# Rail Ridership Response to Planned Highway Closures

## a Revealed-Preference Analysis of the Dutch Network

**Hugo von der Thüsen**

Master's Thesis - TU Delft - July 2026

Hugo von der Thüsen (4862600): *Rail Ridership Response to Planned Highway Closures: A Revealed-Preference Analysis of the Dutch Network,* Master's Thesis, © July 2026

Master's thesis to obtain the degree of Master of Science in Transport, Infrastructure and Logistics at Delft University of Technology. To be defended publicly on 15 July 2026.

SUPERVISORS:

| | |
|---|---|
| Dr.Ir. N. (Niels) van Oort (*Chair*) | TU Delft |
| Dr.Ir. M. (Maarten) Kroesen | TU Delft |
| M. (Mahsa) Movaghar | TU Delft |
| M. (Menno) de Bruyn | NS |

# SUMMARY

## INTRODUCTION

The Dutch highway network is ageing. A large share of the infrastructure was built between the 1960s and 1980s and is now deteriorating and nearing the end of its design life. Combined with a Dutch policy preference for preventative maintenance, this translates into a steady and predictable stream of planned highway closures.

The Netherlands operates one of the densest intercity rail networks in Europe, and on a subset of well-served corridors (e.g. Amsterdam–Utrecht) rail travel times are already competitive with the car. When a highway is closed and car travel requires a detour, the relative attractiveness of rail increases further. A share of travellers may therefore be expected to switch modes during a closure. NS, the main Dutch railway operator, currently has no reliable method to predict how many extra passengers they can expect during a planned road closure. This makes it difficult to plan for the necessary capacity and risks trains becoming overcrowded or resources being allocated unnecessarily.

The academic literature offers limited research for this question. Travel behaviour under disrupted conditions has been widely studied, but the majority of this work focuses on disruptions within a single mode rather than switches between modes. The few studies that do examine modal shift during road closures rely on stated preference surveys and only focus on single instances of major road closures in an urban context. Large-scale revealed preference research linking road closures to measurable changes in rail ridership is absent from the literature. The Dutch context, with its combination of full-scale road closures and a competitive intercity rail network, is particularly well-suited to studying this phenomenon.

This thesis addresses the gap with two contributions. First, a dataset of full-day highway closures in the Netherlands is constructed from publicly accessible traffic intensity data as no reliable dataset exists. Second, a transferable method was developed and applied, combining closure detection from traffic intensity data, a treatment intensity measure per OD-pair, and a fixed-effects panel design, to causally identify the ridership response using NS OV-Chipkaart data. The method is specified independently of the Dutch datasets that implement it and can be applied in any setting that meets the data requirements set out in this thesis. The central research question is: *what is the causal effect of planned road closures on rail ridership in the Netherlands?*

No reliable archived dataset exists of road closures in the Netherlands, therefore one had to be constructed for usage in this thesis. This was done by using publicly accessible induction loop data to infer when a road did not see any traffic and subsequently mark it as closed if a specific threshold was reached for an entire day. This dataset was verified against various known closures and a private dataset with a list of all archived highway closures.

46 centroids were chosen as origins and destinations and a routing model was created to use Dijkstra's algorithm to find the fastest route between each OD-pair. This undisrupted travel time and route were stored. The routes were compared to the road closure dataset and if a road closure took place on one of the routes, the associated OD-pair was flagged as *exposed* and added to the exposure dataset. A new route was calculated, this time without the road segment that was closed which forced a detour between the OD-pairs. This new, disrupted, travel time was stored alongside the undisrupted travel time and a detour ratio was calculated to determine what the relative extra travel time was.

This detour ratio is used in a fixed-effects panel model, a restricted form of the two-way fixed effects (TWFE) design with calendar time controls, to estimate the causal relation between road closures and rail ridership. The ridership was extracted from automatic fare collection data obtained through the OV-chipkaart system used by NS. A baseline is computed to predict what the ridership should have been on a set day, and it is compared to the actual observed ridership on that day. This normalised ridership is then input into the TWFE-model to estimate the responses of certain characteristics on rail ridership and to determine causality between road closures and rail ridership.

Passenger heterogeneity was estimated by dividing the passengers into six groups, each with distinct characteristics. By adding extra terms to the TWFE model, the individual responses of these groups can be estimated to determine if any of them respond differently to the road closures. Road closure heterogeneity was determined based on the duration and location of the road closure, the last of which has an impact on the mean detour ratio as well as the rail competitiveness on that specific corridor, both of which were also inputs to separate models.

Finally, the estimated effects were applied predictively. A pipeline was built that takes the specification of a planned closure, computes the detour ratio for every OD-pair, and combines it with the estimated coefficients and a baseline to forecast the expected ridership uplift, including confidence intervals and a flag for predictions outside the range of conditions the model was estimated on. The pipeline was validated against ten closure events from 2026, outside the estimation period, on three corridors chosen to test the model under favourable, noisy, and extreme conditions respectively.

215 road closures of longer than two days took place in the study period. The duration of those closures is highly skewed towards the lower end with most of the closures taking two or three days, indicative of weekend closures. 22 closures took longer than a week, on average eleven per year. Roughly 75% of all OD-pairs were exposed to some form of road closure, with the sparsely populated North of the country seeing fewer and the densely populated centre-West seeing more closures. The dataset constructed is cross-validated and can be used for future research.

Road closures of a duration longer than eight days see a significant increase in rail ridership which increases when the detour ratio between specific OD-pairs becomes higher. Figure 1 shows a 3.7% increase in ridership found at the median detour ratio for 8+ day closures. Shorter road closures see no significant change in ridership. Rail competitiveness has small and non-significant effects on the ridership although the direction of the increase is in line with expectations and literature.

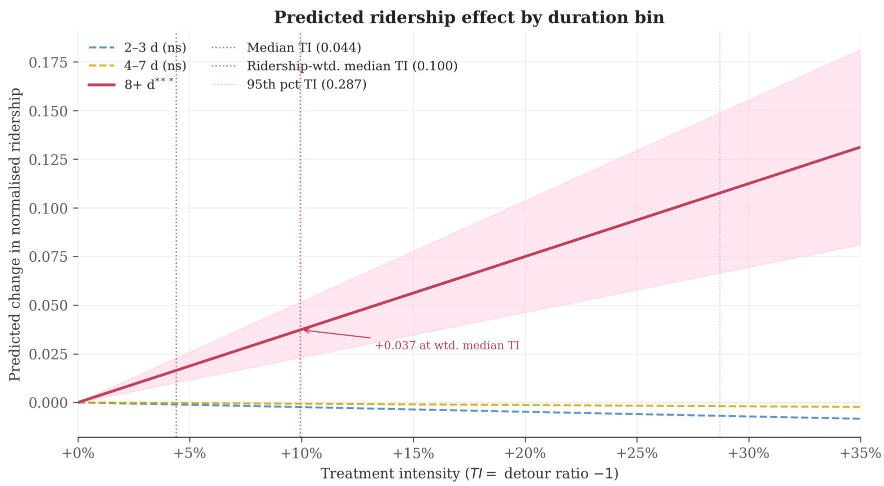


Figure 1: Main finding of this research. Rail ridership increases when the relative detour by car gets longer but only for closures with a duration longer than seven days.

Passenger types respond to the closures in different magnitudes. Incidental travellers see the highest response, almost double the aggregate. Business and student travellers, who do not pay for their journey, also see a larger than aggregate response to road closures. People travelling for work who are paying for their own journey as well as passenger who are likely on a social or recreational trip, respond less strongly to road closures.

The validation shows that the estimated effects transfer to closures the model has never seen, within limits. A 44-day closure of the A28 was predicted almost exactly (+8.3% forecast against +6.7% observed). Seven short closures of the A12 confirmed that short-duration closures produce no response distinguishable from normal day-to-day variation. The ten-day closure of the Vlaketunnel in the A58, the only motorway connection to Zeeland, produced a response two to three times larger than forecast, showing that in corridors without road alternatives the model's forecasts act as a lower bound.

## DISCUSSION

The finding that only road closures of eight or more days produce a significant increase in rail ridership is consistent with the theoretical prediction that mode choice is habitual and resistant to change. For shorter closures, travellers have a large choice of within-mode strategies to cope with the closure ranging from accepting the delay to working from home or cancelling trips altogether. Only once the closure extends into the second working week do the travellers feel the need to change their coping strategy and therefore possibly see modal change as a viable alternative. The robustness analysis supports this as alternative duration bin boundaries produce virtually identical coefficients.

The estimated ridership increase is modest in absolute terms. At the ridership-weighted median treatment intensity of 10%, the predicted increase is 3.7%, increasing to 11.5% at the 95th percentile. This is lower than the effects reported by Kattan et al. (2013) and Tennøy and Hagen (2021), which is expected for two reasons. First, those studies examined closures of 14 months, far beyond the lower boundary duration range studied here. Second, those studies relied on stated preference methods, which are known to overstate behavioural responses.

An important limitation is that the approach limits observations to AFC-data and therefore only something can be said about the rail ridership, not the car traffic counts and therefore not about the modal shift. The results presented are therefore limited to an increase in rail ridership and nothing can be said about modal shifting or decreases in traffic. Treatment intensity and railway travel time ratio is based on free-flow car traffic and therefore not completely representative of real-life scenarios. A detour will likely see more congestion and therefore become even less attractive than stated in this research. The estimated ridership responses are therefore likely a lower bound on the full effect.

The out-of-sample validation strengthens the causal interpretation of the results: effects that were artefacts of the estimation sample would not reproduce observed ridership during unseen closures. It also defines the domain in which the model can be used: forecasts are reliable under conditions resembling the estimation sample, and act as flagged lower bounds in extreme corridors that the sample does not represent.

## CONCLUSION

The goal of this thesis was to answer the question of what the causal effect of planned road closures on rail ridership is in the Netherlands. Three contributions were made to address this question. First, a dataset of 215 full-day highway closure events was constructed from publicly available traffic intensity data spanning 882 closure days across 2024 and 2025. This dataset is cross-validated and available for future research. Second, a revealed-preference causal analysis was conducted using NS OV-Chipkaart automated fare collection data, applying a fixed-effects panel design to a panel of 1,743 origin–destination pairs. Third, the causal estimates were embedded in a validated prediction pipeline that forecasts the expected ridership uplift of a planned closure before it begins.

The main finding is that planned highway closures cause a measurable increase in rail ridership but only under specific circumstances. The closure must be at least eight days long. Shorter closures produce no detectable change in ridership in either direction. For closures of eight days or longer, the estimated ridership increase at the ridership-weighted median treatment intensity is 3.7%, rising with the severity of the detour. The duration is therefore the governing characteristic of the response with the detour ratio, and thereby the location, playing another important role.

Passenger heterogeneity is substantial. Incidental travellers respond most strongly, with an estimated effect nearly double that of the aggregate. Business travellers with employer-paid cards and students, both of whom face no marginal cost at the point of travel, also respond above the aggregate. Commuters paying for their own subscription and social or recreational travellers respond below the aggregate. These differences are statistically significant and have direct implications for operational planning. Travellers most likely to board during a long closure are also those most sensitive to targeted interventions such as route-specific promotions and coordination with employers and educational institutions.

The revealed preference estimates cannot establish whether the additional rail trips represent travellers switching from the car. This remains the key open questions which could be answered by a large-scale stated preference survey. Future research should also examine whether the ridership increases observed during the closures are lasting and therefore lead to increases in ridership beyond the duration of the road closure. For NS, the practical result is that the expected ridership response to a planned closure can now be forecast in advance, with the pipeline itself indicating when a forecast falls outside the conditions under which its accuracy was demonstrated.

Practical recommendations to railway operators all start with good communication between the operators and the highway authorities, noting well in advance when and where large-scale highway closures are going to take place. If the closures are expected to take longer than a week, serious consideration has to be given to reinforcing the railway operations by increasing capacity especially during peak hours. If that is not a possibility, the operators should communicate with relevant large employers and educational institutions affected by the closure as these groups of passengers account for a large share of extra passengers which can be influenced.

*Cities have the capability of providing something for everybody, only because, and only when, they are created by everybody.*

*— Jane Jacobs, 1961*

# ACKNOWLEDGEMENTS

The past eight years have been an absolute joy. I have gone from a rather shy high school boy, unable to understand my rather chaotic and easily distracted self, to someone who feels ready to finally work and be a professional. But I can not attribute this all to myself as many people have helped me get to where I am today and in those eight years I sure have come across my fair share of amazing friends, colleagues, teachers, and my family.

I would like to start off by thanking the professional committee behind this thesis. Firstly, I want to thank Maarten, who taught me the ins and outs of travel behaviour and helped me with it in the context of this thesis. Secondly, I want to thank Niels, not only for helping me get this opportunity at NS, but also for making me always take a step back and put my findings in a broader context. Thirdly I want to thank Menno for his valuable insights into travel behaviour from a more practical perspective and for pushing me to examine every facet of this behaviour during my thesis. Lastly, I want to thank Mahsa for our weekly meetings where I was taught not only about how highways work and how to examine them, but most importantly how to tackle a large piece of research such as this. I believe these meetings made this thesis into a much more professional report than I could have hoped.

Although Menno was my main supervisor at NS, I also want to thank my colleagues there who were not only very welcoming of me and my fellow interns, but also made me feel like my research would actually become of use later on. I hope to see longer trains during road closures in the future! Besides the regular staff at NS, I felt like I was also part of a team of five TU Delft graduate interns who, at times, helped me get out of my thesis-slump. I want to thank my neighbour Thijs in particular for being my daily sparring partner during our, rather frequent, coffee breaks.

Then on to my friends. Honestly, I don't know where to begin as I could not have imagined having such a large group of friends around me during my wonderful time at this faculty, but let's start chronologically. I still remember my first-years weekend and the opening week of the university like it was yesterday. The fun I had during those first years is as invaluable as the friends I made during those first few weeks. Those friends helped me through the roughest time of my studies, the COVID-period, a period I think most of my fellow students do not like to be reminded about. My fifth year was likely the year in which I had the most fun, and I must admit that it was not because I was studying. During this year I did a board year which turned out to be invaluable to my personal development. I want to thank all of my fellow board members, who have obviously become some of my best friends, for this amazing year together. I couldn't keep away from committees and associations so during my master's I joined Dispuut Verkeer and made a whole new group of friends with whom I could finally talk about my main interest, public transportation. Jokes aside, I had some great two years of my Master degree with you all.

For the past three years there has been one person who has pushed me to strive the most, and that is my girlfriend, Martha. I can't begin to count the number of times I was unable to see the wood for the trees and you were there to help me out. I'm also grateful to my wonderful family for all the warm times spent together, and to my dog for all the warm cuddles.

Lastly, but most importantly, I want to thank my parents. For me this might be the end of eight years of studying, but for you this might be the end of 25 years of raising me. I might not have made this journey as easy as it could have been, but if there is one thing I have learned, it's that you will always be there for me, no matter what. Thank you for supporting me, mentally, financially, and lovingly.

Thank you all
Love, *Hugo*

# NOMENCLATURE

| | |
|---|---|
| AFC | Automated Fare Collection |
| CI | Confidence Interval |
| DiD | Difference-in-Differences |
| HRN | Hoofdrailnet |
| HWN | Hoofdwegennet |
| LoS | Level of Service |
| NS | Nederlandse Spoorwegen |
| NDW | Nationaal Dataportaal Wegverkeer |
| NWB | Nationaal Wegenbestand |
| OD | Origin-Destination |
| OLS | Ordinary Least Squares |
| PT | Public Transport |
| RP | Revealed Preference |
| RUM | Random Utility Maximisation |
| RWS | Rijkswaterstaat |
| SP | Stated Preference |
| TI | Treatment Intensity |
| TTR | Travel Time Ratio |
| TWFE | Two-Way Fixed Effects |
| VoT | Value of Time |

CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1 INTRODUCTION

Roads require constant maintenance, and the Dutch road network is no exception. A large part of this network was built between the 1960s and 1980s and is now nearing the end of its design life, with many structures having already had their replacement date pushed back. Vehicles have become heavier over the years, both because people are buying larger cars and because the shift towards electric vehicles means that these cars have to lug heavy battery packs with them. These heavier vehicles place more strain on infrastructure that was not designed for the traffic of today. Rijkswaterstaat (RWS), the administration responsible for the main road network, has estimated that over €34 billion is needed to keep this ageing infrastructure in working order in the coming years (Rijkswaterstaat, 2024). Combined with a Dutch emphasis on preventative maintenance, this means a steady and predictable stream of planned road closures for the next few years, often closing down lanes or entire sections of highway.

The Netherlands is also a dense country with an extensive and frequent railway network where, on a subset of well-served corridors (e.g. Amsterdam–Utrecht), rail travel times are already competitive with the car. During a road closure that alternative becomes more attractive still as driving the same route takes longer. In principle, a closure could therefore push a share of travellers onto the train. However, the railway network is itself busy, both on the tracks and in the trains, and its timetable is rigid. It cannot absorb a sudden influx of passengers without planning and investment ahead of time. An ongoing policy push towards more sustainable travel adds to the relevance of understanding when and how travellers move between the two modes.

A road closure forces a detour that lengthens the car journey. For initially long journey a short closure somewhere along the way can be mitigated by taking a different route which often does not significantly increase the travel time; however shorter journeys are relatively highly impacted by highway closures. For some travellers this tips the balance towards the train as it now becomes a more attractive option due to the higher reliability and lower travel time compared to the car. This switch from one mode to another is called "modal shift" and is the key question when it comes to operational response during road closures.

Because these closures are planned and announced in advance, travellers can reconsider their mode rather than just react under pressure. Therefore planned road closures are fundamentally different from sudden incidents which often occur when the traveller is already in their vehicle and has no other choice than to continue using that mode. These planned road closures are also easier to identify as their timing and locations are known which creates a good environment for a causal analysis.

## 1.1 PROBLEM STATEMENT

Currently NS responds to highway closures on an ad-hoc basis. All road closures are communicated in advance by Rijkswaterstaat which are sometimes accompanied by requests for impact mitigation measures such as supplying NS-business cards for travellers affected by the road closures or discounts on affected parallel railway lines. However, not all road closures get such treatment and almost none get an operational response which increases capacity on the railway lines which serve as the alternatives to the highways during the closure.

The reason for this asymmetry in response is that NS can not clearly predict if they can expect more passengers during planned road closures on the main roadway network in The Netherlands. This poses a problem to planners who need to know if they should assign more capacity to certain routes during road closures or not, who those extra travellers could be, and if they might be of monetary value.

