last models will add a fixed effects model. The Poisson analysis and a negative binomial regression analysis have been used. These analysis are used because they fit the dataset well. The coefficients in the table should be read as following. If the coefficient is above 0, for example 0.300, then for every one unit increase in the independent variable a 30 percent increase in the depended variable is estimated. When the coefficient is negative This analysis is used because the count data is a random variable that is above 0 and the occurrence of the counts is independent of each other. The difference between the Poisson and negative binomial regression is that the second model is expected to be better for the used dataset. This is a dataset with a great amount of zeroes.. However, this will be tested. Table 4 gives the various models that will be used.

4.2.1 Models

There are two intervention and control groups used in this thesis. Presented in this chapter are mainly the results that use the intervention group were cycling highways were completed in 2020. However, the results were the intervention group contains routes that were completed in 2019 are also included. The models will use both the Poisson analysis and the negative binomial distribution analysis. All models try to estimate the average treatment effect of the completion of cycling highways. However as will be seen the models differ in variables and analysis. The first model can be considered a classic or standard difference-in-difference model. In this model a comparison is made between routes were a cycling highway is completed are compared to routes were no cycling highway is completed . Secondly, the count of cyclists is compared before and after the construction of the cycling highway is analyzed. Lastly, the effect of the years is taken into account. In the second model the difference-in-difference analysis is also used, however, in this model the variables are added. This include firstly the amount of inhabitants within 500 meters of the counting point. Secondly distance in kilometers to the nearest café, restaurant, secondary school, train station, highway and daily supply store is used. On top of this the time of day, weekday and month are used. An independent variable is used for the type of traffic on the cycling highway. This variable is a binary variable. A 0 means that traffic on the route is mixed traffic. A 1 means that the traffic on the cycling route is not mixed. Lastly weather variables are introduced. This include a variable for rain or no rain in the last hour, amount of sunshine (in 0.1 hours) in the last hour, and categories for temperature are used. The last model is a fixed effects model. Fixed effects assume that there are effects within a group that are specific and constant to the group. This is in contrast with a random effects model were differences between groups are assumed to be random stochastic effects. In this model the counting points are grouped according to their cycling highway route. It is thus assumed that the effects in the cycling groups are constant.

Table 5 Model used in the analysis

Model	**Model 1**	**Model 2**	**Model 3**
Method	Classic difference in difference	Difference in difference with variables	fixed effects
Environmental variables	None	inhabitants, distance to locations, mixed traffic, weather, time of day, weekday, month	inhabitants, distance to locations, mixed traffic, weather, time of day, weekday, month
Type of analysis	Poisson and negative binominal regression analysis	Poisson and negative binominal regression analysis	Poisson and negative binominal regression with fixed effects

4.2.2 Results

The results of this thesis are presented in two main tables. Table 5 presents the results using the Poisson analysis and table 6 the results from the negative binomial regression analysis. Tables 6 and 7 show the results using the 2019 intervention group. Lastly figures 2 and 3 show the difference in cyclist in years before and after completion.

Model 1

Firstly the results of the first model will be discussed. The model used the classic difference-in-difference method to analyze the treatment effect of the completion of cycling highways. From column 1 in table 5 the effect of the completion of the cycling highway can be seen. Using this method and the Poisson analysis an 44,6% percent increase in cyclist is estimated when a cycling highway has been constructed compared to before the route was completed. When looking at the first row of the first column the difference between the control and the intervention group is 31,6%. This is the difference in cyclist between the control and the intervention group before construction of the cycling highway. This thus could show that the routes with the most potential were chosen to construct first. The difference when using the negative binomial regression analysis is similar. With a difference of 31,5% before the completion of the intervention and a 44,6% difference after completion. The pseudo R^2 of the Poisson analysis (0.0671) is better than that of the negative binomial analysis (0.0059) and might therefore be the better method, caution is however needed when interpreting this statistic. The pseudo R^2 tries to explain the correlation between the predicted and the observed values.

Model 2

In the second model, variables are added to the model. This are possible cofounders that might influence the count of cyclists. The results of the second model show that adding the variables significantly alters the results. The difference between the intervention and control group here is -1,28%. This shows that the influence of the variables is large when looking at the count of cyclists on a particular route. After the completion of the cycling highway the difference is 66,39%. This means that there is a 66,39% increase in cyclist per hour after the completion of the cycling highway. This is an interesting result. It means that the effect of completion of a cycling highway is greater than expected when using the standard model. When using the binomial regression model, the analysis shows a similar pattern. From a difference of 4,99% to a difference of 57,16%. Again, the pseudo R^2 is better for the Poisson model. The difference with the 2019 intervention group are greater in this model. Using the Poisson model, the difference between the control and intervention group is 25,39%. The completion of a cycling highway in this model equates to a 36.12% increase in the cycling count. In the 2019 model the pseudo R^2 value is also better. When looking at the variables that have been added in this model a significant effect can be seen for a weekday or a weekend day. Counts are similar within the week with a drastic drop for the weekend: In column 3 a 39,65% decrease can be seen for Saturday while Sunday sees a 48,74% decrease. The nighttime also decreases the amount of cyclists with a peak during the morning and evening rush hours. Distance to locations of interest seems to be of limited effect on the count of cyclists. In column 3 for example a one-kilometer increase in distance to a supermarket only decreases the count of cyclists with -1,2%. A small effect is also seen with the number of inhabitants. This could be due to that only a 500-meter proximity to the counting locations is used. In a 500-meter radius within counting points the amount of inhabitants is not high. This means that there a bigger radius might be needed. No mixed traffic does seem to have a significant effect. If there is no mixed traffic an 11,2% increase in cyclist can be observed. Similarly, if there is rain in the hour during the count a 20% decrease in cyclists is seen. Weather on a whole seems to be a significant contributor to the number of cyclists. An increase in temperature increase the amount of cyclists. However, the increase is less when the temperature rises over 30 degrees. This could be due to the heat.

Model 3

The effect of completion of a cycling highway in the third model is 24,65% . This is different from the previous models. The effect is estimated to be smaller in this model. The weather effects in this model are also significant. Other variables have less effect here as well except for the weekend days. On these days there is a significant decrease in cyclists.