### 1.1.1 *Research gap*

Travel behaviour under network disruptions has been studied extensively but predominantly within-mode. Existing work examines how travellers re-route, re-schedule, or cancel trips within the disrupted mode (Zhu and Levinson, 2012), while modal change remains largely unexamined. The few studies that do address modal shift during road closures rely on stated preference surveys performed on single cases in urban context (Kattan et al., 2013; Tennøy and Hagen, 2021). Large-scale revealed preference evidence linking road closure to measured changes in rail ridership is absent from the literature. The Netherlands offers conditions to close this gap with its frequent planned highway closures, competitive intercity rail network, and nationwide AFC data. Chapter 2 reviews this literature in detail. The research questions below follow from it.

## 1.2 RESEARCH QUESTIONS

The main research question of this thesis is:

> **What is the causal effect of planned road closures on rail ridership in the Netherlands?**

To answer this question, the following sub-questions are addressed:

1. Which origin–destination pairs are exposed to planned road closures?
2. What is the estimated causal impact of exposure to planned road closures on rail ridership?

3. Which passenger categories exhibit disproportionate increases in rail ridership during road closures?

4. How does the impact of planned road closures on rail ridership vary across different road closure characteristics?

5. How can the estimated causal effects be used to predict rail ridership during planned road closures?

## 1.3 scope

This research is bounded along two dimensions: physical scope, which defines the transport network and infrastructure under study, and temporal scope, which defines the period of observation. Together, these boundaries reflect the availability of data and the gap identified in the literature, namely the absence of revealed-preference evidence on car-to-rail modal shift during planned highway disruptions on longer interurban corridors in the Netherlands.

### 1.3.1 *Physical Scope*

The analysis is restricted to rail services operated by NS on the Hoofdrailnet, the main railway network as designated by the Dutch government. This restriction is imposed by the AFC dataset, which records only OV-Chipkaart transactions on NS-operated services. Consequently, travel by other rail operators, as well as travel by bus, tram, and metro, falls outside the scope of this study.

The operator identity of each trip is determined by the tap-in and tap-out gates, which are operator-specific. An observed transaction therefore unambiguously identifies NS as the carrier, even on corridors where multiple operators provide service. On the Enschede–Almelo corridor, for instance, both NS and Blauwnet operate regular services; only transactions recorded in the NS AFC system are included. It is assumed that the AFC data correctly reflects NS travel and that passengers did not inadvertently board a service operated by another carrier.

The road network is similarly restricted to Dutch public roads recorded in the NDW dataset. International corridors and roads not covered by NDW fall outside the scope of this analysis.

### 1.3.2 *Temporal Scope*

The primary observation window is the last two calendar years (2024 - 2025). Within this time-window both the SHiVi data used for road work identification and the NS AFC-data is highly detailed and the networks analysed stay the same with only minor network adjustments.

## 1.4 THESIS OUTLINE

Chapter 2 discusses the relevant literature on this subject and identifies the relevant research gap. In chapter 3 the methodology is laid out starting with the data preparation and the conceptual framework. After this the methodology for the exposure modelling, where the road closures are identified, is presented and finally the causal analysis is explained. The results are presented in chapter 4 including an interpretation and answering of the sub questions. In the discussion chapter 5, the results are related back to the literature and a high-level interpretation is formulated to understand the impact of the findings on mobility in the Netherlands. Lastly, in chapter 6 the main findings are drawn together and the research question is answered.

# 2 LITERATURE REVIEW

## 2.1 LITERATURE SEARCH STRATEGY

A structured literature matrix was developed to synthesise the selected studies across dimensions such as disruption type, data source and behavioural response. The matrix is provided in Appendix A.

Keywords used for this research are listed in table 1. Besides keywords in Scopus and Google Scholar, AI-based tools were used to iteratively refine search terms and identify potentially relevant foundational studies. All suggested references were subsequently verified through Scopus and Google Scholar and evaluated based on relevance, methodological quality and citation impact before inclusion in the literature review.

Table 1: Boolean search strings used in Scopus and Web of Science

| **Topic** | **Search string** |
|---|---|
| Road works and modal shift | "road works" AND ("mode shift" OR "modal shift") |
| Planned disruptions | "planned disruption" AND "travel behaviour" |
| Capacity reduction | "capacity reduction" AND ("mode choice" OR "modal shift") |
| Travel time reliability | "travel time variability" OR "travel time reliability" |
| Elasticity | "travel demand elasticity" AND ("rail" OR "public transport") |
| Random utility theory | "random utility" AND "mode choice" |
| AFC / revealed preference | "automated fare collection" AND "mode choice" |

## 2.2 LITERATURE REVIEW

Dense, intermodal transport networks give people freedom to choose how they move around, be that by public transportation, car or active modes such as cycling and walking. These networks also give people the option to change their journey if they want or if they are forced because of a disruption on their usual route and mode. Network disruptions can broadly be divided into two categories, planned and unplanned. Unplanned disruptions are caused by unforeseen circumstances such as adverse weather, accidents or equipment malfunctions. Although they can be mitigated by making a network more robust, they can never be fully prevented. Planned disruptions are disruptions in the network of which it is known beforehand that they will occur. These come mainly in the form of maintenance to the infrastructure but can also be due to events or other scheduled interventions (Cats et al., 2017).

Within those categories more variations exist. Disruptions could come in the form of a total closure, say a bridge collapse or the repaving of a highway if it is performed all at once. Disruptions can also be partial, for example due to accidents, hedge trimming or small-scale flooding events (Cats and Jenelius, 2018). The disruptions also vary in duration with long-term disruptions due to collapses of maintenance having prolonged effects on the system whereas short-term disruptions only produce a short shock (Sullivan et al., 2010).

While both planned and unplanned disruptions impose inconvenience on travellers, they differ fundamentally in the travel behaviour they evoke. Unplanned disruptions occur unexpectedly and give the traveller little to no time to adapt or re-plan their journey (Zhu and Levinson, 2012). This sudden disruption triggers short-term, reactive responses which can lead to less predictable decisions made under time pressure. Planned disruptions do give travellers the time to respond, especially when communicated broadly and well in advance. Besides giving the traveller the option to re-route in the same mode, it also allows them to explore options beyond their current travel habits more thoroughly. This can vary from a change of mode, an earlier or later departure and arrival time or even cancelling or rescheduling the activities completely. Planned disruptions can also take longer than unplanned ones, although this is not always the case, which can force travellers to overhaul their travel habits as they will have to deal with the new circumstances for an extended period.

## 2.3 BEHAVIOURAL FOUNDATIONS OF MODE CHOICE

Travellers choose their preferred mode based on the attributes associated with each alternative. While the analyst cannot predict an individual's choice due to unobserved factors influencing utility, choice probabilities can be modelled using Random Utility Maximisation (RUM) theory. Under RUM, individuals are assumed to select the alternative that yields the highest utility, where utility consists of an observed component and a random error term (McFadden, 1974; Ben-Akiva and Lerman, 1985). The observed component represents the utility that can be measured, such as the attributes of the alternative (price, travel time) and the characteristics of the decision maker (age, income). Captured in the error term are all factors influencing the decision that are not observed, such as individual tastes, preferences or attributes that were not measured. Planned road works can therefore be interpreted as an exogenous change in the generalised cost of car travel, altering the relative utility between car and rail alternatives.

People value the time differently. Those who are willing to incur significant delays before changing their journey have a lower value of time (VoT) than those who are willing to pay for time-savings during their journey. Travel time is typically converted into monetary units through this VoT representing the rate at which travellers are willing to trade money for time savings (Small and Verhoef, 2007).

Small et al. (2005) find that travellers place significant monetary value on reliability, not just on travel time. Travellers value reducing one minute of travel time variability almost as much as reducing one minute of average travel time. Road works introduce a large amount of variability in the reliability of a journey although travellers do adapt after having experienced the unreliability (Jamous and Balijepalli, 2018). Road works can therefore be assumed to change the generalised cost of travel by increasing the unreliability of a journey.

Road works also present an opportunity for long-term change. Travel mode is habitual, meaning that travellers tend to gravitate towards the modes which they are used to thereby reinforcing their status quo (Gärling and Axhausen, 2003). Such habit can make voluntary modal change difficult, even when alternatives are objectively competitive. Disruptions can break these habits as they force people to adapt their journey possibly including a modal change. Prolonged capacity reductions may therefore act as a trigger for breaking habits, increasing the likelihood of sustained modal switching. This can also work the other way around. Long-term track work can force passengers to find alternative travel modes for the time being which might show them that other options exist which they could use as their permanent new mode of travel. The aggregate ridership has been shown to decrease after long-term track closures, but individual riders have also been shown to reduce their trip frequency, likely having formed new modal habits or acquired private transport (Krijgsman van Spangenberg et al., 2026).

## 2.4 TRAVEL BEHAVIOUR UNDER NETWORK DISRUPTIONS

Travel disruptions require people to adapt their journey from their usual route and mode to accommodate the new circumstances. People can respond by changing their route, adjusting travel times to avoid congestion, switching modes or cancelling their journey altogether (Zhu and Levinson, 2012).

Although travellers can respond to disruptions in various ways by altering their journey, not all alterations are equally relevant in every context. This thesis focuses on the modal shift of travellers during planned road capacity reductions, and in particular on shifts from car to the train. Other responses such as trip cancellation, re-routing and changing of departure and arrival times are not within the scope of this research.

Chorus et al. (2006) note that it is hard to change mode choice as opposed to other travel factors due to its habitual nature. Mode is driven by different factors such as status and vehicle ownership which have strong influence on the final choice made.

## 2.5 PLANNED VERSUS UNPLANNED DISRUPTIONS

Travel behaviour under unplanned disruptions has been widely studied in the transport literature. Unexpected events such as accidents, infrastructure failures and adverse weather conditions require immediate adaptation and often trigger short-term behavioural responses, including rerouting, departure time shifts and trip cancellation (Zhu and Levinson, 2012). Because these disruptions occur without prior warning, travellers frequently rely on heuristics or real-time information when making decisions under time pressure (Chorus et al., 2006). From a network perspective, the robustness of the transport system determines its ability to absorb such sudden capacity losses (Cats, 2016).

Planned disruptions differ fundamentally in that they are typically communicated in advance. The availability and quality of information can significantly influence behavioural responses, as travellers incorporate expected delays into their decision-making process (Chorus et al., 2006). Empirical evidence suggests that advance communication about planned road works can lead to measurable behavioural adjustments, with a share of travellers altering departure times, routes or travel modes in response to anticipated congestion (Hillcoat, 2018). Evidence from a three-month rail line closure in Copenhagen shows that prolonged planned disruptions can induce substantial behavioural adaptation, including persistent changes in travel patterns even after service restoration (Eltved et al., 2021). Unlike unplanned incidents, which often result in temporary adjustments, planned capacity reductions may create conditions for more deliberate reassessment of travel behaviour. However, many existing studies focus on behavioural responses within a single mode and do not explicitly examine modal switch.

## 2.6 BEHAVIOURAL ADAPTATION AND TRAVELLER CONSTRAINTS

Some travellers are inflexible with their journey, they have little freedom to change their travel behaviour in terms of route, mode, time and cost. These travellers are often described as captive travellers. Evidence from unplanned public transport strikes shows that such constraints can lead to high rates of trip cancellation, particularly among travellers without access to a private car (Van Exel and Rietveld, 2001). Captivity is not a fixed characteristic. Instead, it varies by context, for example by trip purpose, time of day, and day of the week. Studies of planned disruptions indicate that travellers during peak periods and on weekdays tend to exhibit lower behavioural flexibility, suggesting higher levels of effective captivity even when disruptions are known in advance (Yap and Cats, 2022).

## 2.7 PLANNED ROAD CLOSURES AND PUBLIC TRANSPORT DEMAND

Travel disruptions to PT networks can have great effects on road congestion often because the road network is already near capacity during peak hours and the few extra cars on the road can tip it over its breaking point (Lo and Hall, 2006; Spyropoulou, 2020).

A 14-month road capacity reduction in Calgary studied by Kattan et al. (2013) was shown to cause an increase in PT ridership by 6.6% which coincided with a total drop in trips of 4.4% and a forward shift of departure time by 15 minutes. A study by (Tennøy and Hagen, 2021) of a 14 month tunnel closure in Oslo showed that public transport usage increased from a modal share of 40% to 49% during and even after the closure. These results were gathered from a stated preference (SP) survey.

Much research has been done on travel behaviour due to PT disruptions. Yap and Cats (2022) studied planned tram disruptions in Amsterdam where a lower demand response was observed during peak hours and on weekdays. This shows that captive passengers are less likely to change their journey due to planned disruptions. The study looked solely at re-routing without modal change and used machine learning to analyse the revealed preference (RP) data from the Dutch OV-Chipkaart Automated Fare Collection (AFC) system.

Studies examining travel behaviour under disruptions draw on two broad categories of empirical evidence. Stated preference (SP) surveys ask respondents how they would behave under hypothetical scenarios, making them useful for studying contexts where observed data is scarce. However, SP methods are subject to hypothetical bias: respondents systematically over-report socially desirable behaviours such as switching to public transport, and their stated intentions do not reliably translate into actual travel decisions (Ben-Akiva and Lerman, 1985). Revealed preference (RP) data, on the other hand, captures what travellers actually do, making it the gold standard for behavioural inference. AFC systems such as the Dutch OV-Chipkaart provide large-scale RP data that has been used to study travel behaviour under planned PT disruptions (Yap and Cats, 2022; Yap et al., 2018), but has not yet been applied to study modal shift in response to road closures.

## 2.8 GAP

A substantial body of research has examined how road and rail disruptions influence travel behaviour, establishing that response varies by disruption type, traveller characteristics and the available information (Zhu and Levinson, 2012; Cats and Jenelius, 2018). However, the majority of this work focuses on redistributing demand within a single mode; studying re-routing, departure time shifts, or trip cancellation, rather than examining modal change.

Studies that do focus on planned disruptions demonstrate that advance communication enables more deliberate behavioural adaptation than unplanned events (Hillcoat, 2018; Eltved et al., 2021). Yet even within this literature, the focus remains predominantly on within-mode responses. Eltved et al. (2021) study passenger redistribution following a rail line closure and Yap and Cats (2022) analyse re-routing behaviour during planned tram disruptions in Amsterdam, neither examines whether traveller switch to an entirely different mode as a result.

The few studies that do examine modal shift in the context of road closures (Kattan et al., 2013; Tennøy and Hagen, 2021) provide indicative evidence that public transport ridership increases during prolonged road closures. However, these studies rely on stated preference surveys, which are subject to hypothetical bias and do not reflect actual travel decisions. Large-scale revealed preference evidence linking specific road closure events to measurable changes in rail ridership remains absent from the literature.

The Dutch context is particularly well-suited to studying this phenomenon, and yet remains understudied. The combination of a dense highway network nearing end-of-life, a broad maintenance programme generating frequent planned closures, and a competitive intercity rail network creates conditions under which modal shift is both plausible and relevant. Furthermore, the availability of nationwide AFC data from the OV-Chipkaart system provides an opportunity to study revealed travel behaviour at scale. While the Dutch AFC system has been exploited for studying behaviour during planned PT disruptions (Yap and Cats, 2022; Yap et al., 2018), the question of whether planned road closures generate measurable increases in rail ridership remains unaddressed in the Dutch context.

Previous research has focussed on case-study based examples (Yap and Cats, 2022; Yap et al., 2018; Tennøy and Hagen, 2021; Kattan et al., 2013), limiting the conclusions to a narrow set of corridors and types of disruptions. As a result, the literature offers limited understanding of how responses vary across different road work characteristics, network contexts, and the relative attractiveness of available alternatives. This makes it difficult to determine under which conditions planned road closures are most likely to induce a shift towards rail.

This thesis addresses the gap found in the literature by combining road works data and revealed preference AFC data to causally identify the effect of planned highway closures on rail ridership in the Netherlands. Unlike prior work, it does not rely on stated preferences, does not restrict behavioural responses to within-mode re-routing, uses a large set of road closure instances to determine responses based on characteristics, and explicitly targets the car-to-intercity rail modal shift on longer interurban corridors.

# 3 METHODOLOGY

In this chapter the methodology for estimating the causal effect of highway closures on rail ridership is presented and for applying the resulting estimates predictively. The method is designed to be transferable. Each section first describes the generic approac, which can be applied in any setting that meets the data requirements set out in 3.3, and then its implementation in the specific Dutch NS context. Figure 2 summarises the methodology, three data streams are processed into a single analysis panel, on which the causal effects are estimated and subsequently applied predictively. The following sections describe each stage in the order shown in the methodological overview, concluding with the out-of-sample validation.

## 3.1 METHODOLOGICAL OVERVIEW

Figure 2 summarises the methodology of this thesis. Three data streams are processed in parallel, firstly the daily traffic intensities, from which executed road closures are detected. Secondly the road network, on which baseline and detour routes are computed to obtain a treatment intensity per OD-pair. Lastly the AFC ridership data, which is normalised against pair-specific baselines. The three streams converge into a single analysis panel at the OD-pair $\times$ day level, on which the causal effect of closures on rail ridership is estimated using a two-way fixed effects model. The resulting estimates feed a prediction pipeline that forecasts the ridership response to a planned closure before it takes place, and which is validated against closures outside the estimation sample.

The closure dataset produced by the first step is both an intermediate product and a contribution in its own right: it is cross-validated against a non-public archive of planned closures and is available for future research.

Table 2 collects the parameters and specification choices of the method. Each choice is motivated in the section indicated in the last column. Together, the figure and the table specify what is required to reproduce this analysis or to apply the method in another context: the processing structure, the choices to be made, and, with the data requirements in Section 3.3, the data needed to make them. The following sections describe each stage of the figure in order.

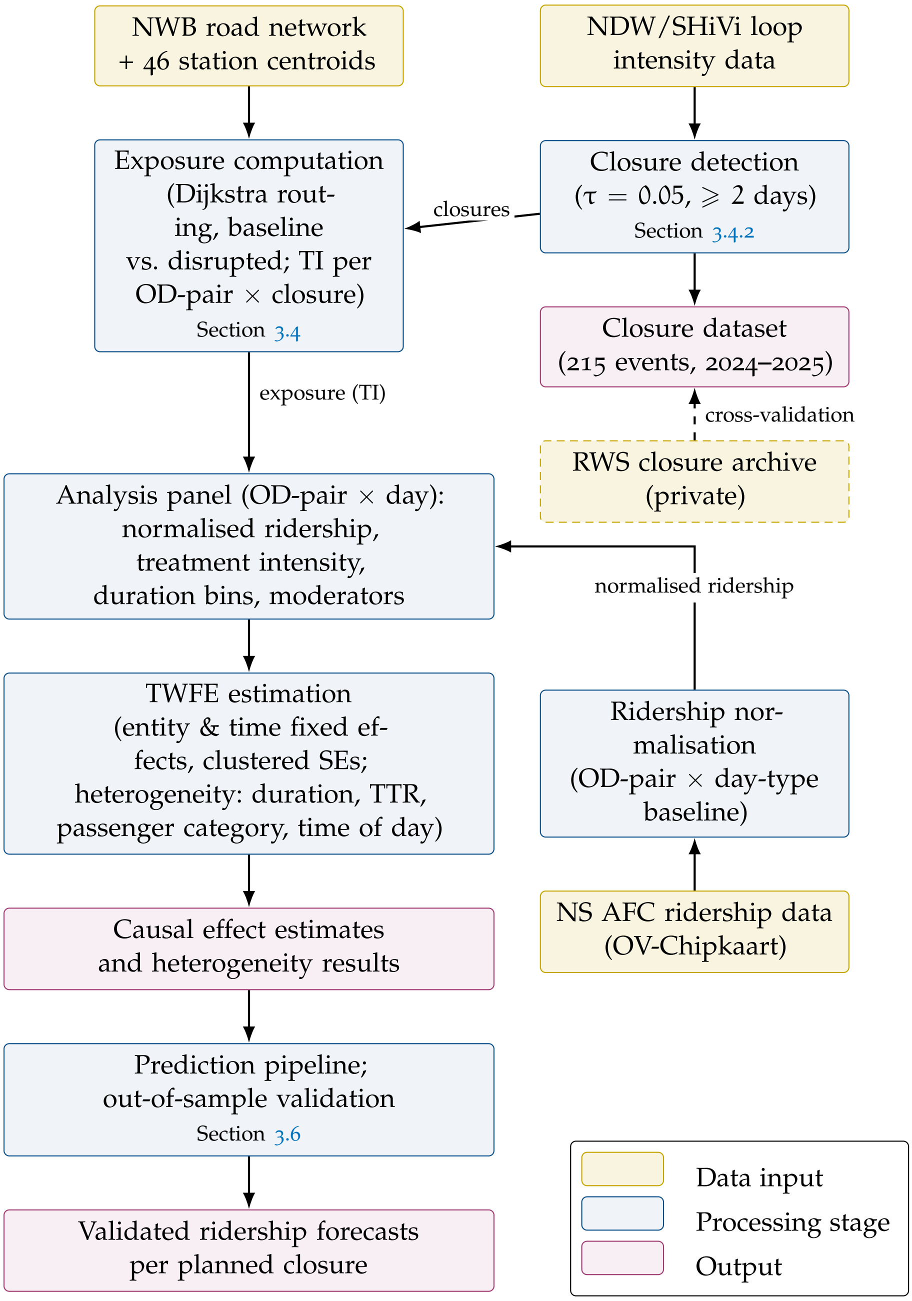


Figure 2: Overview of the methodology. Three data streams — traffic intensity, the road network, and AFC ridership — are processed into a single analysis panel on which the causal effects are estimated. The estimates feed the prediction pipeline, which is detailed in Figure 9. Section numbers indicate where each stage is described.

Table 2: Method parameters and specification choices. Each parameter is motivated in the section indicated; the values shown are those used in this study and constitute the full set of choices required to apply the method in another context.

| Parameter | Value used | Set in |
|---|---|---|
| Closure detection threshold ($\tau$) | 0.05 × baseline intensity | §3.4.2 |
| Minimum closure duration | 2 consecutive days | §3.4.2 |
| Closure duration bins | 2–3, 4–7, 8+ days (ref. 2–3) | §3.5.3.2 |
| Ridership baseline | pooled median per OD-pair × day type | §3.5.1.1 |
| N-road free-flow speed | 80 km/h (calibrated) | §3.4 |
| Time-of-day boundaries | peaks 06:30–09:00, 16:00–18:30 | §3.5.6 |
| Prediction-year trend correction | 4 weeks pre-closure vs. same period prior year | §3.6.2 |
| Extrapolation threshold | 95th pct. of estimation-sample TI (= 0.287) | §3.6.2 |

## 3.2 CONCEPTUAL FRAMEWORK

Planned road closures are an exogenous intervention in the road network altering the characteristics of the existing network thereby changing the routing options. The resulting detours increase travel times and congestion on alternative roads. In response, travellers adapt their behaviour, including modal shifting, which can be seen in the ridership obtained from the AFC-data. This main causal chain is shown in the centre of Figure 3. Every other element of the framework is annotated with how this study addresses it: controlled, absorbed by the fixed effects design, examined in a heterogeneity analysis, or acknowledged as unobserved.

The strength of the causal effect is shaped by moderators. The severity of the intervention itself depends on trip distance where the closure adds a roughly fixed detour time, which represents a large relative increase for short trips but a small increase relative increase for long trips, so the same closures produces different treatment intensities across OD-pairs. The attractiveness of the behavioural response depends on the quality of the rail alternative, measured as the travel time of rail relative to road, which is examined in the TTR heterogeneity analysis. The tendency to adapt further depends on behavioural determinants such as habits, attitudes, vehicle ownership, and financial constraints, all unobservable characteristics in the available data. Finally, the observable modal shift is bounded by the current modal share on a pair where rail already carries most travel, less room for shifting remains. Pair-specific, time invariant moderators of this kind are absorbed by the entity-fixed effects.

Confounders affect the treatment and independently affects the outcome of the main causal chain. If not controlled for it creates spurious association. In the case of road works and travel, seasonality is the main confounder. Road works are often planned in specific periods and always outside of the "bouwvak" which are the three weeks in summer when no construction takes place. Seasonality also affects people's travel behaviour and will therefore have an effect on the observed rail ridership. These confounders are addressed by the additive time fixed effects, complemented by month, day-of-week, and holiday controls. What remains outside the design is discussed in the limitations.

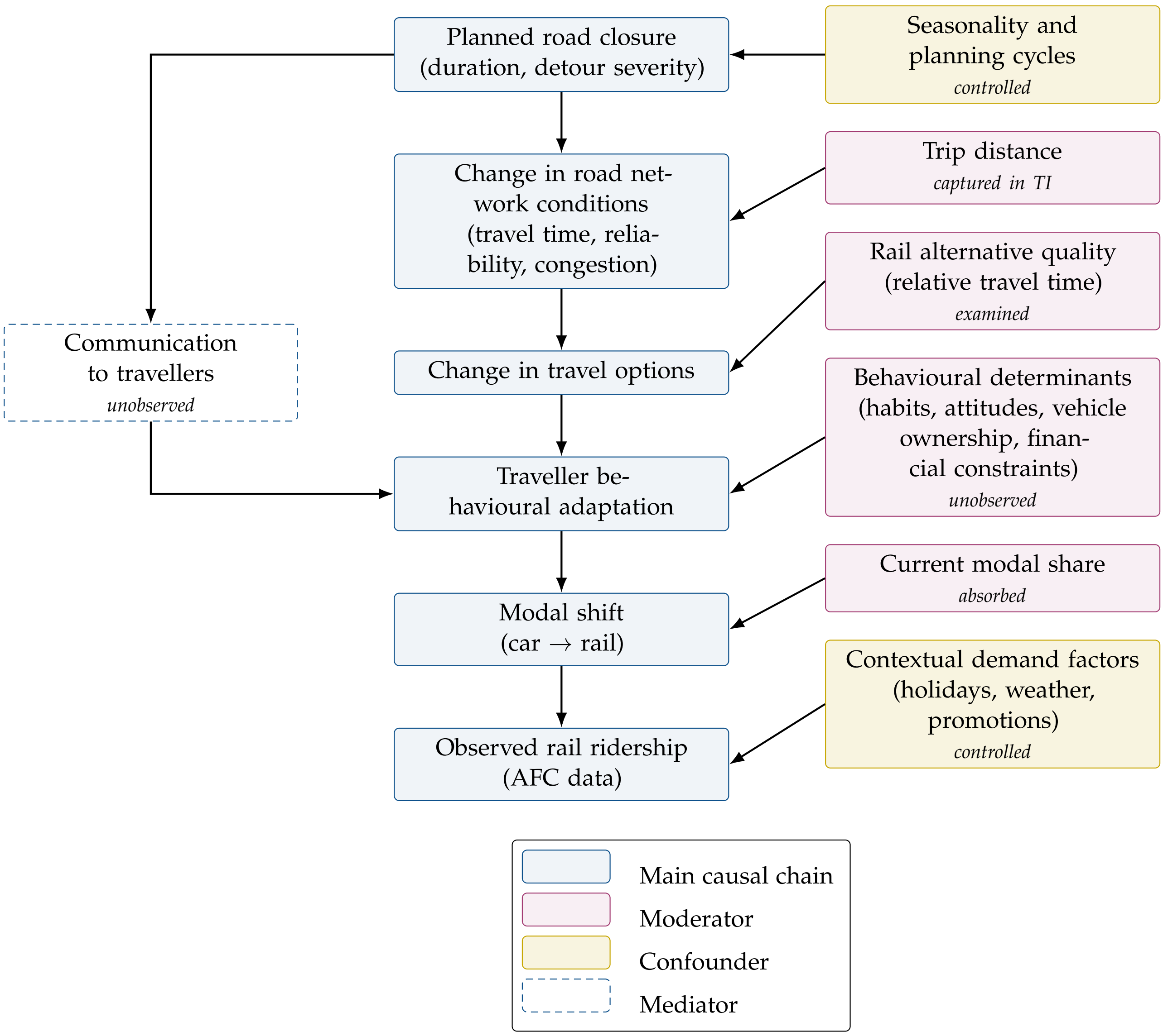


Figure 3: Conceptual framework linking planned road closures to observed rail ridership. The main causal chain runs vertically; moderators and confounders act on it from the right, and communication mediates part of the closure effect. The tag in each element outside the chain indicates how it is addressed: controlled through the fixed effects design and calendar terms, captured or examined in the analysis, absorbed by entity fixed effects, or unobserved and treated in the discussion; details are given in the text.

## 3.3 DATA

The methodology developed in this chapter is not tied to the Dutch context: it can be applied in any setting where three categories of data are available. This section therefore first states the data requirements that the method imposes, independent of any specific dataset, and then describes the datasets used to fulfil each requirement in this study. This separation makes explicit which parts of the data foundation are generic to the method and which are specific to the Dutch and NS implementation.

### 3.3.1 *Data requirements*

The method imposes four data requirements.

First, a record of *when and where full road closures occurred*. The dataset must include the start and end time of the closures and the location. It also needs to be known if the closure was for the entire road or only partial as only fully closed roads are valid for this model. The observations must cover the full study period and the entire network of interest, so that closure detection is not biased towards particular regions or periods.

Second, a *routable representation of the road network*, with travel times or free-flow speeds per segment, from which fastest routes between origins and destinations can be computed both for the undisrupted network and for the network with closed segments removed. This is what turns a detected closure into a treatment intensity per OD-pair.

Third, *rail ridership observations per origin–destination pair per day*, spanning both closure and non-closure days so that a baseline can be constructed for every pair. For the heterogeneity analyses, the ridership observations must additionally be separable by time-of-day period and by a passenger classification; because smart-card data is typically anonymised, this classification must be derivable from card or product attributes rather than from individual travel histories.

Fourth, an *independent record of closures* against which the inferred closure dataset can be cross-validated. This was necessary for this specific research as the method to create the road closure dataset needed to be verified. If the road closure dataset is already verified and can be trusted, this data does not need to be included.

Any national or regional context offering these four categories of data supports the method. The remainder of this section describes the Dutch datasets used here; Table 4 summarises them and maps each to the requirement it fulfils.

### 3.3.2 *Data used*

#### *SHiVi traffic intensities (requirement 1)*

The SHiVi dataset is an open dataset published by the NDW and contains daily traffic intensities for all roads on the main Dutch highway network (HWN) from 3 January 2019 onwards. The underlying measurements are collected by loop detectors embedded in the road surface at intervals of roughly 300–500 metres, aggregated from five-minute observations to daily values. The network is divided into SHiVi segments, each representing a directional road section between successive junctions, where a junction may include an on-ramp, off-ramp, road fork, or intersection. Because no archived dataset reliably records whether planned road closures were actually executed, the SHiVi data is used to infer historical closures from substantial reductions in observed traffic intensity, as described in Section 3.4.2.

*Nationaal Wegenbestand (requirement 2)*

The NWB is an open dataset maintained by the Dutch national and regional governments and describes the entire Dutch road network at the individual road-segment level. For this study, the motorway network is extracted from the NWB and converted into a directed routing graph, in which each segment carries a free-flow travel time. Provincial N-roads serving as natural detour routes are added to the graph with a calibrated free-flow speed of 80 km/h. The graph is used to compute fastest routes between all OD-pairs under both undisrupted and disrupted conditions, as described in Section 3.4.

*NS AFC data (requirement 3)*

Rail ridership is measured using automated fare collection (AFC) data from the OV-Chipkaart system, provided by NS. Although the underlying system registers individual tap-ins and tap-outs, the data used in this study is an aggregated origin–destination flow table: for every combination of travel date, origin station, destination station, and time-of-day period, it records the number of travellers per card-product category. No individual card identifiers are included, so travellers cannot be followed over time; the card-product attribute is therefore the only dimension along which passengers can be classified, and the six passenger categories used in the heterogeneity analysis are derived from it, as described in Section 3.5.5.1. The extracted data spans 2023–2025. The fields used in this study are listed in Table 3.

Table 3: Fields of the NS AFC origin–destination flow data used in this study

| Variable | Description |
|---|---|
| Travel date | Calendar date of travel |
| Origin station | Station where the journey began |
| Destination station | Station where the journey ended |
| Time-of-day period | Morning peak, evening peak, off-peak, or weekend (Table 8) |
| Card-product category | Product classification from which the passenger categories are derived |
| Traveller count | Number of travellers for this combination |

*RWS closure archive (requirement 4)*

A non-public archive of planned highway closures, provided by Rijkswaterstaat, is used exclusively to cross-validate the closure dataset inferred from the SHiVi intensities. The archive records planned rather than executed closures, which is precisely why it cannot replace the inferred dataset as an analysis input, but it provides an independent reference against which the detection method's hit rate is assessed.

Table 4: Overview of datasets used in this study

| Dataset | Req. | Open | Distributor | Purpose in Study |
|---|---|---|---|---|
| SHiVi | 1 | Yes | NDW | Inference of executed road closures (timing, location) from traffic intensities |
| NWB | 2 | Yes | Government | Routing graph for baseline and detour routes; treatment intensity per OD-pair |
| AFC (OV-Chipkaart) | 3 | No | NS | Rail ridership per OD-pair, day, time-of-day period, and passenger category |
| RWS closure archive | 4 | No | RWS | Cross-validation of the inferred closure dataset |

## 3.4 EXPOSURE MODELLING

To model exposure two questions need to be answered: What is defined as exposure and what are the units of analysis that are being exposed? In this study an origin-destination-pair (OD-pair) is considered exposed when planned road works affect one or more road segments on the relevant road-based route between the origin and destination, thereby changing the generalised cost of car travel. Therefore the unit of analysis is the OD-pair and the exposure variable is the road closure. The output of this section will be a list of road closures and their affected OD-pairs which will feed into the causal analysis in the next section.

### 3.4.1 *Definition of OD-pairs and spatial units of analysis*

Rail ridership is observed between origins and destinations on a station-to-station basis. By contrast, highways connect towns, cities, and regions at a broader spatial scale, often serving areas that are associated with multiple railway stations. As a result, station-level ridership data and corridor-level road disruption data are not directly comparable in spatial terms. A common spatial unit is therefore required to link the two.

For this reason, the spatial unit of analysis is defined as a set of aggregated origin-destination (OD) connections between selected study areas. Exposure is defined at the OD-pair level because the relevance of a road closure depends on the full connection between an origin and a destination. A closure near one city does not affect all journeys starting or ending there in the same way; its relevance depends on whether that closure lies on the relevant route between the two places. The unit of analysis is thus explicitly not the individual trip, as modelling personal route choice is not the objective of this study.

Each origin and destination is represented by a single study centroid. This is necessary because the routing model requires one origin point and one destination point for each connection. These centroids represent broader urban or station catchment areas rather than individual stations or travellers. An OD-pair is therefore defined as the ordered connection between two such centroids. Representing study areas by centroids inevitably simplifies variation within those areas, but it makes national-scale routing feasible and ensures that road and rail travel conditions can be compared consistently.

Municipal boundaries are used as a first approximation of catchment areas. For most study areas, these boundaries correspond reasonably well to the area that can be considered functionally linked to the relevant rail node and surrounding highway access. A small number of exceptions require a different approach. For stations located between multiple urban areas, such as Ede-Wageningen and Driebergen-Zeist, the catchment area is defined as the aggregation of the surrounding municipalities served by that station. Schiphol Airport is treated as a separate special case, where the airport itself is used as the representative location rather than the municipality in which it is located.

To determine the representative point of each catchment area, a population-weighted centroid is calculated rather than a simple geometric centre. This ensures that the centroid better reflects where travellers are likely to originate or end their journeys. A purely geometric centroid may be distorted by large uninhabited areas within a municipality and would therefore provide a less realistic basis for national-scale route assignment. This is most important for municipalities such as Rotterdam which include large swathes of uninhabited industrial areas that should not count towards centroid positioning as can be seen in figure 4.

Determining what centroids should be analysed affects the number of OD-pairs quadratically and with it the compute time. Not every possible place in the Netherlands can or should be included and the study focuses on locations where interaction between rail and motorway travel in meaningful. The selected centroids are intended to capture the main spatial structure of interurban travel relevant to the research question. Therefore centroids located far from highways, not served by NS or with low population numbers, should not be studied.