The third model includes fixed effects. With a fixed effect model tries to control for time in varying cofounders within the group. This model also includes the increase of cyclist not just after completion but considers a lagging effect. This lagging effect is shown in figure 6 for the Poisson analysis and in figure 7 for the negative binomial regression analysis. The adding of this variable and the fixed effects means that some variables are dropped in this model. As can be seen the mixed traffic variable is dropped in this model because it is the same within the cycling route and therefore the fixed effect group. The fixed effect model drops all the non-time varying predictors from the model because the fixed effects model itself tries to predict this. This is the same for the first row because this again the same in the group. Lastly the variable after completion is dropped because this is now replaced with a dynamic range. Figure 6 and 7 include the years from 6 or more years before completion to 6 or more years after completion. However as can be seen in these figures three years before to one year after completion have been combined. This is due to the dataset. The dataset does not contain counts for the RijnWaalpad the two years leading up to completion of the route. This cycling highway route is a significant portion of the observations and thus skewed the results in these two years. This also the case for the F12 fast cycling route. There is no data for these years because there are no counts for the year 2013 and 2014. With combining these years, the results are assumed to better represent the reality.

The results from this model are interesting. Before completion the number of cyclists is fairly stable before completion of the cycling highway. This indicates that there is no anticipating effect of the cycling highway. Which was an assumption in the model. The year after completion only a small increase in cyclists is seen, while the effect increases the years after the intervention is complete. This indicates that there is a lagging effect when it comes to the intervention. When using the Poisson model, the amount of cyclists per hour on the counting stations peaks at 5 years after the intervention. A 39,81% increase in cyclists can be seen then. However, after this year, and after four years when using the negative binomial regression model, a decrease of cyclist can be seen. This could be due to two factors. The first one is the nature of the dataset. Not many. This could be attributed to the coronavirus pandemic. Because a significant amount of cycling highways was completed in 2014 and 2015 a decrease in 2020 could be because of the pandemic.

Table 6 Results from the Poisson analysis using the 2020 intervention group.

	1. Classic difference in difference Poisson	3. Variables added Poisson	5. Counting points fixed effects Poisson
Cycling highway <=2020	0.3197***	-0.0128***	Same within group
After completion	0.4462***	0.6639***	0.2465***
Year (Ref. 2010)			
2011	0.2303***	0.1616***	0.1627***
2012	0.4219***	0.2127***	0.4219***
2015	0.0201***	-0.3585***	0.3173***
2016	-0.0104***	-0.5122***	0.2296***
2017	-0.0767***	-0.4408***	0.2676***
2018	-0.0005	-0.3985***	0.2563***
2019	0.2130***	-0.1248***	0.3644***
2020	0.2181***	-0.2487***	0.2963***
Month (Ref. may)			
June		0.0592***	0.1246***
August		-0.4038***	-0.2374***
September		-0.1068***	0.01845***
October		-0.2472***	-0.0411***
November		0.0089***	0.1554***
December		-0.1369***	-0.0297***
Weekday (Ref. Monday)			
Tuesday		0.1005***	0.0996***
Wednesday		0.0429***	0.0461***
Thursday		0.0683***	0.0699***
Friday		0.0041***	0.0048***
Saturday		-0.3965***	-0.3945***
Sunday		-0.4874***	-0.4861***
Time (ref. 00:00-01:00)			
01:00		-0.4712***	-0.4716***
02:00		-0.7589***	-0.7578***
03:00		-0.5549***	-0.5538***
04:00		0.3140***	0.0315***
05:00		1.543***	1.5430***
06:00		2.886***	2.8854***
07:00		3.0812***	3.0799***
08:00		2.2501***	2.2482***
09:00		2.1614***	2.1581***
10:00		2.2592***	2.2553***
11:00		2.4473***	2.4429***
12:00		2.5616***	2.5573***
13:00		2.7134***	2.7091***
14:00		2.7666***	2.7620***
15:00		2.8633***	2.8597***
16:00		2.8869***	2.8843***
17:00		2.3390***	2.3982***
18:00		2.0857***	2.0857***

19:00		1.8765***	1.8769***
20:00		1.6209***	1.6209***
21:00		1.3990***	1.3984***
22:00		0.9894***	0.9892***
23:00		0.5344***	0.5340***
Distance to closest			
Supermarket		-0.0125***	-0.0070***
Daily supplies		-0.0049***	-0.0828***
Café		0.0043***	-0.0064***
Restaurant		0.0370***	-0.0585***
Highway		0.0266***	0.0624***
Train station		0.0109***	0.0161***
Secondary education		0.0018***	-0.0069***
Mixed traffic		0.1121***	Same within group
Amount of inhabitants within 500 meters		-0.0001***	-0.0003***
Weather			
Rain in the last hour		-0.2070***	-0.2131***
Amount of **sunshine in the last hour**		0.0117***	0.0120***
Temperature (Ref. <5 degrees C)			
5 – 10		-0.1122***	-0.0879***
10 – 15		0.0178***	0.0344***
15 – 20		0.1379***	0.1614***
20 – 25		0.2944***	0.3155***
25 – 30		0.3837***	0.4088***
> 30		0.2424***	0.3024***
Amount of observations	375,545	375,545	375,545
Pseudo R^2	0.0671	0.5349	

The significance is shown in *,**,*** and stands for 10,5,1 percent respectively.

Table 7 Results from the negative binomial regression analysis using the 2020 intervention group.

	2. Classic DID negative binomial	4. Variables added Negative binomial regression	6. Counting points fixed effects Negative binomial regression
Cycling highway <=2020	0.3151***	0.0499***	Same within group
After completion	0.4789***	0.5716***	0.2843***
Year (ref. 2010)			
2011	0.2015***	0.0993***	0.2700***
2012	0.4223***	0.2014***	0.4842***
2015	-0.0572***	-0.3533***	0.5011***
2016	-0.0751***	-0.0529***	0.5152***
2017	-0.1845***	-0.3966***	0.5740***
2018	-0.1156***	-0.3551***	0.4542***
2019	0.1345***	-0.0503*	0.5781***
2020	0.1659***	-0.2007***	0.6508***
Month (ref. may)			
June		0.0799***	0.0014
August		-0.4330***	-0.1829***
September		-0.2149***	-0.0086
October		-0.4320***	-0.2724***
November		-0.1444***	0.2022***
December		-0.2799***	0.0365***
Weekday (ref. Monday)			
Tuesday		0.1097***	0.1152***
Wednesday		0.0838***	0.0775***
Thursday		0.1385***	0.0895***
Friday		0.1358***	0.0632***
Saturday		-0.0475***	-0.3038***
Sunday		-0.0666***	-0.4314***
Time (ref. 00:00-01:00)			
01:00		-0.4668***	-0.2557***
02:00		-0.7562***	-0.3775***
03:00		-0.5680***	-0.0740***
04:00		0.2920***	0.4679***
05:00		1.4820***	1.1176***
06:00		2.8323***	2.0626***
07:00		3.0404***	2.2746***
08:00		2.2458***	1.7538***
09:00		2.1794***	1.7286***
10:00		2.2832***	1.8049***
11:00		2.4644***	1.9719***
12:00		2.5833***	2.0743***
13:00		2.7421***	2.2257***