A starting point for centroid selection is a list of 32 destinations used internally by NS for market share assessment. These destinations were selected to cover 95% of travel within the NS network. 25 of these are individual cities, six are parts of cities and one is an airport (Schiphol). The six city parts are two for each of the cities of Amsterdam, Rotterdam and Utrecht. For this study those six were

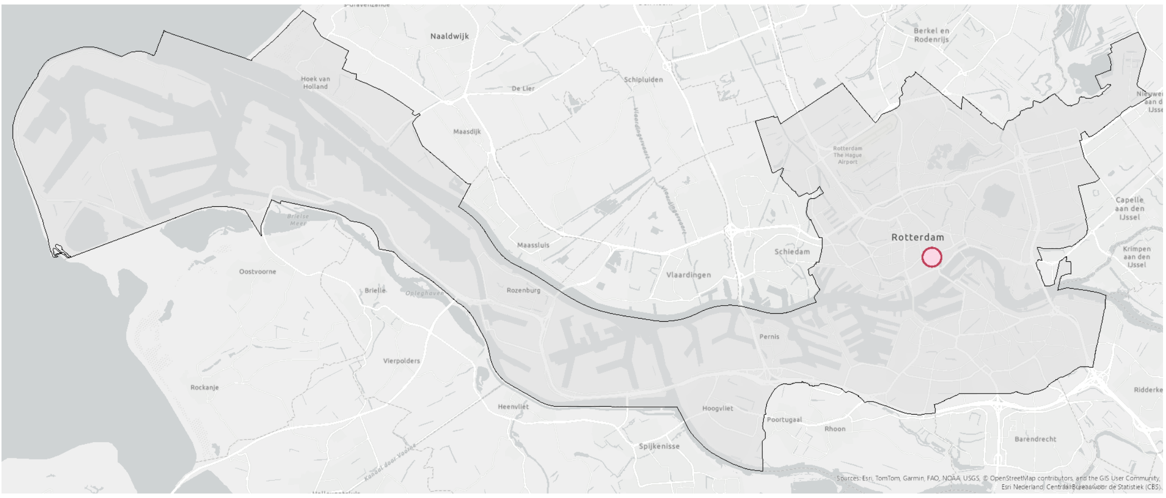


Figure 4: Example of centroid location calculation for Rotterdam. Centroid is clearly located in the populated area, excluding the uninhabited port from centroid positioning.

combined into their respective city to form three nodes. This leaves 29 centroids from this source.

The initial set of 29 centroids was not sufficient to provide adequate coverage of the national NS network for the purposes of this study. While these centroids captured many of the main urban areas, they did not yet include all locations needed to represent the broader structure of the network. Additional centroids were therefore selected based on three further criteria. First, all terminal locations on the NS network were included, such as Den Helder, Heerlen, and Tiel. Second, important transfer nodes were added, particularly where transfers take place between NS services and, in some cases, services operated by other carriers. Examples include Roosendaal, Geldermalsen, and Hoorn. Third, locations associated with some of the busiest stations in the Netherlands were added in order to ensure that major passenger demand centres were also represented. This led to the inclusion of cities such as Gouda and Deventer, among others.

The final set of centroids therefore combines major urban areas, network termini, important transfer nodes, and high-demand rail locations. All selected centroids contain at least one station on the NS railway network and are located within 10 km of a motorway. Den Helder, although not on a conventional motorway, is included as it is the terminus of a main NS line as well as being connected to the motorway network via high-grade N-roads. A map of all selected centroids is shown in Figure 8 The total number of centroids is 46 which leads to $46^2 = 2116$ OD-pairs. A table of all the centroids, their coordinates, reason for inclusion and abbreviation can be found in appendix C.

All ordered OD-pairs are included, this means that A-B and B-A are distinctly different pairs as road closures can affect journeys in either direction independently. Same-origin same-destination pairs are excluded as internal journeys will not be modelled. In a later step it may be found that some OD-pairs have negligible rail relevance because train travel times are uncompetitive or ridership is effectively absent. This distinction will be made later on in the methodology.

### 3.4.2 *Road Closure Identification*

The definition of exposure for this study relies on the presence of road closures on a section of highway within the studied network. Here road closure is defined as the full closure of a highway in one direction for at least a day. Road closures due to road works can often take place at night during low-traffic hours which are incidentally the times that trains do not run. The choice for a full day is made based on the characteristics of road closures and to avoid introducing noise into the data with smaller, unplanned, incidents. One full day is seen as one calendar day.

No single dataset was found that was sufficiently complete and detailed for the purposes of this study. Various datasets which describe planned road works, both in the future and archived, are available, but they are not sufficiently detailed enough for this study. The available datasets lacked a reliable indication of whether closure actually occurred and the archived version did not include the severity levels of the road works. It is therefore necessary to create a new dataset of previous road closures based on other data.

For this study the data chosen is the SHiVi open dataset of the NDW (Nationaal Dataportaal Wegverkeer and Rijkswaterstaat, 2019). This dataset includes the intensities for all roads on the main Dutch highway network (HWN) for all days from 3 January 2019 aggregated from five-minute observations to daily values. The underlying input consists of loop detector measurements, which record vehicles passing over detectors placed at intervals of roughly 300-500 metres. The dataset is organised into SHiVi segments, which correspond to road sections between successive junctions, where a junction may include an on-ramp, off-ramp, road fork, or intersection. Traffic intensity is defined as the average hourly traffic volume on a SHiVi segment over the selected aggregation period. NDW applies a pre-filter to remove technical irregularities from the requested data. Because SHiVi segment intensities are derived from multiple loop detectors, the influence of an individual faulty detector is reduced at the segment level.

Limitations in the export functionality of data from NDW's Dexter portal require a selection of relevant SHiVi sections. This selection is based on the graph network and the centroids, such that at least one segment is included between every centroid and every highway junction. Because SHiVi data is directional, one segment is included per direction and segments are always parallel. A simplified example of selected sections for the A12-highway is presented in figure 5.

SHiVi is used because closures can be inferred from strong reductions or near-zero traffic volumes on affected segments. As the data is aggregated for an entire calendar day, a full road closure in one direction produces near-zero reading in the daily intensity sum for the affected segment. This property makes SHiVi well-suited for detect-

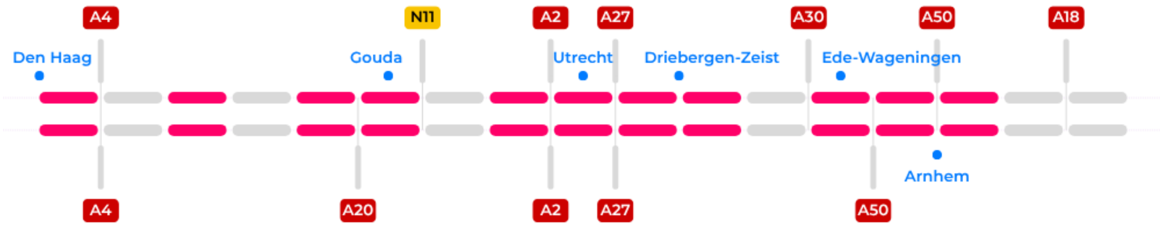


Figure 5: Simplified example of selected SHiVi road segments on the A12 highway. Highlighted in magenta: selected road segments. Grey segments were not selected.

ing the kind of planned, full-day highway closures that are the focus of this study.

*Traffic intensity baseline*

The closure detection equation (eq. 1) requires an estimate of the expected traffic intensity under normal, undisrupted conditions. This baseline intensity $I^{\text{baseline}}_{s,t,m}$ is determined separately for each road segment using a three-tier procedure. The primary baseline (*tier 1*) is the median observed intensity for the corresponding weekday and calendar month. This first tier is computed using all clean observations, that is, days that are not public holidays and whose intensity has not been flagged as suspicious in a preliminary pass. A tier-1 cell is considered reliable when it contains at least three clean observations and its median is at least 20% of the segment-wide weekday median, ensuring that a long closure which dominates a cell does not produce near-zero baselines which would obscure further detections.

When a period contains structurally low-traffic periods such as school holiday, or too few observations of a specific day exist, for example during consecutive weekend closures, a seasonally-adjusted fallback baseline is used (*tier 2*). This fallback is computed as the product of the global weekday median (e.g. all Mondays in 2025) and a monthly factor which is derived from all weekdays in that calendar month. Because the monthly factor is based on a pool of all weekdays in that month, it is stable even when multiple weekend closures take place consecutively. If the monthly factor itself cannot be computed due to insufficient clean observations, the global weekday median is used as a last resort (*tier 3*). The tiered baseline is used only for detection purposes and does not affect the stored event data.

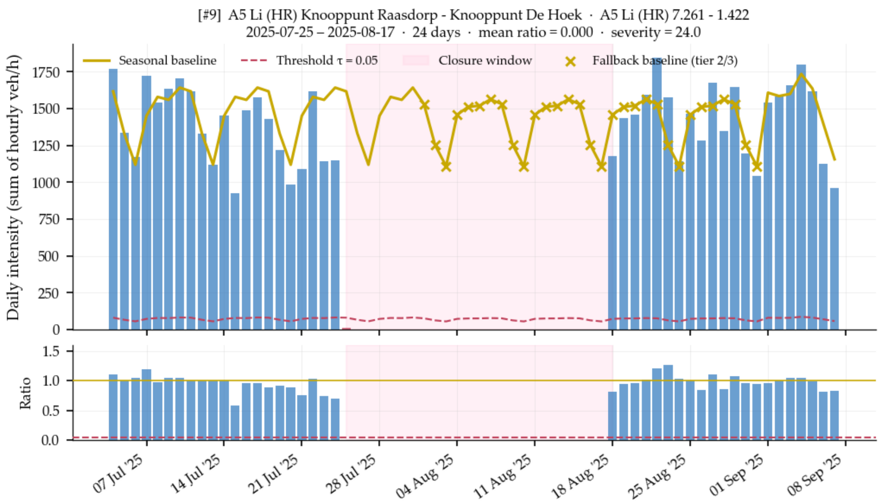


Figure 6: Example of highway intensities including computed baseline, identified road works

Figure 6 shows a graph of observed intensities on a single segment of the A5 in blue with the computed baseline intensities in yellow. The baseline has several X-marks which indicate that the baseline for that day was computed using the tier 2/3 calculations as there was insufficient data to calculate using tier 1. The red dotted line indicates the threshold below which a day is identified as closed. The red-shaded area is a detected road closure period which has been manually confirmed to have taken place.

*Closure detection and filtering*

A close look at the SHiVi data shows that multiple single-day outages are present in the datasets. These were recognised as outages as cross-referencing it with online searches for those closures did not return any results. This is problematic as the detection algorithm would flag it as a road closure. For this reason a minimum duration filter is applied to the detected closure events, requiring that a segment must be closed for at least two consecutive days before the period is classified as a road closure.

Following automated detection, the full set of identified closure events was inspected visually by examining the daily intensity time series for each detected event alongside its surrounding context window. A small number of events were identified as data artefacts rather than genuine road closures. These were all either implausibly long zero-intensity periods or identified road closures which, after manual inspection, did not return any associated road works. This manual inspection is based on a private list of road works obtained from Rijkswaterstaat which summarises all road works and events that took place in a certain week. This dataset was not used for the actual road work identification firstly because it is a private dataset and secondly because the dataset is visual in nature as it is compiled of a series of maps with road works drawn overtop.

$$\text{Flag}_{s,t} = \begin{cases} 1, & \text{if } \frac{I_{s,t}}{I^{\text{baseline}}_{s,t,m}} \leqslant \tau_{\text{closure}} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where:

- $\text{Closure}_{s,t}$ is a binary indicator equal to 1 if segment $s$ is classified as closed on day $t$, and 0 otherwise;
- $I_{s,t}$ is the observed traffic intensity on segment $s$ on day $t$;
- $I^{\text{baseline}}_{s,t,m}$ is the baseline traffic intensity on segment $s$ on day $t$ in month $m$ under normal, undisrupted conditions;
- $\tau_{\text{closure}}$ is the minimum proportional reduction in traffic intensity required for a segment to be classified as closed.

Identifying the road closures starts from the intensity starts with flagging every suspicious day using equation 1 which determines if a road has experienced an intensity drop of $\tau_{\text{closure}}\%$ lower than the baseline intensity for that day. Days with road closures are automatically combined with neighbouring days to form consecutive road closures where only road closures of at least two consecutive days are stored.

Table 5: Output variables of the road closure identification step

| **Variable** | **Description** | **Attribute name in dataset** |
|---|---|---|
| Road closure ID | Unique identifier for each detected road closure event | rc_id |
| Start date | Day on which the road closure starts | start_date |
| End date | Day on which the road closure ends | end_date |
| Closure duration | Total length of the road closure, derived from start and end time | n_days |
| Location / affected segment | Exact spatial location of the closure including direction | verkeersbaan_id |
| Baseline intensity | Expected traffic intensity under normal conditions for the affected segment | mean_baseline |
| Observed intensity | Recorded traffic intensity during the closure period for the affected segment | mean_intensity |
| Intensity reduction | Relative reduction in intensity compared with the baseline intensity | mean_ratio |

The output of this step is a dataset of identified road closures and their relevant characteristics which are summarised in table 5. The start and end time are recorded and used to determine the overall length of the road closure. The exact location is stored and used in the next step to generate alternative paths. The baseline intensity and the recorded intensity are also stored.

Of the 222 closure events detected across the full Dutch highway network in 2024 and 2025, 215 (97.0%) were successfully matched to a geometry in the assignment network and included in the analysis. The remaining 7 events (3.0%) across two road groups could not be matched and were excluded. Four events on the A16 and A20 involve short junction connector segments whose boundary nodes are absent from the assignment network, and six events concern the N200, a parallel service road alongside the A2 that is not included in the network. These exclusions represent a limitation of the current network construction and may result in a small number of genuine closures being omitted from the analysis. The A73 is excluded in its entirety due to a lack of reliable data points. This is acceptable as trains from a different operator run parallel to this highway and therefore it can be expected that people changing modes will take a different operator.

### *Sensitivity analysis*

The choice of $\tau_{closure} = 0.05$ and $d_{min} = 2$ are validated through a sensitivity analysis in which $\tau_{closure}$ is varied from 0.00 to 0.20 while all other parameters are held constant.

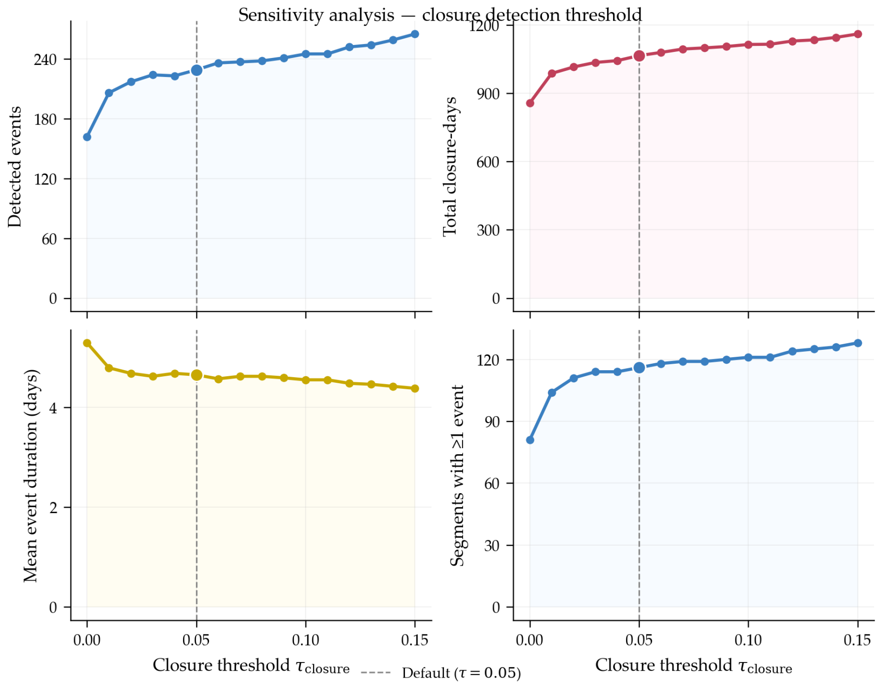


Figure 7: Sensitivity analysis results of intensity threshold

Figure 7 shows the results of the sensitivity analysis. Between $\tau = 0.06$ and $\tau = 0.11$ the number of detected events rises from 236 to 245, an increase of just 9 events. Besides, the mean event duration remains constant at approximately 4.6 days. Below this range the event count rises steeply, consistent with expectations as 100% decreases in intensity are not the norm due to "shoulder-traffic" at the beginning and end of the day disturbing the data. Above $\tau = 0.11$ mean event duration begins to decline, indicating that shorter, weaker signals are entering the dataset. The baseline specification of $\tau_{closure} = 0.05$ sits at the conservative boundary of the stable region, ensuring that only near-total intensity drops are classified as closures while remaining robust to the precise threshold value chosen.

### 3.4.3 *Road Network Construction and Route Generation*

To be able to link road works to node-pairs the road segments which are used to drive between OD-pairs have to be identified. No publicly available dataset was found that directly provides the level of detail and attribute structure required for this study. A directional graph is therefore constructed and used to identify plausible routes between OD-pairs. A more detailed route assignment would be possible, however, for the scope of this research a simplified approach is adequate.

First, a directed, weighted graph of the Dutch highway network is constructed. All A-roads are included and a selection of N-roads is also added to the graph. These N-roads were selected based on road characteristics similar to A-roads including grade-separation, dual carriageways, and ownership. Some N-roads are necessary to complete the network, especially in rural regions, where A-roads might not be continuous as they are interrupted by N-roads, these do not always conform to the characteristics of an A-road. This directed graph consists of nodes, mainly junctions, and edges, representing the roads. Each edge represents a road segment between two nodes with characteristics of the road and a unique identifier which is used to determine where on the network this edge is. Directionality is included to reflect one-way carriageways and direction-specific routing. The graph is based on the Dutch National Road Database (Nationaal Dataportaal Wegverkeer, 2024). The full network is shown in figure 8.

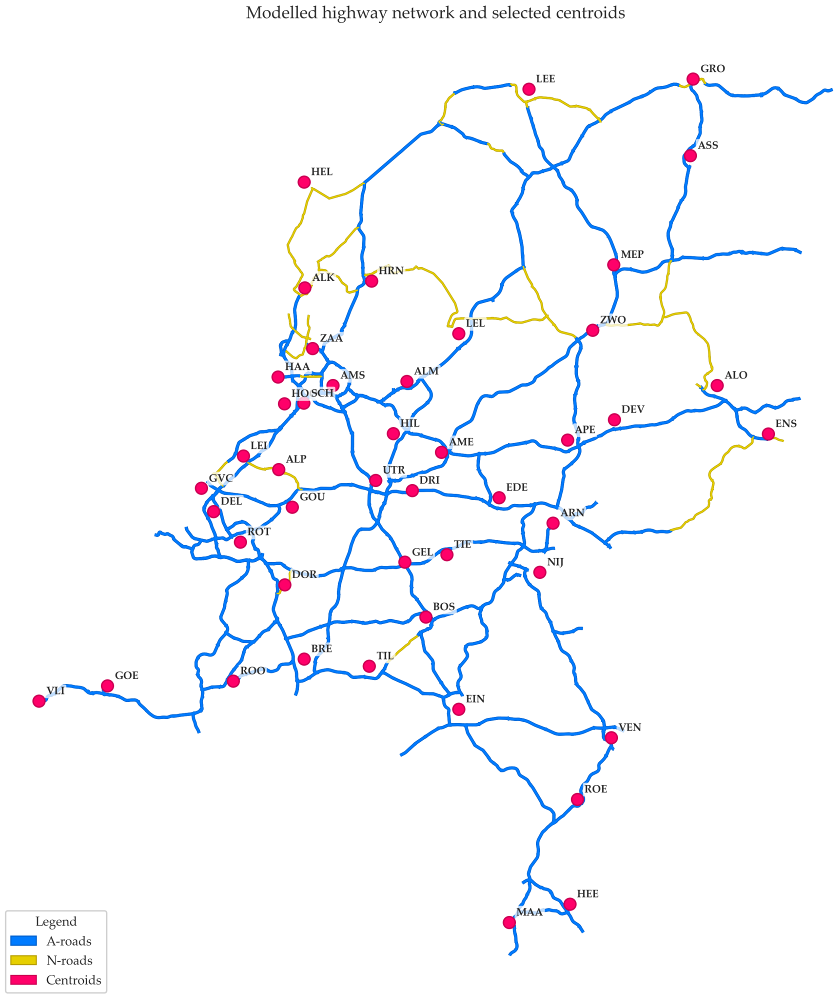


Figure 8: Map of modelled Dutch highway network with all included A- and N-roads as well as the selected centroids (abbreviated)

Weights are given to each road segment to convert the graph from a distance-based measurement to a time-based graph. Time is the main factor in travel behaviour and forms a fair comparison with travel times by public transport. For this conversion speeds under free-flow conditions were assumed which are 100 km/h for A-roads, 80 km/h for N-roads and 30 km/h for access and egress travel. The majority of local access roads in the Netherlands are legally limited to 30 km/h. Most distributor roads are limited to 50 km/h but are subject to signalised intersections. The exact value also has limited sensitivity on the results since access/egress distances in the model are short by design. Manual comparison with travel time estimates from Google Maps suggests that the assumed free-flow speeds produce plausible baseline travel times, generally within 10% of benchmark estimates.

Access and egress are defined as the car travel required to connect each centroid to the modelled highway network. Because only highways are included in the network graph and local roads are excluded, access and egress travel times cannot be derived directly from the graph itself and must therefore be approximated. This is done by calculating the shortest distance from each centroid to the nearest suitable on-ramp and converting that distance into travel time using an assumed average speed of 30 km/h. This represents a simplification of local access conditions and should therefore be taken into account when comparing car travel with the first- and last-mile components of the public transport alternative.

For each OD-pair, the routing method first identifies a set of feasible access and egress nodes on the modelled road network. For each centroid, the nearest access nodes are selected per road, subject to a maximum access distance. All feasible OD access-node combinations are then evaluated, and their total travel time is calculated as the sum of origin access time, network travel time, and destination access time. This produces a ranked set of candidate access-node pairs, from which the fastest combinations are retained for further route generation.

For the best access-node pair, the fastest route is calculated on the directed weighted graph using Dijkstra's algorithm. This route provides the baseline travel time between the two centroids under undisrupted conditions. However, restricting the analysis to only the single fastest route would not adequately reflect driver behaviour, as travellers may also choose slightly longer but still realistic alternatives.

Additional candidate routes are therefore generated using repeated shortest-path search. For a limited number of promising access-node pairs, the algorithm saves shortest simple paths in ascending order of travel time. Candidate paths are retained only as long as their total travel time remains within a predefined threshold relative to the fastest route. In this way, the method identifies a limited set of plausible alternatives without requiring a full enumeration of all possible routes.

To avoid storing multiple versions of effectively the same route, each candidate path is converted into a route family based on its sequence of mainline roads. A route family is a list of consecutive A/N-roads with no regard for distance travelled on each road, e.g. a trip from Den Haag to Eindhoven would have multiple route families including *A12-A13-A16-A58-A2* but also *A12-A2*. Short terminal stubs are removed from this sequence so that minor differences near the beginning or end of a trip do not create artificial distinctions between otherwise similar routes. If multiple candidate paths belong to the same route family, only the fastest representative of that family is retained.

The remaining route families are then filtered further. For shorter trips, alternative routes are retained if they remain within both a relative and an absolute travel-time threshold compared with the fastest route. Finally, highly similar routes are removed using an overlap criterion based on the proportion of shared mainline road length. This ensures that only meaningfully distinct alternatives are kept.

The result is, for each OD-pair, a limited set of plausible route families under undisrupted conditions. These route families form the basis for linking road closures to affected OD-pairs in the subsequent step.

### 3.4.4 *Road Work Exposure Linking and Disrupted Routing*

The last step in the exposure modelling is linking road closures to OD-pairs and subsequently calculating the new fastest travel time when the original undisrupted route is no longer available. This step outputs the road closures, their affected OD-pairs and the new travel times.

For every road closure instance the road segment is compared to the road segments used for any of the calculated shortest-paths for all OD-pairs. If that segment overlaps with one of the paths, that OD-pair is stored and registered as exposed. Next, the affected road segment is removed from the graph and the routing algorithm is recalculated for the new network. This time multiple shortest paths are recorded to account for various detours travellers might take. For each detour, or disrupted path, the travel time is compared to the travel time in the undisrupted condition. This ratio is called the *travel time ratio* and is stored for every detour route to be used in the causal analysis. For example, a ratio of 1.00 means no effect on travel time, a ratio of 1.60 means travel time is 60% higher under the disruption.

## 3.5 CAUSAL ANALYSIS

The goal of a causal analysis is to estimate the effects of planned road works on rail ridership and establish that the effect is indeed causal and not mere correlation. Establishing causality requires identifying changes in ridership that can be attributed to road works rather than to other contemporaneous factors. This is done by comparing the time frame during which the road works took place to a time frame without road works, while controlling for structural differences between origin–destination pairs and for common time-varying factors. In addition, ridership trends for exposed OD-pairs are compared to those of non-exposed OD-pairs during the same time period. By exploiting variation across OD-pairs and over time, and under clearly stated assumptions, this framework allows identification of the causal impact of road works on rail demand.

### 3.5.1 *Data preparation*

Before estimating the causal model, the raw AFC-data must be prepared to form a panel suitable for fixed-effects estimation. A corridor-specific baseline is constructed against which the observed ridership can be normalised, allowing treatment effects to be expressed as proportional deviations from typical demand rather than as absolute passenger counts. Second, an observation threshold is applied to exclude corridors with insufficient data to support reliable fixed-effect estimation.

#### 3.5.1.1 *Baseline ridership*

An initial specification compared ridership during closure periods directly against the same period one year prior. This approach was abandoned in favour of a baseline constructed from the median ridership across all clean observations for each OD-pair × day-type cell, as the year-on-year comparison is sensitive to structural trends in ridership across the study period.

To enable comparison of ridership changes across OD-pairs with structurally different demand levels, observed ridership is normalised against a pair-specific baseline. The baseline for each observation is defined as the median observed ridership for the corresponding (OD-pair, part of day, day of week, month) combination, computed over all clean observations in the study period. Clean observations exclude extended school holiday periods, the summer holiday, Christmas holiday, May holiday, and spring holiday, during which ridership is structurally depressed and not representative of normal travel demand. Single public holidays such as Koningsdag and Hemelvaart are not excluded from the baseline computation, as they occur infrequently enough that the median is robust to their influence; these days are instead controlled for separately in the causal model. Where a given cell contains fewer than four clean observations, a fallback baseline is used computed over the same (OD-pair, part of day, day of week) combination without the month dimension, pooling across all months to increase stability. Normalised ridership is then defined as the ratio of observed to baseline passengers, such that a value of 1.20 indicates ridership 20% above the typical level for that OD-pair, part of day, and day of week.

#### 3.5.1.2 *Observation threshold*

Not all OD-pairs in the network have sufficient observations to produce a reliable baseline and stable fixed effect estimates. Corridors with very few recorded journeys tend to have sparse and volatile normalised ridership values, as a small number of observed trips can produce extreme ratios relative to the baseline. Including these corridors would add noise to the panel without meaningfully contributing to the causal analysis.

To address this, a minimum observation threshold is applied. OD-pairs with fewer than 1,587 total observations across the study period are excluded from the analysis. This threshold corresponds to the median observation count across all OD-pairs in the network, retaining the top 50% of corridors. After applying the threshold, 1,743 OD-pairs remain in the panel. The threshold is applied consistently across all model specifications including the passenger category and fare paid analyses to ensure that the set of corridors entering each model is the same.

### 3.5.2 *Identification strategy*

Estimating the causal effect of road closures on rail ridership requires more than a simple comparison of ridership during closure periods. This would constitute correlation, that ridership tends to be higher when a closure is active, but correlation alone cannot establish causality. The gap between the two is the possibility of confounding where factors exist that simultaneously influence both the occurrence of closures and the level of ridership. Road closures are concentrated in summer months, when ridership patterns differ structurally from the rest of the year. Exposed OD-pairs may also systematically differ from unexposed OD-pairs in ways that can affect ridership independently of any closure, for example in their baseline demand level or the competitiveness of rail on the corridor. A before-after or treated-vs-untreated comparison that does not account for these differences would attribute this to the closures where in fact it is driven by other underlying characteristics.

A Two-Way Fixed Effects (TWFE) design solves this by exploiting variation along two dimensions simultaneously. Entity fixed effects absorb all time-invariant characteristics of each OD-pair, which makes identification come entirely from within-corridor changes over time. Time effects absorb shocks common to all OD-pairs, such as seasonal patterns and day-of-week effects. In this study the time effects are restricted: rather than a fixed effect for every calendar date, common shocks are parametrised through additive day-of-week, month, part-of-day, holiday, and year components. This restricted TWFE specification assumes that common shocks follow the calendar structure of demand. The restriction is tested by re-estimating the model with a full set of date fixed effects as a robustness check (Section 4.4.5). What remains after the fixed effects have been applied is the variation that can be attributed to the closure itself. For this to be valid, one assumption must hold: in the absence of treatment, exposed and unexposed OD-pairs would have followed the same ridership trajectory. This parallel trends assumption and the event study used to assess its plausibility are described in the subsections below.

#### 3.5.2.1 *Parallel trends assumption*

The core requirement for Difference-in-Differences (DiD) causal inference is the parallel trends assumption. This assumption states that, in the absence of treatment, the average outcomes of treated and control units would have followed the same trends over time. This enables the control group to serve as a valid counterfactual for the treated group, such that any post-treatment divergence between the two can be attributed to the treatment itself rather than pre-existing differences in trajectories (Callaway and Sant'Anna, 2021). In the restricted specification used here, the assumption is conditional on the calendar structure. In this, treated and untreated OD-pairs must follow parallel trajectories after accounting for day-of-week, monthly, holiday, and annual patterns. A one-off common shock coinciding with treatment would violate this conditional form. The date fixed effects robustness check relaxes this restriction.

The parallel trends assumption is fundamentally untestable in the post-treatment period, because the counterfactual outcome for treated units is never observed. However, it can be assessed indirectly by examining whether treated and control units followed similar trends *before* treatment began. If pre-treatment trends diverge systematically, this suggests that the two groups were already on different trajctories prior to any treatment, undermining the credibility of the DiD estimate (Roth et al., 2023).

Roth et al. (2023) emphasise that pre-trends has become standard practice in DiD precisely because the parallel trends assumption cannot be verified directly. However, they caution that failing to reject pre-trends does not prove the assumption holds, it only provides evidence consistent with it. The power of pre-trends tests is often limited particularly whent reatment and control groups are similar in the pre-period by design (Roth, 2022). In this study, the event study serves as the primary analysis for assessing whether the parallel trends assumption is plausible. The event study plots estimated coefficients for each relative time period around the closure event, allowing visual and statistical inspection of whether pre-closure trends are flat and do not follow trends. It also serves as a way to infer if any anticipatory behaviour takes place prior to the road closure. These two outcomes contradict each other as a flat pre-trend will mean no anticipatory behaviour takes place and vice versa.

#### 3.5.2.2 *Pre-trends event study*

To assess the plausibility of the parallel trends assumption, an event study is estimated around each road closure event. The event study replaces the binary exposure indicator $E_{ijt}$ with a set of relative-time indicators $D^k_{ijt}$, each equal to one when OD-pair $(i, j)$ is $k$ days relative to the start of a closure event. The estimated coefficients $\hat{\beta}_k$ trace the average ridership deviation for exposed OD-pairs relative to unexposed pairs at each day surrounding the closure. Under the parallel trends assumption, the pre-period coefficients ($k < 0$) should be jointly indistinguishable from zero (Roth et al., 2023).

The event window is set to $k \in [-14, +14]$, covering the two weeks before and after the beginning of a closure. The reference period is $k = -1$, normalising ridership to the day immediately preceding the closure, so all estimated coefficients represent deviations relative to that baseline. For closure events shorter than fourteen days, only the available post-treatment periods are included.

To isolate the closures from preceding or succeeding closures, only isolated closure events are included in the event study. A closure event is classified as isolated if no other closure affects the same OD-pair within a window of 30 days prior to or after the event. This prevents contamination of the pre- and post-period by overlapping treatments, which would bias the time coefficients and make the pre-trends unreliable (Callaway and Sant'Anna, 2021).

Because the severity of a closure may affect whether a ridership response is detectable at all, the analysis is repeated across four minimum detour ratio thresholds $\tau_{min} \in \{1.00, 1.05, 1.10, 1.20\}$. This acts as a filter for any closure–OD-pair combinations where the detour ratio is below $\tau_{min}$. Restricting to more disruptive closures reduces sample size and makes the event study unreliable. The baseline specification uses $\tau_{min} = 1.05$, which retains 836 isolated events across 2580 unique OD-pairs. The sensitivity analysis results are reported in table 11.

### 3.5.3 *Model Specification*

The causal effect of planned road closures on rail ridership is estimated using a restricted TWFE model, with entity fixed effects and calendar time controls. The model exploits variation across OD-pairs and over time by comparing ridership on exposed corridors during closure periods to ridership on the same corridors outside of closure periods, and to ridership on unexposed corridors during the same time periods. Treatment intensity is measured as a continuous variable capturing the severity of disruption to car travel, interacted with duration bins to allow the treatment effect to vary across closure lengths. The following sections describe the estimating equation, the duration bin specification, the entity and time fixed effect structure, and the standard error approach in order.

#### 3.5.3.1 *Estimating equation*

The estimating equation is specified as:

$$\begin{aligned}\tilde{Y}_{ijt} = \alpha_{ij} + \lambda_w + \mu_m + \nu_p + \xi_h + \rho_y + \beta_1 \cdot \mathrm{TI}_{ijt} + \\ \beta_2 \cdot \left(\mathrm{TI}_{ijt} \times \mathbb{1}[d_{ijt} \in [4,7]]\right) + \beta_3 \cdot \left(\mathrm{TI}_{ijt} \times \mathbb{1}[d_{ijt} \geqslant 8]\right) + \varepsilon_{ijt}\end{aligned} \tag{2}$$

where $\tilde{Y}_{ijt}$ is normalised ridership for OD pair $(i, j)$ at time $t$; $\mathrm{TI}_{ijt}$ is treatment intensity; $\alpha_{ij}$ is the entity fixed effect for OD pair $(i, j)$ interacted with part-of-day; $\lambda_w$, $\mu_m$, $\nu_p$, $\xi_h$, and $\rho_y$ are additive fixed effects for day-of-week, month, part-of-day, holiday period, and year respectively; and $\varepsilon_{ijt}$ is the error term. The fixed effect structures are discussed in sections 3.5.3.3 and 3.5.3.4. The reference category is closures of 2–3 days, so $\beta_1$ captures the effect for short weekend closures, while $\beta_2$ and $\beta_3$ capture the additional effect for medium (4–7 days) and extended ($\geqslant$ 8 days) closures relative to that reference.

Normalised ridership is defined as:

$$\tilde{Y}_{ijt} = \frac{Y_{ijt}}{\bar{Y}_{ij,s}} \tag{3}$$

with $Y_{ijt}$ the observed passenger count and $\bar{Y}_{ij,s}$ the median ridership for OD pair $(i, j)$ in day-type cell $s$ over all clean observations.

Treatment intensity is defined as:

$$\mathrm{TI}_{ijt} = (\mathrm{DR}_{ijt} - 1) \cdot \mathbb{1}\left[\mathrm{exposed}_{ijt} = 1\right] \tag{4}$$

with $DR_{ijt}$ the detour ratio for OD pair $(i, j)$ during closure event $t$, and $\mathbb{1}[\cdot]$ the indicator function. $TI_{ijt} = 0$ for all unexposed observations.

#### 3.5.3.2 *Duration bins*

The relationship between closure duration and ridership behaviour is not assumed to be linear or log-linear. Traveller behaviour is expected to change discretely as closures often coincide with set durations corresponding to the days of the week. This is visible in figure 11 which shows a high number of 2-3 day closures and a steep decline as the closures get longer. If the treatment intensity (TI) is included in the function as a continuous variable with a normal or logarithmic duration, e.g. as $TI \times \log(\text{duration})$, it would be asserted that the effect of treatment intensity increases with the logarithm of duration, a specific curve. For this reason a binned interaction specification is preferred over a continuous term, as it allows for the treatment effect to vary freely across duration classes without imposing a functional form on the duration-response relationship.

Three duration classes are defined. The reference category covers closures of 2-3 days, corresponding to the practice of scheduling intensive highway maintenance across weekend windows to minimise disruption to weekday commuters. This category accounts for the large majority of detected closure events. The second bin covers 4-7 days, capturing closures contained within a single working week. The third bin covers 8 days and longer, capturing closures that span at least at least two consecutive working weeks and therefore confront commuters with disruptions across multiple weeks.

The bin boundaries were validated against two empirical checks. First, the distribution of closure instances shows a natural concentration at 2-3 days with a long tail towards the higher durations. Secondly, alternative bins were estimated including breaks after 3, 6 and 13 days, 3 and 9 days, and after 7 days. In all cases the pattern was identical. Bins below 8 days produced non-significant coefficients and bins of 8 days or longer produced large significant positive effects, confirming that the results are not sensitive to the precise lower boundary choice.

The dataset contains one closure exceeding 30 days: a 62-day closure of the Coentunnel near Amsterdam. To verify that this single event does not drive the 8+ result, the preferred specification was re-estimated excluding all OD-pairs exposed to the Coentunnel closure. The coefficient on the 8+ bin was unchanged to four decimal places, confirming that the result is not sensitive to the single extended closure.