14:00		2.7937***	2.2296***
15:00		2.8804***	2.3761***
16:00		2.8809***	2.3529***
17:00		2.3781***	1.9320***
18:00		2.0500***	1.6761***
19:00		1.8350***	1.4823***
20:00		1.5762***	1.2548***
21:00		1.3403***	1.0724***
22:00		0.9264***	0.7463***
23:00		0.5197***	0.3249***
Distance to closest			
Supermarket		-0.0126***	-0.0045
Daily supplies		-0.0004	0.0035***
Café		0.0039***	0.0033***
Restaurant		0.0372***	0.0000
Highway		0.0320***	-0.0009***
Train station		0.0127***	-0.0025***
Secondary education		0.0136***	0.0027***
Mixed traffic		0.3024***	Same within group
Amount of inhabitants within 500 meters		-0.0002***	-0.000***
Weather			
Rain in the last hour		-0.2363***	-0.2206***
Amount of s**unshine in the last hour**		0.0106***	0.0124***
Temperature (ref. <5 degrees C)			
5 – 10		-0.01318	-0.0167**
10 – 15		0.0385***	0.0797***
15 – 20		0.1433***	0.2047***
20 – 25		0.3295***	0.3657***
25 – 30		0.3832***	0.4573***
> 30		0.3538***	0.3883***
Amount of observations	375,545	375,545	375,545
Pseudo R^2	0.0059	0.0850	

The significance is shown in *,**,*** and stands for 10,5,1 percent respectively.

Table 8 Difference over the years

2020 Poisson regression change over the years in %		**2020 Negative binomial regression change over the years in %**
Years	**Difference**	
-6 or more	-1.9135***	-2.8174***
-5	-4.4471***	-0.6061***
-4	-3.0422***	-6.0912***
-3 to -1	-3.5836***	-1.3564***
0	0	0
1	0.2157***	0.2979***
2	13.7126***	14.2444***
3	22.6611***	23.4898***
4	31.1490***	29.1022***
5	39.6812***	15.1772***
6 or more	28.9386***	8.8864***

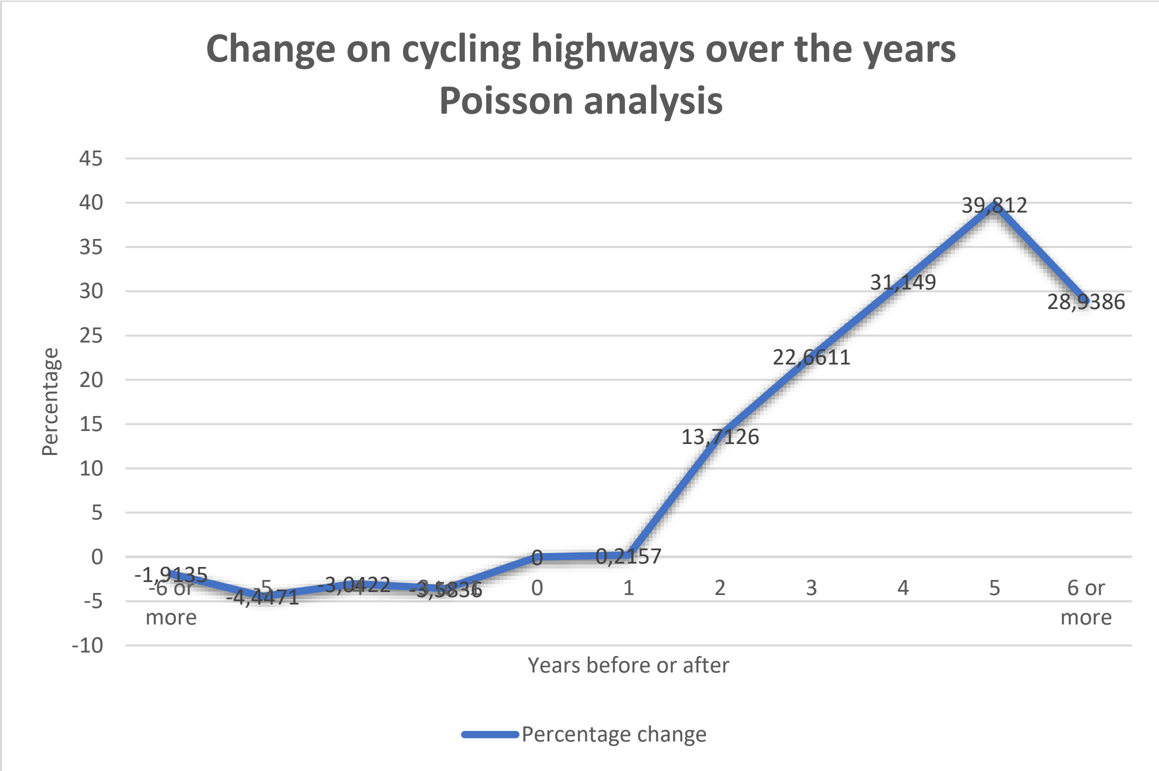


Figure 4 Change in cyclists on cycling highways years before and after completion using the Poisson analysis.