#### 3.5.3.3 *Entity fixed effects*

OD-pairs differ structurally in ways that are unobservable, some corridors have high baseline demand because they connect major cities, others are thin because they serve rural areas. In addition, not all

Table 6: Duration bin definitions used in the interaction specification.

| Bin | Duration | Operational interpretation |
|---|---|---|
| Reference | 2–3 days | Weekend maintenance |
| | 4–7 days | Single working week |
| | 8+ days | Multiple working weeks |

cities are created equal, some are large attractors of workforce, others attract mainly students, and many more variations. Without controlling for these differences, a treated OD-pair with high baseline demand could appear to respond more strongly simply because it is busier, not because the closure had a large effect.

Entity fixed effects address this by absorbing all time-invariant characteristics of each corridor. The fixed effect $\alpha_{ij}$ is assigned to each OD-pair $(i, j)$ interacted with part-of-day, such that identification comes purely from within-corridor variation over time — comparing the same corridor to itself before and during treatment rather than across corridors. This ensures that structural differences in demand, distance, and baseline rail attractiveness do not confound the estimated treatment effect.

#### 3.5.3.4 *Time fixed effects*

Rail ridership fluctuates with the time of day, day of week, month of year, and holiday period in ways that affect all OD-pairs simultaneously. A closure that falls on a Monday in January would, without time controls, be implicitly compared against a baseline that includes Saturdays in August which would produce a spurious treatment effect driven by seasonal and weekday patterns rather than the closure itself. Time fixed effects address this by absorbing these common shocks, ensuring that treated observations are compared only against structurally equivalent time periods.

The time fixed effects $\gamma_s$ is implemented as a set of additive dummy variables rather than fully interacted calendar date fixed effects. A fully-interacted specification would assign a separate fixed effect to every calendar date in the panel, absorbing all common time variation including the treatment variation itself. The additive approach controls for systematic patterns without absorbing date-level variation that is correlated with treatment assignment.

Concretely, $\gamma_s$ indexes unique combinations of day-of-week (7 levels), month (12 levels), part-of-day (morning peak, off-peak, evening peak, and weekend), and summer holiday periods (normal, *summer holiday*, and *peak summer holiday*). This last one is included due to the nature of the Dutch school holidays which are staggered based on region. During the *peak summer holiday* weeks all regions are free whereas during the regular *summer holiday* weeks only one or two of the regions have summer school holidays planned. The reference categories are *Monday*, *January*, *Off-peak*, and *Normal period* respectively. This specification is equivalent to the day-type fixed effects used in the baseline construction (section 3.5.1.1), ensuring consistency between the normalisation step and the causal model.

This calendar parametrisation is a restriction relative to the canonical TWFE specification, which includes a fixed effect for every calendar date and thereby absorbs any common shock non-parametrically, regardless of its origin. The restriction reduces the parameter count from one effect per date to 24 calendar components, at the cost of assuming that common shocks follow the weekly, monthly, holiday, and annual structure of demand. A one-off national event affecting ridership on a single date is not absorbed by the calendar terms. To verify that this restriction does not drive the results, the main specification is re-estimated with a full set of date fixed effects, replacing all calendar controls. The results are reported in Section 4.4.5.

#### 3.5.3.5 *Standard errors*

In a panel dataset, observations from the same OD-pair recorded on consecutive days are not independent of each other. Ridership on the AMS→UTR corridor on a Tuesday is likely similar to ridership on the same corridor on the Wednesday of the same week, because both observations share the same underlying corridor characteristics and are subject to the same persistent unobserved factors. Standard OLS standard errors assume independence across all observations and do not account for this within-corridor correlation. As a result, they tend to be too small, making results appear more statistically significant than they truly are (Bertrand et al., 2004).

To correct for this, standard errors are clustered at the entity level which is the OD-pair interacted with part-of-day. Clustering allows for correlation between observations within the same entity over time, without making any assumption about the structure of that correlation. Independence is only assumed across different entities. All reported standard errors, confidence intervals, and significance tests in this study use entity-clustered standard errors, which is standard practice in panel data models of this type (Cameron and Miller, 2015).

### 3.5.4 *Travel time ratio heterogeneity*

OD pairs are divided into quartiles based on the rail competitiveness ratio $TTR_{ij}$, where $TTR_{ij} = t^{car}_{ij}/t^{rail}_{ij}$ is the ratio of baseline car travel time to rail Level of Service (LoS) travel time.

A value of $TTR_{ij} < 1$ indicates that car travel is faster than rail for that OD-pair which is the case for the majority of the OD-pairs (1,505 of 1,533). A value of $TTR_{ij} > 1$ indicates that rail is faster than the car under undisrupted conditions. This comes with the caveat that the comparison is made between car travel time under free-flow conditions and rail travel time as a measure of LoS. This is not a fair comparison as the LoS is weighted by experienced travel time as well as adding the waiting time for the train at the beginning and at transfer points. The decision to use LoS is made because the travel time by train is not only the in-vehicle time and including the waiting times and transfer times is necessary to give the full picture of how long a journey by train would actually take. This also indicates why frequency is such an important factor for travellers when considering which mode they take as a higher frequency will lower the travel time even if the in-vehicle time stays the same.

Besides the waiting time, LoS also assigns weights to the different types of time which the traveller experiences. These weights have been calculated by De Keizer et al. (2015) and are used to calculate the LoS between every train station pair in the country. A minute spent in the train is equal to 1 minute of LoS-time. Transfers cost penalties to the LoS with the reference 1 transfer being equal to no penalty, 0 transfers giving a negative −22.63 minutes penalty, and 2 transfers adding 15.62 minutes of penalty. The duration of this transfer also matters and is calculated by multiplying the transfer time with a weight. This weight is different depending on the duration of the transfer being equal to −1.89 for each minute between 2–5 min. and +1.67 for each minute above 5 minutes. The type of interchange also matters as a cross-platform interchange has no penalty whereas a cross-station interchange adds a 7.22 minute penalty. The waiting time at the beginning of the journey depends on the ideal departure time of the passenger. If the passenger wants to depart at time $x$, every minute that the train departs earlier or later than that ideal departure time is equal to a weight. The weights used are derived from the rooftop model as described in Wardman (2004), following Bates (2003) where the exact values used by NS are not public.

For every OD-pair the LoS travel time is calculated using these weights and stored next to the nominal undisrupted travel time by car. By dividing the car travel time by the LoS train travel time, the Travel Time Ratio, or TTR, is constructed. Although this ratio is comparing free-flow to LoS, it is still deemed useful and fair for this research as the comparison is the same for every OD-pair and is only used to estimate whether a more competitive railway service has an effect on the estimated rail ridership under disrupted conditions.

The theoretical motivation for this heterogeneity analysis follows directly from the random utility framework outlined in Section 2.3. When a road closure increases the generalised cost of car travel, the probability of switching to rail depends on the relative utility of the two modes. On corridors where rail is already close to competitive with car, a marginal increase in car travel time is more likely to tip the traveller towards rail than on corridors where car is substantially faster. The TTR therefore moderates the treatment effect.

Two complementary specifications are used to test this. First, a continuous interaction between treatment intensity and $\log(\mathrm{TTR}_{ij})$ is added to Specification 4:

$$\tilde{Y}_{ijt} = \alpha_{ij} + \lambda_w + \mu_m + \nu_p + \xi_h + \rho_y + \beta_1 \cdot \mathrm{TI}_{ijt} + \sum_{k \in K} \theta_k \left(\mathrm{TI}_{ijt} \times D^k_{ijt}\right) + \theta_{\mathrm{TTR}} \left(\mathrm{TI}_{ijt} \times \log(\mathrm{TTR}_{ij})\right) + \varepsilon_{ijt} \quad (5)$$

where $\theta_{TTR}$ captures the marginal change in the treatment effect per unit increase in $\log(TTR_{ij})$. The log transformation is applied because $TTR_{ij}$ is a ratio variable, for which a log scale is symmetric around TTR = 1 and better reflects the diminishing behavioural response to changes at extreme values of the ratio.

Second, a pooled model is estimated in which OD pairs are assigned to one of four equal-sized quartile groups based on their associated TTR value. Quartile interaction dummies Q2, Q3, and Q4 are added to treatment intensity, with Q1 (the least competitive quartile for rail) serving as the reference:

$$\tilde{Y}_{ijt} = \alpha_{ij} + \lambda_w + \mu_m + \nu_p + \xi_h + \rho_y + \beta_1 \cdot TI_{ijt} + \sum_{k \in K} \theta_k \left(TI_{ijt} \times D^k_{ijt}\right) + \sum_{q \in \{Q2,Q3,Q4\}} \delta_q \left(TI_{ijt} \times \mathbb{1}[q_{ij} = q]\right) + \varepsilon_{ijt} \quad (6)$$

where $\delta_q$ measures the additional treatment effect for OD pairs in quartile q relative to Q1. This specification complements the continuous interaction by allowing the TTR effect to differ non-linearly across the competitiveness distribution and producing coefficient estimates that are directly comparable across quartiles.

### 3.5.5 *Passenger heterogeneity*

Passengers travel with different motives and demographic backgrounds which can have an influence on their travel behaviour and therefore their response to road closures (Yap and Cats, 2022). To capture this heterogeneity a passenger type analysis is performed by estimating the model specified in equation 2 but with a subset of the passengers divided by type. Passenger type is derived from the product used to travel, which serves as a proxy for the traveller's relationship with the operator and their likely travel motive. This categorisation is a generalisation as a single card type may be used for multiple travel purposes. In addition to passenger type, journeys are classified by whether the traveller paid a marginal fare at the time of travel, derived from the combination of card type and time-of-day, to assess whether price sensitivity moderates the modal shift response.

#### 3.5.5.1 *Passenger categories*

Card type is one of the recorded data points in the AFC-data. For the selected OD-pairs in the given time period, roughly 55 types of cards are available whose classification can be used to determine the type of passenger. This card type, although aggregated, can say something about the most likely reason for travel of that specific passenger. Someone who checked-in using a weekend-unlimited pass on a Sunday is likely travelling for social or recreational purposes whereas someone using a company-issued business card during the Tuesday morning peak is probably travelling for work.

The mapping from card type to travel purpose category follows a product-logic argument: each card type is designed for a specific market segment. Business cards, for example, are bought and distributed by employers, making work travel the overwhelmingly dominant use case. Weekend-unlimited passes only provide discounts during the weekend and will thus most likely not be used for work or school travel if used on a weekend day. Student cards are issued exclusively to students actively enrolled in a study programme. This classification was constructed in consultation with NS staff familiar with the subscription portfolio and its market.

In the AFC-data, no distinction is made between a subscription used within or outside of its valid time-period. For this reason, the time-period of the check-in is also included in the data query. By combining the card type and the time-period, it can be inferred if the passenger paid for their journey and, for a few cards, if the journey was likely for purpose A or B. For example, a weekend-free card used on the weekend is likely for social-recreation purposes whereas if it is used on a weekday, the passenger will be paying full-fare and therefore it is classified as an incidental journey.

It should be noted that this mapping constitutes an informed approximation rather than a revealed observation of travel purpose. Wrongful classification is likely at the margins, for example occasional personal use of a business card, which is sometimes allowed by the employer or a passenger who has a weekend subscription because they work on the weekends.

It is also important to note that this classification blurs the line between travel purpose and passenger type. By classifying every trip made using student cards into the same category, this category will say more about who used the card than what their journey purpose was. Classifying business cards under one umbrella does the opposite, it likely says more about the travel purpose than the passenger as anyone could have a business card, but most can only use it for work-related purposes.

Six panels are constructed based on six categories of travel card types. The six categories are:

1. **Incidental**: Passengers paying full fare
2. **Employer-paid commute**: Passenger travelling using a card issued by their employer
3. **Student**: Passengers travelling using a student card
4. **Personal commute**: Passenger travelling using a personal subscription likely used for work journeys
5. **Social / Recreation**: Passengers travelling using a subscription likely used for social / recreative purposes
6. **Other**: Catch-all category for other card types

The distinction is made between *employer-issued* and *personal subscriptions* for work-related travel because it differs in who is ultimately paying for the journey. For business-card travellers, the price is almost always paid for by the employer creating a lower barrier to entry than for the work subscriptions where the employee themselves pays for the trip. International tickets, group tickets, and unrecognised cards among a few others, fall into the *other* category. A full list of all card types and their classifications can be found in appendix D.

#### 3.5.5.2 *Fare paid classification*

Beyond passenger type, journeys are classified based on if the passenger taking the journey paid or not. This does not mean that they evaded fares, rather it looks at if they have a subscription which offers them free travel for their journey or not. A journey is classified as fare-paid if the card type used requires payment for that specific trip by the traveller. For example, the holder of a *Dal Vrij* subscription travels free during off-peak hours but pays during peak hours. The classification therefore depends on the combination of card type and the time period in which the journey took place, and is derived from a lookup table mapping each of the card types to a fare status for each of the four time periods: Morning Peak, Off-Peak, Evening Peak, and Weekend. No distinction is made between a subscription offering free travel where the subscription is paid for by the passenger as opposed to one paid for by the government (student cards) or the employer (business cards).

This distinction is theoretically motivated by the economics of mode choice. A traveller who pays full fare for each rail journey faces a direct cost comparison between rail and car at the time of the trip. When a road closure increases the time cost of car travel, the relative attractiveness of rail increases, potentially tipping the decision in favour of rail. A traveller whose rail journey is already covered by a subscription faces zero marginal cost regardless of the closure, meaning the closure affects only the car side of the cost comparison. The fare paid classification therefore provides a direct test of whether marginal cost sensitivity drives the observed modal shift response.

The fare paid analysis is estimated using the same fixed-effects specification as the main model (equation 2). The dependent variable is a trips-weighted average of normalised ridership across all card types within the fare group for each OD-pair, date, and time period cell. Entity fixed effects and additive time dummies are estimated independently within each fare group.

#### 3.5.5.3 *Passenger heterogeneity model specification*

To account for these new groups, a new model is specified. The panel is split into six and the model is adapted to only include the data for a single passenger category at a time. The comparison with the previous year is also made within the same category meaning that the student passenger numbers for year A are compared to student passenger numbers for year $A - 1$. For this, the same logic is used as in main fixed-effects model meaning that a travel day is compared to a comparable travel day one year before, accounting for fluctuations stemming from holidays and seasonality. For each group a separate baseline is calculated using the same logic as in section 3.5.1.1. The model specification used for this analysis is:

To estimate treatment effects by passenger group, the main fixed-effects specification (Equation 2) is re-estimated on six non-overlapping subsamples, each comprising all observations belonging to a single passenger group $g \in \mathcal{G}$. The estimating equation for group $g$ is:

$$\begin{aligned}\tilde{Y}^{(g)}_{ijt} = \alpha^{(g)}_{ij} + \lambda_w + \mu_m + \nu_p + \xi_h + \rho_y + \beta^{(g)}_1 \cdot \mathrm{TI}_{ijt} + \\ \beta^{(g)}_2 \cdot \left(\mathrm{TI}_{ijt} \times \mathbf{1}[d_{ijt} \in [4,7]]\right) + \beta^{(g)}_3 \cdot \left(\mathrm{TI}_{ijt} \times \mathbf{1}[d_{ijt} \geqslant 8]\right) + \varepsilon_{ijt}\end{aligned} \tag{7}$$

where all variables are defined as in Equation 2, and the superscript $(g)$ indicates that the model is estimated on the subsample of observations for passenger group $g$ only. The normalised dependent variable $\tilde{Y}^{(g)}_{ijt}$ uses a group-specific baseline $\bar{Y}^{(g)}_{ij,s}$, computed as the median ridership for group $g$ on OD pair $(i,j)$ in day-type cell $s$, ensuring that the treatment effect is measured relative to the group's own typical ridership level. Entity and time fixed effects are estimated within each subsample independently. Estimating separate models rather than a single pooled model with interaction terms avoids imposing a common fixed effect structure across groups with different underlying demand patterns, and gives results that are directly interpretable as group-specific treatment effects without reference to a baseline group.

The total treatment effect for group $g$ at duration bin $k$ is:

$$\beta^{(g)}_{\text{total},k} = \mathrm{TI}_{ijt} \cdot \left(\beta^{(g)}_1 + \beta^{(g)}_k\right), \tag{8}$$

where $\beta^{(g)}_k = 0$ for the reference bin (2–3 days), $\beta^{(g)}_k = \beta^{(g)}_2$ for 4–7 day closures, and $\beta^{(g)}_k = \beta^{(g)}_3$ for closures of 8 or more days. Standard errors are clustered at the entity level in all specifications.

This changes the input of the model when calculating the group-specific causal effects, shown in table 7.

Table 7: Disaggregated OD–time dataset by passenger group

| OD | Time | Group | Ridership | Baseline | Exposure |
|---|---|---|---|---|---|
| A–B | $t_1$ | Student | 542 | 520 | 0 |
| A–B | $t_1$ | Work | 120 | 80 | 0 |
| A–B | $t_1$ | Incidental | 67 | 20 | 0 |
| A–B | $t_2$ | Student | 600 | 557 | 1 |
| A–B | $t_2$ | Work | 180 | 97 | 1 |
| ... | ... | ... | ... | ... | ... |

#### 3.5.5.4 *Pairwise Wald tests*

To formally test whether the differences in treatment effects between passenger categories are statistically significant, pairwise Wald tests are conducted on the 8+ day total treatment effects estimated from the separate TWFE models. While the separate models establish whether each category responds significantly to long road closures in isolation, they cannot determine whether the differences between categories are themselves statistically distinguishable. The Wald test addresses this directly.

For each pair of categories $(g, h)$, the null hypothesis is:

$$H_0 : \beta_{\text{total},g} = \beta_{\text{total},h} \tag{9}$$

where $\beta_{\text{total},g} = \beta_{\text{base},g} + \beta_{8+,g}$ is the total treatment effect at the 8+ day duration bin for category $g$. The test statistic is:

$$t = \frac{\hat{\beta}_{\text{total},g} - \hat{\beta}_{\text{total},h}}{\widehat{\text{SE}}_{\text{diff}}} \tag{10}$$

where the standard error of the difference is approximated as:

$$\widehat{\text{SE}}_{\text{diff}} = \sqrt{\widehat{\text{SE}}_g^2 + \widehat{\text{SE}}_h^2} \tag{11}$$

This approximation is conservative as it assumes zero covariance between the two estimates. Since the categories are estimated on separate sub-panels with independent entity fixed effects, the true covariance is negligible, and setting it to zero slightly overstates the standard error of the difference. This means the resulting t-statistics are a little bit smaller than the true values, making it harder to reject $H_0$ and ensuring that significant results are robust to this assumption.