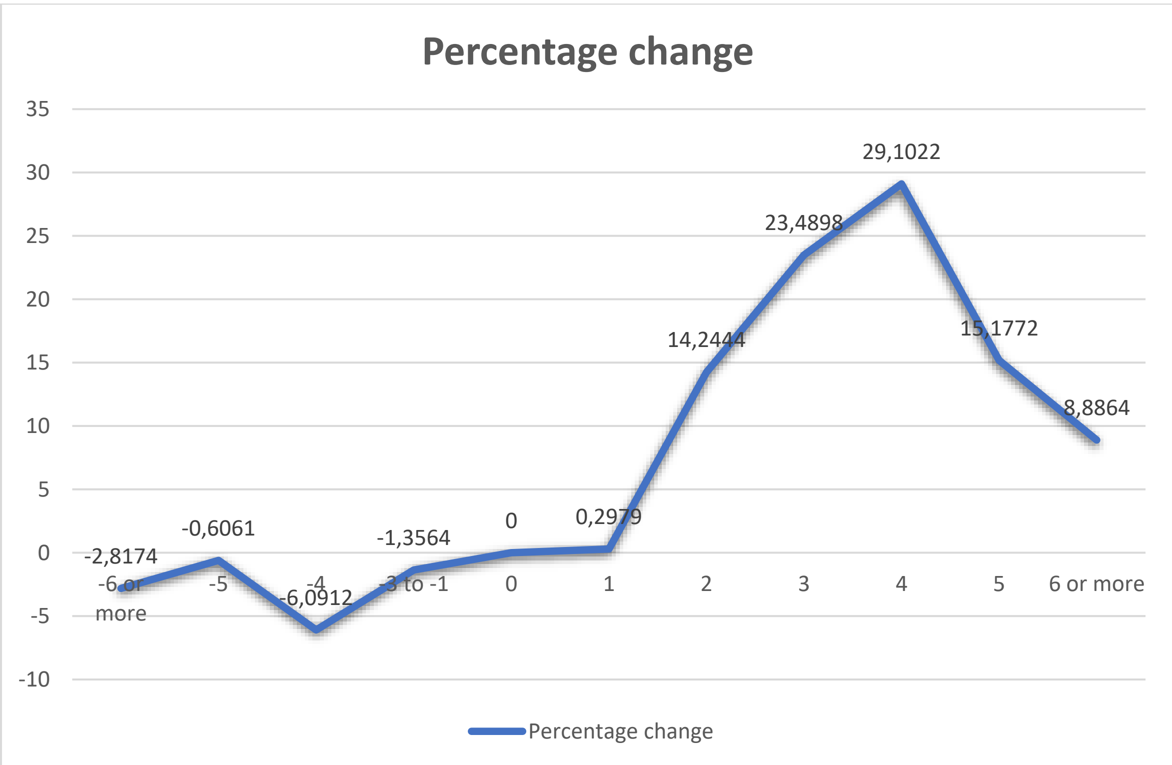


Figure 5 Change in cyclists on cycling highways years before and after completion using the negative binomial regression analysis.

5 Conclusion

5.1 Conclusion

The conclusion of this thesis is threefold. Starting with the effect of cycling highways. The goal of this research was to try and answer the question if the construction of cycling highways can influence the amount of cyclist. The answer to this questions the models in this research show that a cycling highway influences the number of cyclists and that the number of cyclists increases after the completion of a cycling highway. The methods and analysis used make this statement plausible. The significance levels in the analysis also support the claim. The most complex model used shows that cycling highways might increase cycling levels 39,8% five years after completion.

This research is one of the few studies that has shown that has empirically shown the effect of cycling highways. The results from this research can be compared to other research in the field. Skov-Peterson et.al. (2017) found in their study on cycling highways an increase between 6% to 71% in cycling counts on weekdays. When accounting for induced cycling they found an increase between 4% and 6%. Positive effects are thus also found in this study. Other studies on cycling infrastructure that also use cycling counts also estimate that infrastructural intervention have positive effects. Three other studies using count data also found positive effects.. Firstly Hong et al. (2020) evaluated large infrastructure interventions in Glasgow and found that three infrastructural interventions found an increase between 12% and 18% after the completion of the cycling infrastructure. Secondly Heesch et al. (2016) found an 69% increase in cyclist after opening the last stage of the new infrastructure. Thirdly Rissel et al., (2015). In this study two bike counting stations were used to measure the effect of a new bikeway. They found an increase of 23% and 97%. However the self-reported cycling frequency of participants did not increase. The expectation is that most cyclist therefore do not come from induced cycling but other routes. All the four researches thus found an increase in cycling counts after the completion of cycling infrastructure. However the increase differs per research. This might be due to the difference in methods used.

This research shows the ex-post the effect of cycling highways and is therefore interesting both scientifically and for society. Scientifically it can contribute to the knowledge about cycling highways and their effects. Societally this study can contribute to the existing knowledge on cycling highways. It can help answer questions about effectiveness of the policy measures. Policy makers could use this study to evaluate current cycling highways and possibly inform future decisions on cycling highways.

However, some uncertainties still exist. Firstly, the volume of increase is still somewhat uncertain. Unobserved and unknown cofounders might explain some difference in cycling counts. While the model uses various variables and the fixed effect model to try to compensate for cofounders some time varying cofounders might still exist. As said fixed effects model capture time invarying cofounders but not time varying. For example changes in promotion of the cycling highways or investments in other transportation modes were not taken into account. When a highway or public transit line is constructed along a route were a cycling highway is constructed some people may choose to use this other transport option. Further improving on this research both in methodological approaches, dataset and considering possible cofounders might result in different outcomes. On top of this the cycling highways were all situated in the province of Gelderland in the Netherlands. Generalizing results to the Netherlands might be possible but generalizing results to other countries is difficult. In the Netherlands levels of cycling and other factors are similar. Weather conditions and topography for example are similar across the country. Cofounders, biases and other variables are different in other countries. For example, overall cycling levels are relatively high in the Netherlands already and this might not be the case in other countries. This means that the effect of large infrastructural interventions influence the amount of cyclists in different ways.

Lastly an ex-post evaluation design can be used to research the effectiveness of cycling highways. As said in the relevance of this study not enough ex-post research is done on cycling infrastructure in general and cycling highways specifically. This means that future research could use a similar design and improve upon the design to effectively evaluate cycling highways and possibly other cycling interventions.

5.2 Recommendations

Three main recommendation can be drawn from this research and its conclusion. Firstly, more research is needed. This research focused on a limited set of cycling highways in one province of the Netherlands. This makes generalizability difficult. To be able to make conclusion about cycling highways in general more research in other provinces and countries should be done. On top of this more data would increase the strength of studies in the future. Some effects of the intvention could be better understood if more data is available. This can be both on more routes and more years. Future studies should thus try to include as many years and routes as is feasible in their study. However the strenght of this study is not without merit. Various variables on weather and demographics have been taken into account. The methods used also contribute to the statistical significance of the results.

Secondly future research could improve upon this research design. As this is one of the few studies that uses an ex-post evaluation design on cycling highways improvements are possible. Further research could repeat this research and try to replicate these results. Further research could also try to replicate the results in other countries as the context is different in other countries.

The third recommendation is to use more ex-post research when evaluating cycling highways. This recommendation is important for the societal relevavnce of this research. As this research has shown this can add value and increase knowledge about the effectiveness of interventions. This is particularly interesting for policy makers for multiple reasons. Firstly, it increases the accountability of policy. These cycling highways, as is most infrastructure, is funded with public money and the intention is mostly to use this money effectively. Ex-post research can be the basis for an effective cost-benefit analysis of built infrastructure. On top of this many cycling highways are still being planned. Decisions on these cycling highways could be better informed with more ex-post research on existing cycling highways.

6 Discussion

The last part of this thesis is the discussion on this research. Starting with the discussion on the methodology. The methodology used in this thesis is as said fairly new when it comes to evaluating cycling highways. This means that there might be some concepts that were applied that might not have been applied without error. Besides this other analysis methods might have been better to use. The second point of discussion is the dataset. It has already been said that the dataset was not perfect. More variables on demographics could be added or improved. Such as the 500-meter used for inhabitants. The dataset also did not contain data for every counting point for every year. This introduces noise in the analysis as was seen with the missing data on the RijnWaalpad and the drop after five years. A more complete dataset could have resulted in more accurate estimates of the effects of cycling highways. Secondly more routes in the Netherlands might also have contributed to more generalizability. Ex-post research

References

Aittasalo M, Tiilikainen J, Tokola K, Suni J, Sievänen H, Vähä-Ypyä H, et al. Socio-ecological natural experiment with randomized controlled trial to promote active commuting to work: process evaluation, behavioral impacts, and changes in the use and quality of walking and cycling paths. Int J Environ res Public Health. 2019;16(9):1

Alkin, M., (2012). Comparing Evaluation Points of View. *Evaluation Roots*, 4–11. https://doi.org/10.4135/9781412984157.n1

Alkin & Rossi (2012). *My views of evaluation and their origins. Evaluation Roots.* https://doi.org/10.4135/9781412984157.n1

Barker, C., Pistrang, N., & Elliott, R. (2016). Research methods in clinical psychology : an introduction for students and practitioners (Third). John Wiley and Sons.

Buehler, R., & Dill, J. (2016). Bikeway Networks: A Review of Effects on Cycling. *Transport Reviews*, *36*(1), 9–27. https://doi.org/10.1080/01441647.2015.1069908

Buehler, R., & Pucher, J. (2012). Cycling to work in 90 large American cities: New evidence on the role of bike paths and lanes. Transportation, 39(2), 409–432. https://doi.org/10.1007/s11116-011-9355-8

Buekers, J., Dons, E., Elen, B., & Panis, L. I. (2015). Health impact model for modal shift from car use to cycling or walking in Flanders: application to two bicycle highways. Journal of Transport & Health, 2(4), 549-562.

Busso, M., Gregory, J., & Kline, P. M. (2010). ASSESSING THE INCIDENCE AND EFFICIENCY OF A PROMINENT PLACE BASED POLICY. *National burea of economic research*, *53*(9), 1689–1699.

CBS, 2018. Data on cycling in the Netherlands. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84710ned/table?ts=1612528416706

Crane, M., Rissel, C., Standen C., Ellison, A., Ellison, R., Wen, L. M., Greaves, S., Longitudinal evaluation of travel and health outcomes in relation to new bicycle infrastructure, Sydney, Australia. Journal of Transport & Health, Volume 6, 2017, Pages 386-395, ISSN 2214-1405, https://doi.org/10.1016/j.jth.2017.07.002.

Craig P, Katikireddi SV, Leyland A, Popham F. 2017. Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. Annu. Rev. Public Health 38:39–56

Cronbach, L. J., & Shapiro, K. (1982). Designing evaluations of educational and social programs. Jossey-Bass,.

de Hartog, J. J., Boogaard, H., Nijland, H., & Hoek, G. (2010). Do the health benefits of cycling outweigh the risks? *Environmental Health Perspectives*, *118*(8), 1109–1116. https://doi.org/10.1289/ehp.0901747

Decisio (2012). *MKBA van de fiets.* https://decisio.nl/wp-content/uploads/MKBA-Fiets.pdf

Decisio (2017). *Waarderingskengetallen MKBA Fiets.* https://www.mkba-informatie.nl/mkba-voor-gevorderden/richtlijnen/waarderingskengetallen-mkba-fiets-state-art/

Dill J, McNeil N, Broach J, Ma L. Bicycle boulevards and changes in physical activity and active transportation: findings from a natural experiment. Prev Med. 2014;69(S):S74–S

Dimick, J. B., & Ryan, A. M. (2014). Methods for evaluating changes in health care policy: the difference-in-differences approach. Jama, 312(22), 2401-2402.

Dixon, B. S. (2018). Difference in difference method. *Journal of the American Medical Association*, *312*(21), 2401–2402. https://doi.org/10.1001/jama.2014

Fietsersbond. (2019). *Fietsambities van de provinciebesturen 2019 – 2023*.

Fishman, E., Schepers, P., & Kamphuis, C. B. M. (2015). Dutch cycling: quantifying the health and related economic benefits. American journal of public health, 105(8), e13-e15.

Gemeente Utrecht (2015). *Actieplan Utrecht fiets.*

Gemeente Enschede (2020). *Enschede Fietsstad.*

Gemeente Nijmegen (2018). *Ambitiedocument Mobiliteit.*

Garrard, J., Rissel, C., & Bauman, A. (2012). Health benefits of cycling. City cycling, 31, 55

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. Handbook of qualitative research.

Guyadeen, D., & Seasons, M. (2018). Evaluation Theory and Practice: Comparing Program Evaluation and Evaluation in Planning. *Journal of Planning Education and Research*, *38*(1), 98–110. https://doi.org/10.1177/0739456X16675930

Handy, S., van Wee, B., & Kroesen, M. (2014). Promoting Cycling for Transport: Research Needs and Challenges. *Transport Reviews*, *34*(1), 4–24. https://doi.org/10.1080/01441647.2013.860204

Hanemaayer, D. (2012). A “missing” family of classical orthogonal polynomials. In *Ex ante evaluatie in Nederland*. https://doi.org/10.1088/1751-8113/44/8/085201

Heesch, K. C., James, B., Washington, T. L., Zuniga, K., & Burke, M. (2016). Evaluation of the Veloway 1: A natural experiment of new bicycle infrastructure in Brisbane, Australia. *Journal of Transport and Health*, *3*(3), 366–376. https://doi.org/10.1016/j.jth.2016.06.006

Heinen, E., VanWee, B., & Maat, K. (2010). Commuting by bicycle: An overview of the literature. Trans- port Reviews, 30(1), 59–96.