The approximate standard error for the total effect of each category is derived by propagating the standard errors of the two constituent terms:

$$\widehat{\text{SE}}_{\text{total},g} = \sqrt{\widehat{\text{SE}}_{\text{base},g}^2 + \widehat{\text{SE}}_{8+,g}^2} \tag{12}$$

The test is two-sided and evaluated at the significance thresholds of $p < 0.05$, $p < 0.01$, and $p < 0.001$. All pairwise combinations of interest are tested.

### 3.5.6 *Time-of-day heterogeneity*

The final heterogeneity dimension examines at what time during the day the additional ridership occurs. This requires a finer panel than the OD-pair × day structure used so far. The ridership is disaggregated into the four time-of-day categories morning peak, evening peak, weekend, and weekday off-peak and the panel entity becomes the OD-pair $\times$ time-of-day combination, yielding 6,040 entities. The measure which determines the time-of-day category is the exact time that the passenger "checked-in" using their travel pass. The corresponding times are shown in table 8.

Table 8: Time-of-day categories based on AFC check-in time.

| Category | Days | Check-in time |
|---|---|---|
| Morning peak | Monday–Friday | 06:30–09:00 |
| Evening peak | Monday–Friday | 16:00–18:30 |
| Weekday off-peak | Monday–Friday | all other times |
| Weekend | Friday | after 18:30 |
| Weekend | Saturday–Sunday, Holiday | all day |

Normalised ridership is computed per entity, so that each OD-pair $\times$ period combination is compared against its own baseline, and the entity fixed effects absorb structural level differences between periods within the same OD-pair. The estimation equation extends the duration-bin specification by interacting treatment intensity with the time-of-day categories, with the weekday off-peak as the reference category. Total effects per category are recovered as the sum of the base and interaction coefficients, with standard errors from the full coefficient covariance matrix and clustered at the entity level, consistent with the other heterogeneity analyses.

The interaction is estimated for the 8+ day duration bin only, for two reasons. First, the 2–3 day bin cannot be split meaningfully by construction: it consists predominantly of weekend closures, which leaves the weekday peak interaction cells nearly empty. Second, the 4–7 day bin shows no treatment effect in the pooled specification, and splitting a null effect across time-of-day categories cannot recover a signal that is not present in the aggregate; it only produces four noisier nulls. To confirm that this restriction discards no information, a specification interacting the 4–7 day bin is nevertheless estimated as a robustness check, the results of which are reported alongside the main specification in Section 4.4.8.

## 3.6 PREDICTIVE APPLICATION OF THE CAUSAL ESTIMATES

The previous sections describe how the causal effect of planned road closures on rail ridership is identified and estimated retrospectively. This section describes how those estimates are applied to future road closures. Given the specification of a future planned closure, the estimated coefficients are combined with computed exposure and baseline ridership to produce a predicted ridership uplift per OD-pair. This addresses the fifth sub-question, which asks how the estimated causal effects can be used to predict ridership changes during planned road closures. Two components are required. First, a prediction pipeline that translates a closure specification into per-OD-pair uplift predictions with associated uncertainty (Section 3.6.2). Second, an out-of-sample validation design that tests whether the effects estimated on 2024–2025 closures transfer to closures outside the estimation sample (Section 3.6.3). The corresponding results are presented in Section 4.5.

### 3.6.1 *From causal estimates to prediction*

The coefficients estimated in Section 3.5 are average treatment effects over the closures observed in the 2024–2025 estimation sample. Using them predictively therefore rests on the assumption that a future closure with a given treatment intensity and duration induces a ridership response comparable to that of sample closures with similar characteristics. This assumption is plausible within the range of conditions represented in the estimation sample, but it is not guaranteed to hold outside that range, nor for corridor structures that are underrepresented in the sample. Out-of-sample validation therefore serves two purposes at once. It establishes the scope within which the prediction pipeline can be used operationally, and it constitutes an additional robustness check on the causal estimates themselves: if the estimated effects were artefacts of the estimation sample, they would not be expected to reproduce observed ridership responses to closures the model has never seen.

### 3.6.2 *Prediction pipeline*

The pipeline consists of five stages, summarised in Figure 9. Each stage reuses a component developed earlier in this thesis, the pipeline itself introduces no new estimation.

EXPOSURE A future closure is specified by the affected road segments, the closure dates, and the planned duration. The disrupted network is constructed by removing the closed segments from the NWB routing graph, after which baseline and disrupted fastest routes are computed for every OD-pair, using the same routing procedure and free-flow speed assumptions as in Section 3.4. Treatment intensity then follows as $\mathrm{TI} = \text{detour ratio} - 1$, identically to its definition in the causal analysis. One network refinement was required for the validation cases in Zeeland: the provincial N-roads that serve as the natural detour around the A58 were added to the routing graph with a free-flow speed of 70 km/h, calibrated against Google Maps travel times using the same calibration procedure applied to the motorway network.

BASELINE RIDERSHIP Predicted uplift is expressed relative to expected ridership in the absence of the closure. This baseline is constructed per OD-pair and day type from 2025 AFC data, corrected for year-on-year (YoY) ridership growth by comparing the four weeks prior to the closure to the same four weeks one year earlier and calculating the increase. The baseline ridership is then computed as the baseline for 2025 (per OD-pair, per day) plus the YoY ridership increase.

EFFECT For each OD-pair, the predicted relative uplift is the product of the pair's treatment intensity and the estimated total effect for the applicable duration bin, where the total effect is the sum of the base slope and the bin's interaction coefficient, $\hat{\beta}_{\text{tot}} = \hat{\beta}_{\text{TI}} + \hat{\beta}_{\text{bin}}$, and the bin follows from the planned closure duration. Multiplying the relative uplift by the baseline yields the predicted absolute uplift in trips per day.

UNCERTAINTY Confidence intervals on the predicted uplift are obtained by propagating the estimated coefficient covariance matrix, with standard errors clustered at the entity level as in the main specification, through the linear prediction.

EXTRAPOLATION FLAGGING The pipeline flags every OD-pair whose treatment intensity exceeds the 95th percentile of the estimation sample ($\text{TI} = 0.287$) as an extrapolation. For such pairs the transferability assumption is not supported by the data, and predictions should be interpreted as indicative rather than quantitative. This flag is a design rule of the pipeline rather than a post-hoc qualification: it is applied before any comparison with observed outcomes.

### 3.6.3 *Out-of-sample validation design*

The pipeline is validated against three planned closures that took place in the first half of 2026, outside the estimation sample. The cases were selected to span the input domain of the model rather than to resemble the estimation sample as closely as possible. Each case therefore tests a different part of the transferability assumption.

The first case, a closure of the A28, resembles the conditions under which the coefficients were estimated: treatment intensities within the support of the estimation sample and a corridor in which multiple road alternatives distribute the disruption. It tests interpolation: whether the estimated effects reproduce observed ridership responses under sample-like conditions.

The second case comprises a series of weekend closures of the A12. These shorter closures fall outside of the prediction model as none of them take longer than 7 days. The dense network of highways in the Randstad region also provides plenty of detour options that do not significantly increase the travel time between OD-pairs. This combination should not be accurately able to predict a ridership response.

The third case, the closure of the Vlaketunnel on the A58, represents the opposite extreme. The tunnel is the single motorway bottleneck connecting Zuid-Beveland to the mainland, and its closure produced treatment intensities up to $\text{TI} = 0.885$, roughly three times the 95th percentile of the estimation sample. Moreover, the OD-pairs with the highest treatment intensity also carry a disproportionate share

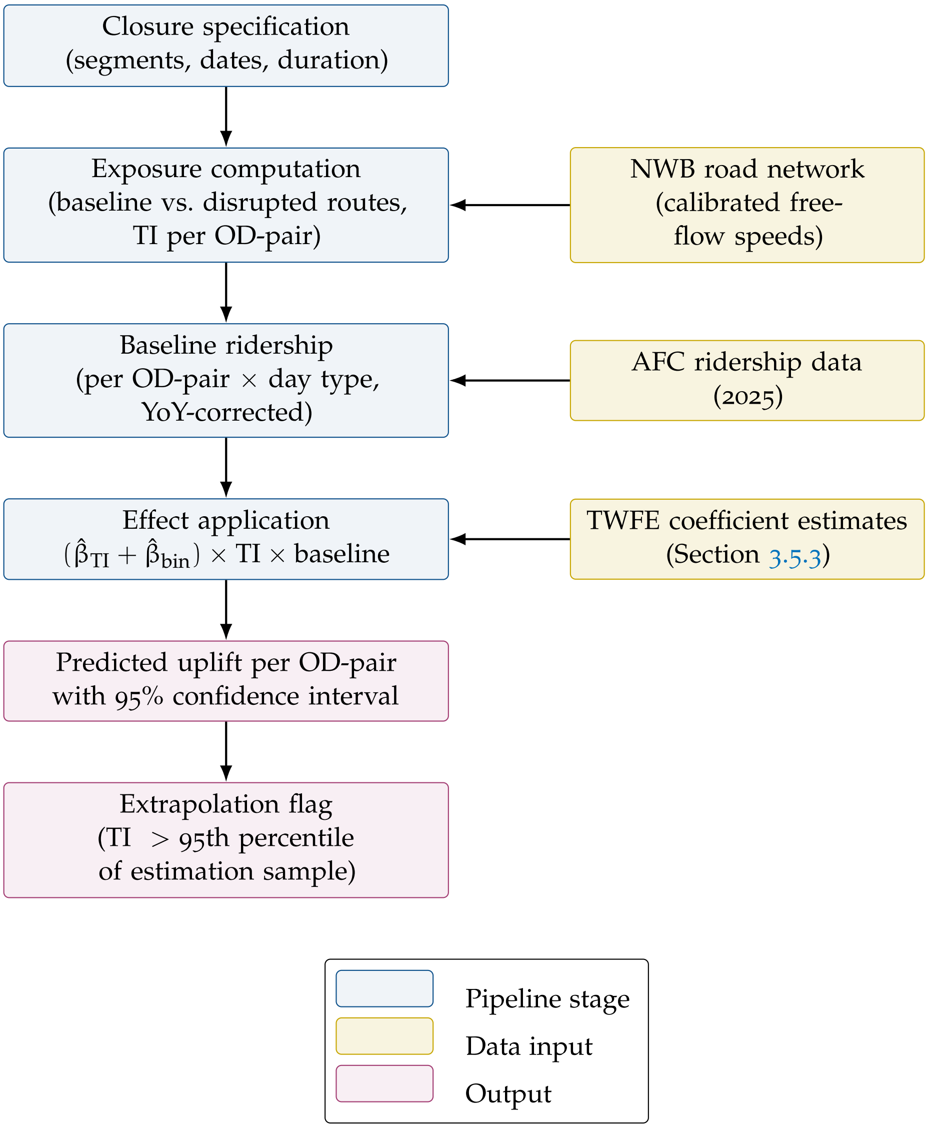


Figure 9: Prediction pipeline translating a planned closure specification into predicted rail ridership uplift per OD-pair. Each pipeline stage reuses a component developed in the causal analysis; no new estimation is introduced.

of demand: the pairs from Goes and Vlissingen towards Breda and Roosendaal account for approximately 21–23% of all mainland-bound ridership from Zeeland. This case tests extrapolation: the behaviour of the pipeline in a single-bottleneck corridor whose exposure lies far outside the estimation sample, where nearly all OD-pairs carry the extrapolation flag defined in Section 3.6.2.

For each case, validation compares the predicted uplift with the observed ridership deviation from the baseline, per OD-pair and aggregated to the closure total. Predictions and observations are compared including their uncertainty: a case is considered consistent with the model when the observed deviation falls within the 95% confidence interval of the predicted uplift. The evaluation criteria are thus fixed before the comparison: within-support cases are expected to reproduce observed responses, low-exposure cases are expected to be dominated by noise, and extrapolation cases are expected to deviate, with the direction and interpretation of that deviation addressed in Section 4.5.

Throughout the validation, the pipeline operates under forward-looking conditions. For each case it uses only information that was available before the closures took place. The validation therefore does not merely test the model, but demonstrates the pipeline in the exact configuration in which it would be applied to future planned closures. The applicability of the model to such closures is established by the validation results themselves, delimited by the scope of validity they reveal.

# 4 RESULTS

In this results chapter, the individual sub-questions will be answered in order to conclusively answer the main research question stated in section 1.2. Some questions will be answered directly by an analysis performed whereas others are answered using a combination of results from different analysis parts. To recap, the sub-questions to be answered and how they are answered are as follows:

- Q1 **Which origin–destination pairs are exposed to planned road closures?:** Answered in section 4.1, the results of the exposure analysis.
- Q2 **What is the estimated causal impact of exposure to planned road closures on rail ridership?:** Answered in section 4.4.1, the global causal analysis.
- Q3 **Which passenger categories exhibit disproportionate increases in rail ridership during road closures?:** Answered in section 4.4.7, the passenger heterogeneity causal analysis
- Q4 **How does the impact of planned road closures on rail ridership vary across different road closure characteristics?:** Answered using characteristics from various analysis phases. Uses a combination of road closure duration (4.4.1), number of affected OD-pairs, detour intensity (4.4.3), and rail competitiveness (4.4.4).
- Q5 **How can the estimated causal effects be used to predict rail ridership during planned road closures?:** Answered by performing validation tests on the obtained estimation results in section 4.5.

This chapter starts with the results of the exposure analysis performed using the methodology as described in section 3.4. As this research created a new dataset for road closures on the highway network in the Netherlands, the new dataset is explored and the important values for the causal analysis are discussed.

Secondly the causal analysis results are presented starting with the results of the main treatment effect based on the road closure duration and the treatment intensity, derived from the detour ratio. After these global results, more specific analysis results are discussed including the heterogeneity of the treatment intensity, rail competitiveness, and passenger types. Lastly, some additional factors which could have an effect on the ridership are discussed.

The last section presents the results of the predictive modelling which is based on the results of the causal analysis. All three cases on which the prediction pipeline is tested are presented and the implications for the predictive capacity of this model are discussed.

All results are presented and interpreted in this chapter. For a global discussion of the implications and limitations of these results, see chapter 5.

## 4.1 ROAD CLOSURE DETECTION

The road closure detection algorithm presented in section 3.4.2 takes as input the SHiVi intensities of 2024 and 2025 aggregated by day. This section describes the resulting dataset. A summary of the results is represented in table 9. The full list of unfiltered road works is available in appendix B.

215 total road closures (events) were recorded of which the majority took place on A-roads. This is an expected figure as N-roads often allow for the capability to reroute against traffic where this is harder to perform on A-roads. 2025 has approximately 25% more events than 2024.

The mean duration of an event was 4.1 days with the median being two days. This mean is inflated by a small number of extended closures, most notably a 62-day closure of the Coentunnel, and ten closures which took longer than a week. By far the most common event type is the two-day closure, which accounts for 159 of the 215 detected events. Of these 159, 150 fall exclusively on weekends, consistent with the Dutch practice of scheduling intensive highway maintenance across Saturday-Sunday windows to minimise disruption to weekday commuters.

Figure 10 shows the monthly distribution of detected closure events across both study years. Activity is low in the winter months, with January and February together accounting for fewer than six events in either year. From March onwards activity increases steadily, peaking in June and July. In 2024 June is the single busiest month with 16 events, while in 2025 both June and July reach 18 events. Notably, August marks a sharp drop relative to the surrounding months in both years, falling to 12 events despite July being the busiest month immediately prior. This dip is consistent with the Dutch *bouwvak*, the three-week summer construction holiday during which road works are prohibited by convention. Activity recovers in August and remains elevated through November before declining again in December.

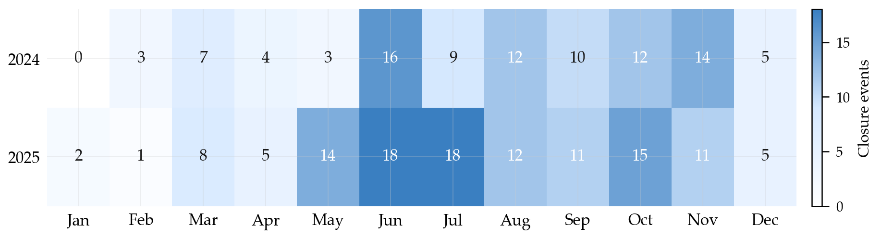


Figure 10: Distribution of road closure events over time. A clear increase in the Summer months is visible with low event counts in the Winter

Table 9: Summary statistics of the detected road closure dataset

| STATISTIC | VALUE |
|---|---|
| *Coverage* | |
| Study period | 2024-2025 |
| Segments monitored | 113 |
| *Closure events* | |
| Total events | 215 |
| A-roads | 208 |
| N-roads | 7 |
| in 2024 | 95 |
| in 2025 | 120 |
| *Closure-days* | |
| Total | 882 |
| weekend | 496 |
| weekday | 386 |

| STATISTIC | VALUE |
|---|---|
| *Duration* | |
| Mean (days) | 4.1 |
| Median (days) | 2 |
| Longest (days) | 62 |
| *Duration distribution* | |
| 2–3 days | 179 |
| 4–7 days | 14 |
| 8–14 days | 9 |
| 15–30 days | 11 |
| >30 days | 2 |
| *Weekend vs. weekday* | |
| Weekend-only events | 150 |
| Weekday events | 65 |
| 2-day: weekend | 150 |
| 2-day: mixed | 9 |

Figure 11 shows the duration distribution of all detected road closure events. As each road closure day is merged with neighbouring closure days, a duration can be calculated based on the consecutive days of closure. Clearly visible is the domination of two-day road closures which is in line with the practice of scheduling road work on weekends. 20 three-day closures took place of which 70% were weekend closures that started on a Friday, one closure which spilled into Monday and the rest were mid-week closures. 14 events took between four and seven days and are often almost week-long closures in the summer months. 13 closures took longer than two weeks, with a large concentration of these events taking place around Amsterdam, both on the A10-ring road and further out.

Plotting the road closures on a map as seen in figure 12 shows a heterogeneous distribution around the country where the events are not concentrated on one road or city but rather spread out evenly with the exception of the North. The sparsely populated and therefore less well connected North of the country sees significantly less full road closures. This could be due to the lower traffic volumes allowing diversions onto the opposite carriageway or more frequent partial road closures. The remaining road closures are shaded by total closure days which shows that most road closures in and around the Randstad area take longer than two days. This could either be a single long closure or multiple weekend closures in succession.