Hong, J., McArthur, D.P. & Livingston, M. The evaluation of large cycling infrastructure investments in Glasgow using crowdsourced cycle data. Transportation 47, 2859–2872 (2020). https://doi.org/10.1007/s11116-019-09988-4

Klobucar, M. S., & Fricker, J. D. (2007). Network evaluation tool to improve real and perceived bicycle safety. *Transportation Research Record*, *2031*, 25–33. https://doi.org/10.3141/2031-04

Krizek, K. J., Barnes, G., & Thompson, K. (2009). Analyzing the effect of bicycle facilities on commute mode share over time. Journal of urban planning and development, 135(2), 66-73.

Matthews, J. N. (2006). Introduction to randomized controlled clinical trials. CRC Press.

Mölenberg, F. J. M., Panter, J., Burdorf, A., & Van Lenthe, F. J. (2019). A systematic review of the effect of infrastructural interventions to promote cycling: Strengthening causal inference from observational data. *International Journal of Behavioral Nutrition and Physical Activity*, *16*(1). https://doi.org/10.1186/s12966-019-0850-1

Oja, P., Titze, S., Bauman, A., De Geus, B., Krenn, P., Reger-Nash, B., & Kohlberger, T. (2011). Health benefits of cycling: a systematic review. Scandinavian journal of medicine & science in sports, 21(4), 496-509.

Panter J, Heinen E, Mackett R, Ogilvie D (2016). Impact of new transport infrastructure on walking, cycling, and physical activity. Am J Prev Med.

Provincie Brabant (2009). *Fiets in de versnelling.*

Provincie Gelderland. (2018). *Koersdocument Duurzame Mobiliteit*.

Provincie Utrecht (2019). *Uitvoeringsplan fiets.*

Pucher, J., Dill, J., & Handy, S. (2010). Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine*, *50*(SUPPL.), S106–S125. https://doi.org/10.1016/j.ypmed.2009.07.028

Rafeedalie, E. (2019). Research: Population and Sample.

Ravallion, M. (2001). The mystery of the vanishing benefits: An introduction to impact evaluation. *World Bank Economic Review*, *15*(1), 115–140. https://doi.org/10.1093/wber/15.1.115

Rayaprolu, H. S., Llorca, C., & Moeckel, R. (2018). Impact of bicycle highways on commuter mode choice: A scenario analysis. *Environment and Planning B: Urban Analytics and City Science*. https://doi.org/10.1177/2399808318797334

Rissel, C., Greaves, S., Wen, L.M. et al. Use of and short-term impacts of new cycling infrastructure in inner-Sydney, Australia: a quasi-experimental design. Int J Behav Nutr Phys Act 12, 129 (2015). https://doi.org/10.1186/s12966-015-0294-1

Rijksoverheid, 2018*. Kabinet: Meer mensen op de fiets.* https://www.rijksoverheid.nl/onderwerpen/fiets/fietsbeleid

Rossi, P. H., Freeman, H., & Lipsey, M. (1998). *Evaluation A systematic approach*.

Skov-Petersen, H., Jacobsen, J. B., Vedel, S. E., Thomas Alexander, S. N., & Rask, S. (2017). Effects of upgrading to cycle highways - An analysis of demand induction, use patterns and satisfaction before and after. *Journal of Transport Geography*, *64*(December 2016), 203–210. https://doi.org/10.1016/j.jtrangeo.2017.09.011

Stappers, N. E. H., Van Kann, D. H. H., Ettema, D., De Vries, N. K., & Kremers, S. P. J. (2018). The effect of infrastructural changes in the built environment on physical activity, active transportation and sedentary behavior – A systematic review. *Health and Place*, *53*(December 2017), 135–149. https://doi.org/10.1016/j.healthplace.2018.08.002

Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2000). *Evaluation Models Viewpoints on educational and human services evaluation*.

Thiemann-Linden, J., & Boeckhout, S. Van. (2012). Cycle Highways. Cycling Expertise, I(12), 3–6.

van Esch, M., Bot, W., Goedhart, W., & Scheres, E. (2017). *Een toekomstagenda voor snelfietsroutes*.

Thiel, S. V. (2010). Bestuurskundig onderzoek: een methodologische inleiding. Bussum: Coutinho.

Thoemmes, F. J., & West, S. G. (2011). The use of propensity scores for nonrandomized designs with clustered data. Multivariate Behavioral Research, 46(3), 514-543.

Wing, C., Simon, K., & Bello-Gomez, R. A. (2018). Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. *Annual Review of Public Health*, *39*, 453–469. https://doi.org/10.1146/annurev-publhealth-040617-013507

Appendix A: Results from 2020

2020 analysis comparing between cycling routes and nearby routes	Classic DID	Variables added Poisson	Variables added Negative binomial regression
Nearby route	-0,0398***	-0,0508***	-0,2741***
After completion	0,5691***	0,6333***	0,7045***
Years effect	Yes	Yes	Yes
Months		Yes	Yes
Day of the week		Yes	Yes
Time of day		Yes	Yes
Weather		Yes	Yes
Locations		Yes	Yes
Observations	375,545	375,545	375,545
Pseudo R^2			

Appendix B: Results from 2019

2019 Poisson regression change over the years in %	
Years	Difference
-6 or more	-2.3793***
-5	-4.7589***
-4	-3.3580***
-3 to -1	-3.7428***
0	0
1	0.2072***
2	13.7914***
3	22.8740***
4	31.7749***
5	39.9204***
6 or more	30.6387***

2019 Negative binomial regression change over the years in %	
Years	Difference
-6 or more	2.5817
-5	2.1711***
-4	1.3795***
-3 to -1	1.3833***
0	0
1	3.7890***
2	16.8327***
3	21.0908***
4	33.7710***
5	7.2606***
6 or more	24.9065***

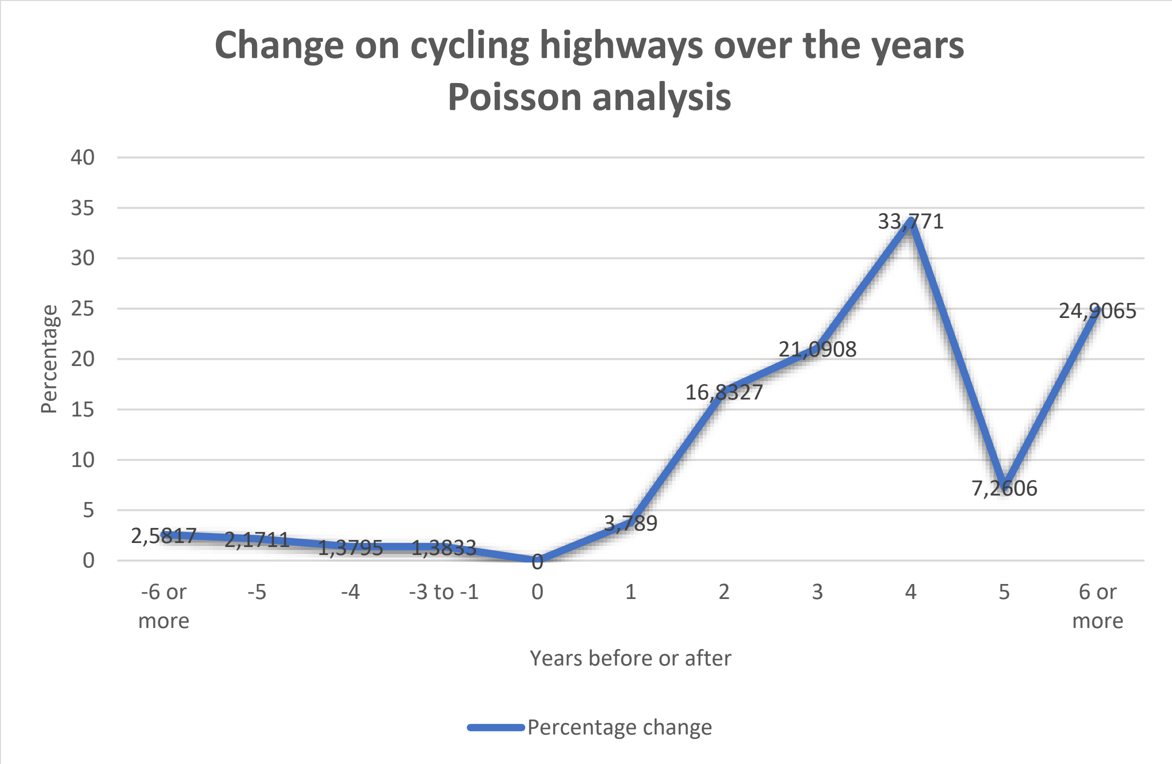
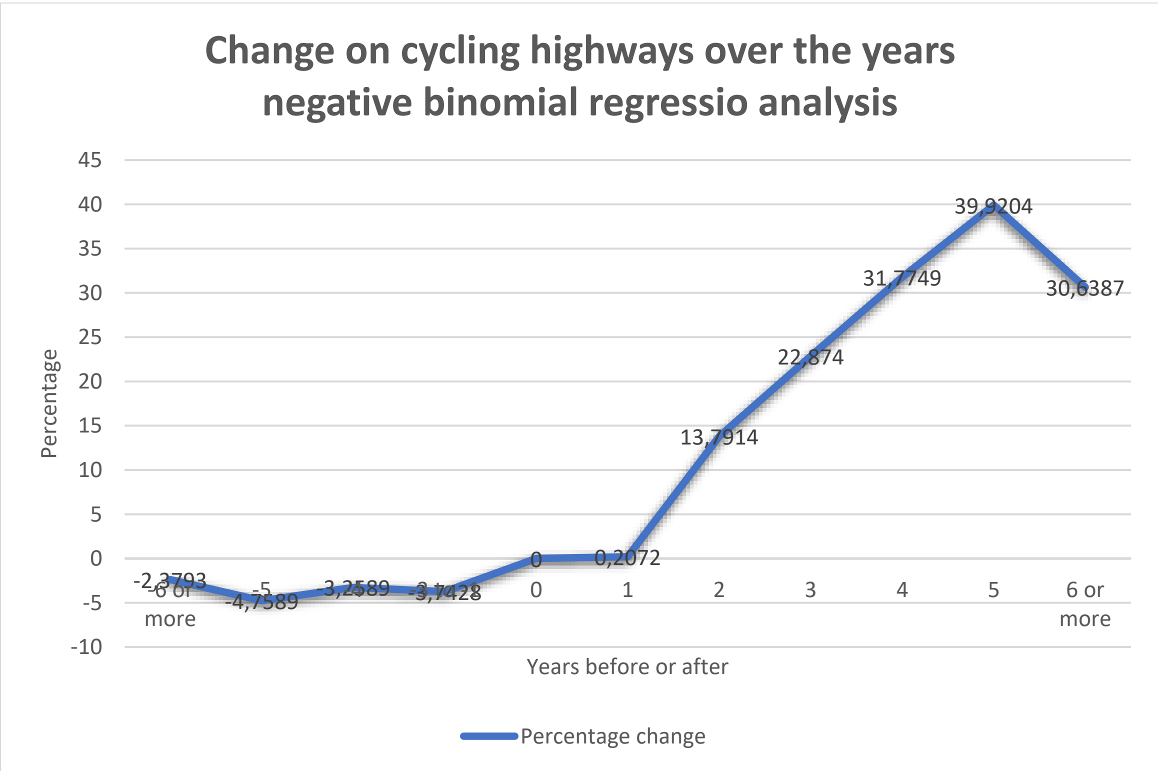


*Table 9 Results from the negative binomial regression analysis using the 2019 intervention group. Source: Own analysis The significance is shown in *,**,*** and stands for 10,5,1 percent respectively.*

2019 Negative binomial regression	2. Classic DID	4. Variables added	6. Counting points fixed effects
Cycling highway <=2020	0.452***	0.3708***	
After completion	0.3455***	0.3271***	See appendix A
Years effect	Yes	Yes	Yes
Months		Yes	Yes
Day of the week		Yes	Yes
Time of day		Yes	Yes
Weather		Yes	Yes
Locations		Yes	Yes
Observations	375,545	375,545	375,545
Pseudo R^2	0.0063	0.0865	

*Table 10 Results from the Poisson analysis using the 2019 intervention group. Source: Own analysis The significance is shown in *,**,*** and stands for 10,5,1 percent respectively.*

2019 Poisson	2. Classic DID	4. Variables added	6. Counting points fixed effects
Cycling highway <=2020	0.4896***	0.2539***	
After completion	0.2942***	0.3612***	See appendix A
Years effect	Yes	Yes	Yes
Months		Yes	Yes
Day of the week		Yes	Yes
Time of day		Yes	Yes
Weather		Yes	Yes
Locations		Yes	Yes
Observations	375,545	375,545	375,545
Pseudo R^2	0.0735	0.5353	



	1. Classic difference in difference Poisson	3. Variables added Poisson	5. Counting points fixed effects Poisson
Cycling highway <=2019	0,4895***	0,2539***	Same within group
After completion	0,2942***	0,3612***	See graph
Year compared to 2010			
2011	0.1927***	0,1323***	0.1574***
2012	0,4197***	0,1908***	0.3889***
2015	0,0640***	-0,2257***	0.3341***
2016	0,0549***	-0,3284***	0.3085***
2017	0,0038***	-0,2788***	0.2674***
2018	0,0950***	-0,2082***	0.2544***
2019	0,3096***	0,0623***	0.3882***
2020	0,3319***	-0,0434***	0.3166***
Month (compared to may)			
June		0,07212***	
August		-0,3862***	
September		-0,0601***	
October		-0,2112***	
November		0,0766***	
December		-0,1454***	
Weekday (compared to Monday)			
Tuesday		0,0986***	0.0123***
Wednesday		0,0334***	-0.0293***
Thursday		0,0601***	0.0147***
Friday		-0,0121***	-0.0389***
Saturday		-0,4117***	0.2595***
Sunday		-0,4913***	0.0155***
Time (compared to 00:00-01:00)			
01:00		-0,4685***	-0.4665***
02:00		-0,7523***	-0.7502***
03:00		-0,5471***	-0.5451***
04:00		0,3128***	0.3125***
05:00		1,5350***	1.5374***
06:00		2,8675***	2.8678***
07:00		3,0621***	3.0544***
08:00		2,2354***	2.2234***
09:00		2,1438***	2.1288***
10:00		2,2434***	2.2250***
11:00		2,4294***	2.4057***
12:00		2,5448***	2.5178***
13:00		2,6943***	2.6685***
14:00		2,7489***	2.7264***

15:00		2,8505***	2.8305***
16:00		2,8778***	2.8566***
17:00		2,3885***	2.3759***
18:00		2,0800***	2.0728***
19:00		1,8725***	1.8663***
20:00		1,6179***	1.6130***
21:00		1,3921***	1.3883***
22:00		0,9852***	0.9843***
23:00		0,5313***	0.5314***
Distance to closest			
Supermarket		-0,0122***	-0.0050***
Daily supplies		-0,0003**	-0.0850***
Café		0,0040***	-0.0089***
Restaurant		0,0346***	-0.0620***
Highway		0,0246***	0.0650***
Train station		0,0010***	0.0170***
Secondary education		0,0034***	-0.0070***
Mixed traffic		0,1083***	Same within group
Amount of inhabitants within 500 meters		-0,0001***	-0,0003***
Weather			
Rain in the last hour		-0,2043***	-0,2391***
Amount of **sunshine in the last hour**		0,0154***	0,0143***
Temperature		0,1138***	0.0370***
Amount of observations	375,545	375,545	375,545
Pseudo R^2	0,0735	0,5353	

	1. Classic difference in difference Negative binomial regression	**3. Variables added Negative binomial regression**	**5. Counting points fixed effects Negative binomial regression**
Cycling highway <=2019	0,4522***	0,3708***	
After completion	0,3453***	0,3271***	See graph
Year compared to 2010			
2011	0,1641***	0,0918***	0.1509***
2012	0,4225***	0,2397***	0.2673***

2015	-0,0355***	-0,2678***	0.3511***
2016	-0,0289***	-0,3691***	0.3141***
2017	-0,1253***	-0,2822***	0.3861***
2018	-0,0439***	-0,2208***	0.3278***
2019	0,2030***	0,0868***	0.4705***
2020	0,2353***	-0,0511*	0.4938***
Month (compared to may)			
June		0,0897***	0.0043***
August		-0,3674***	-0.1794***
September		-0,1270***	-0.0171***
October		-0,3449***	-0.2713***
November		-0,0624***	0.2173***
December		-0,2319***	-0.0056***
Weekday (compared to Monday)			
Tuesday		0,1096***	0.1122***
Wednesday		0,0819***	0.0592***
Thursday		0,1293***	0.0914***
Friday		0,1241***	0.0638***
Saturday		-0,0574***	-0.2965***
Sunday		-0,0752***	-0.4394***
Time (compared to 00:00-01:00)			
01:00		-0,4568***	-0.2430***
02:00		-0,7503***	-0.3762***
03:00		-0,5598***	-0.0731***
04:00		0,2947***	0.4816***
05:00		1,4716***	1.2030***
06:00		2,8159***	2.0921***
07:00		3,0283***	2.2554***
08:00		2,2378***	1.7312***
09:00		2,1762***	1.6942***
10:00		2,2863***	1.7812***
11:00		2,4693***	1.9274***
12:00		2,5899***	2.0338***
13:00		2,7497***	2.1993***
14:00		2,8004***	2.2409***
15:00		2,8857***	2.3327***
16:00		2,8799***	2.3157***
17:00		2,3724***	1.9077***
18:00		2,0436***	1.6613***
19:00		1,8273***	1.4788***
20:00		1,5718***	1.2626***
21:00		1,5718***	1.0723***
22:00		1,3359***	0.7511***
23:00		0,9299***	0.3344***
Distance to closest			

Supermarket		-0,0134***	-0.0019***
Daily supplies		0,0050***	-0.0019***
Café		0,0029***	-0.0046***
Restaurant		0,0353***	-0.0113***
Highway		0,0300***	0.0030***
Train station		0,0126***	0.0043***
Secondary education		0,0163***	0.0031***
Mixed traffic		0,3114***	Same within group
Amount of inhabitants within 500 meters		-0,0002***	-0.0001***
Weather			
Rain in the last hour		-0,2522***	-0.0049***
Amount of s**unshine in the last hour**		0,0130***	0.0169***
Temperature		0,0909***	0.0293***
Amount of observations	375,545	375,545	375,545
Pseudo R^2	0,0063	0,0865	