Together, the 215 detected closure events span 113 unique road segments distributed across the national highway network, providing a

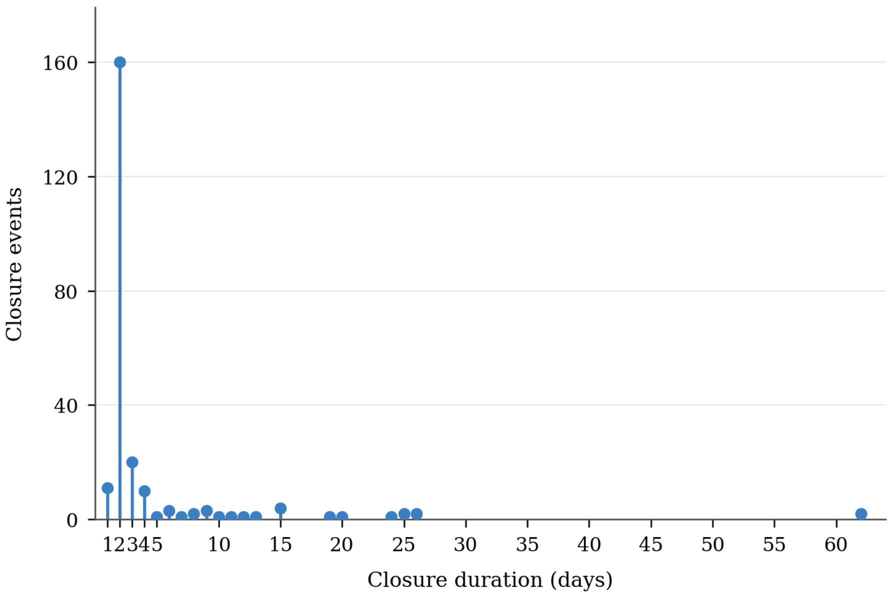


Figure 11: Distribution of road closure duration

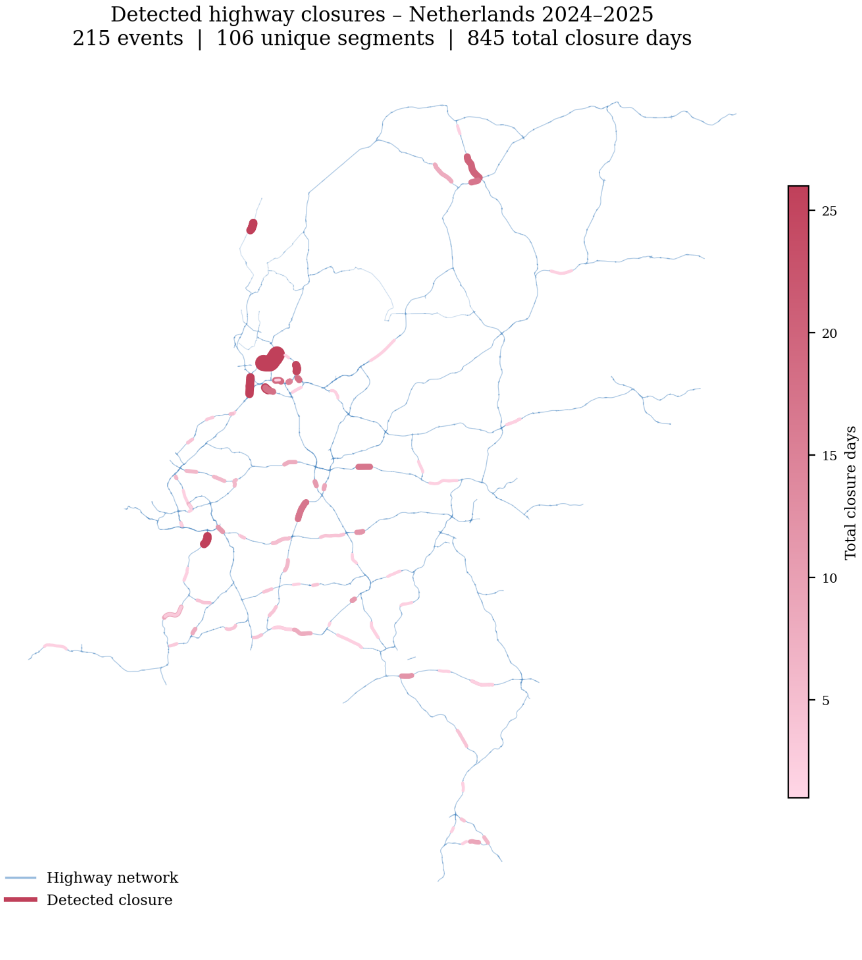


Figure 12: Map of all identified road closure events

geographically diverse treatment dataset for the causal analysis. The concentration of activity in the summer months and the dominance of short weekend closures are both consistent with known Dutch highway maintenance practices, lending face validity to the detection algorithm. The subset of longer closures is of the highest interest as these events could likely lead to a measurable shift in rail ridership.

## 4.2 EXPOSURE MODELLING

The exposure analysis links each of the 91 road segments that experienced at least one closure during the study period to the OD-pairs whose fastest car route passes through that segment, and computes the detour travel time under the disrupted network. A summary of the results is provided in table 10.

*To optimise the code each segment was only evaluated once, even if multiple closures took place on it.*

Table 10: Summary statistics of the exposure analysis output

| Statistic | Value |
|---|---|
| *Coverage* | |
| Unique segments with closure(s) | 91 |
| Roads represented | 28 |
| Unique OD pairs exposed | 1533 |
| OD pairs exposed | 74.1% |
| Unreachable OD pairs | 0 |
| *Travel time* | |
| Mean baseline travel time (min) | 98.7 |
| Mean detour added (min) | 6.7 |
| Mean disrupted travel time (min) | 105.4 |
| Median detour added (min) | 4.2 |
| *Detour ratio* | |
| Mean | 1.0856 |
| Median | 1.0442 |
| 90th percentile | 1.1885 |
| 95th percentile | 1.2868 |
| Maximum | 4.163 |

Of the 2,070 possible ordered OD-pairs in the studied network, 1,533 (74.1%) were exposed to at least one road closure during the study period. This broad coverage reflects the density and interconnectedness of the Dutch main road network. Closures on the A-road system have wide-reaching effects on interurban travel because many OD-pairs share overlapping highway corridors. The 28 roads represented in the detected closure dataset together cover the majority of the national network, leaving only a quarter of OD-pairs entirely unexposed. This spatial variation provides a large and geographically

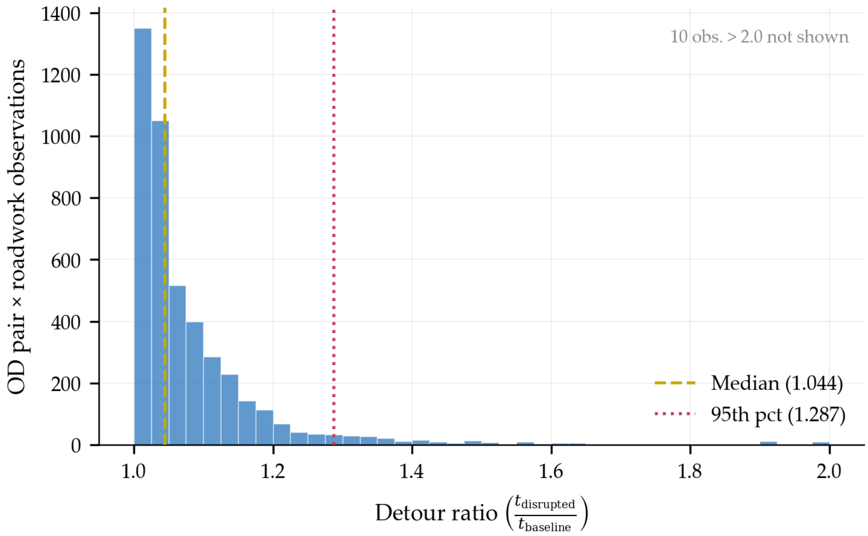


Figure 13: Detour ratio distribution. Higher ratio means a longer detour

diverse set of treated OD-pairs to compare against the untreated remainder. No OD-pair was deemed completely unreachable by any closure event. In every case, at least one viable detour existed within the modelled network, which shows the redundancy of the Dutch highway system.

The detour ratio, defined as the ratio of disrupted to baseline travel time, characterises the severity of each exposure instance. As shown in figure 13, the distribution has a heavy concentration of mild disruptions. The median detour ratio is 1.044, meaning that the typical affected OD-pair experienced an increase in car travel time of approximately 4.4%, equivalent to roughly four minutes on a mean baseline journey of 100 minutes. The mean ratio of 1.086 is pulled upward by a long tail of more severe disruptions: the 90th percentile reaches 1.189 and the 95th percentile 1.287, while the maximum observed ratio of 4.163 corresponds to an extreme case in which the only available detour was more than four times the length of the baseline route, likely caused by a tunnel or bridge closure with no nearby alternative crossing. This is expected as the model runs for every OD-pair including those far apart from each other. As there are over 40 centroids, any road closure will have an effect on a large number of OD-pairs even if the relative distance is far, for these OD-pairs the detour will be short as many alternatives exist. Only for the OD-pairs close to one another does the detour ratio start to get higher.

The road with the greatest number of affected OD-pairs is the A2 (326 pairs), which connects Amsterdam to Maastricht through Utrecht and forms the primary North–South highway axis of the Netherlands. Given this centrality, A2 closures contribute the most to the treated observations in the dataset.

Figure 14 shows the number of OD-pairs affected per road, illustrating that while the A2 is the highest, roads including the A27, A58, A10, and A8 also produce substantial numbers of affect pairs. Figure 15 shows the distribution of detour ratios by road. The severity of disruption varies considerably across roads, reflecting differences in network topology and the availability of alternative routes. Roads with few parallel alternatives, such as those crossing major rivers or passing through urban bottlenecks, tend to produce higher detour ratios than roads in areas with denser alternative routing options.

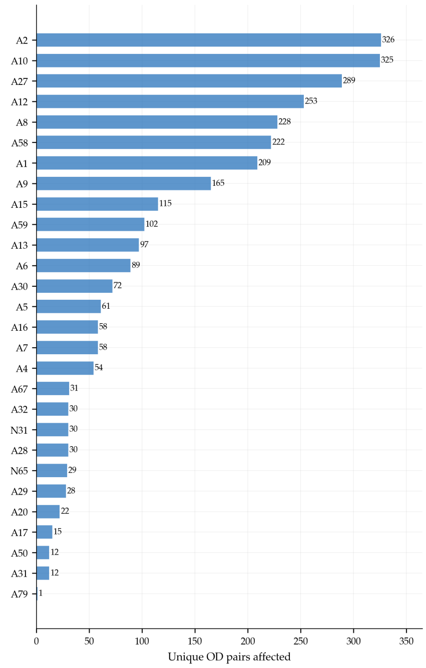


Figure 14: Affected OD-pairs per road

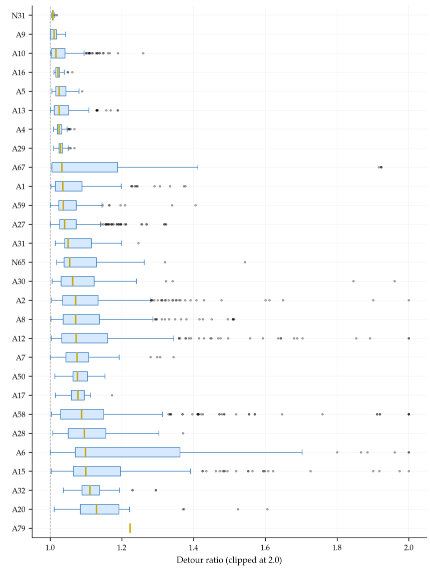


Figure 15: Detour ratio distribution per road

The heterogeneity in both the number of affected OD-pairs and the severity of disruption across roads motivates the road work characteristics analysis in the causal model. The interaction terms in equation 2 are designed to understand this heterogeneity and identify the conditions under which planned road closures are most likely to produce a measurable shift towards rail.

## 4.3 CORRELATION

Before touching causality, correlation between road closures and rail ridership is investigated. Figure 16 shows that there is a weak positive association between closure exposure and ridership as the treated value sits slightly above the control group, however it is not statistically significant. This indicates that pooling together all closures and all treated OD-pairs shows no significant correlation. The dashed line at 1.0 is not the counterfactual as the baseline is a per-cell median while the bar plot shows the mean of a right-skewed ratio, which naturally sits above 1 for both the control and the treated group. The relevant metric is the *difference* between the control and the treated group.

Figure 17 shows a more nuanced result divided into the three duration bins, 2–3, 4–7, and 8+ days. Here the shorter 2–3 and 4–7 day closures still do not show significant results, in contrast to the 8+ day closures which show a significant and positive impact on rail ridership due to road closures. The confidence intervals of the control and treated groups do not overlap and the difference is highly significant ($p < 0.001$).

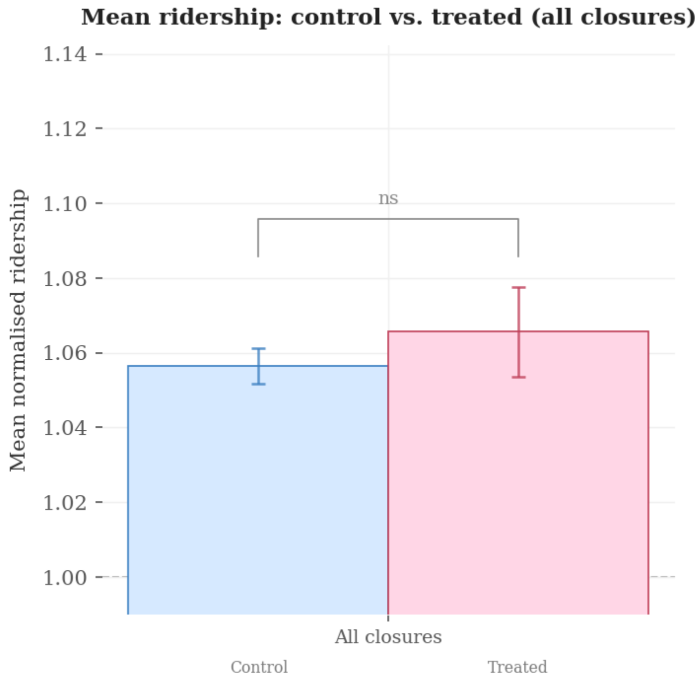


Figure 16: Mean normalised ridership for control versus treated OD-pairs, pooled across all closures. The difference is positive but non-significant.

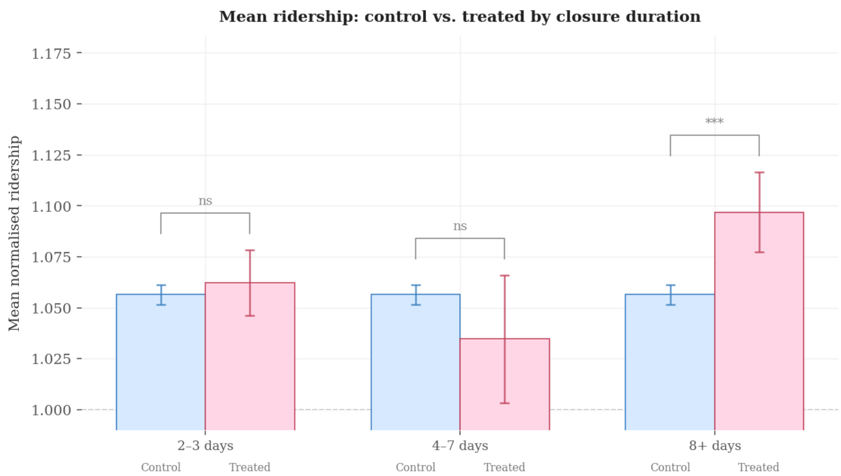


Figure 17: Mean normalised ridership by closure duration bin. Only the 8+ day bin shows a significant difference.

Figure 18 shows the mean deviation of normalised ridership during 8+ day closures on the level of individual OD-pairs. The response is clearly heterogeneous. Very strong positive cells sit alongside strong negative and near-zero cells, with no single uniform pattern across the network. This variation is expected as different corridors serve different travel markets and therefore do not need to respond identically

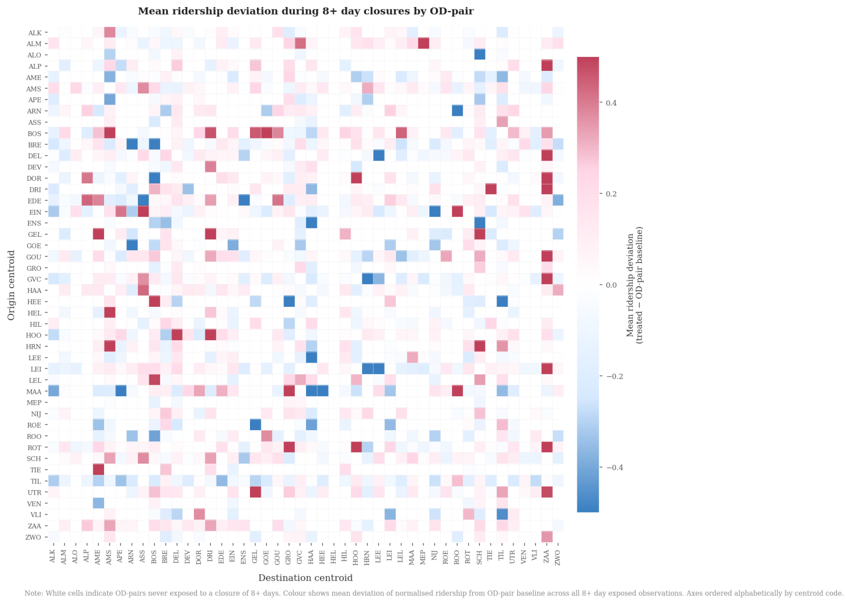


Figure 18: Heatmap of OD matrix normalised ridership values during treated conditions

to a closure. The heatmap does reveal which corridors appear most responsive, which is of direct interest to operators such as NS. Zaandam (ZAA) stands out with a positive deviation toward nearly every destination, possibly reflecting the cluster of long closures around Amsterdam and the Coentunnel through which most of the Zaandam-origin routes pass. The robustness analysis in section 4.4.5 confirms that the aggregate 8+ day effect is unchanged to four decimal places when the Coentunnel closure is excluded, so the result does not depend on it. The most extreme individual cells should be read with caution as they often correspond to OD-pairs exposed to only a small number of long closures.

These correlation results are unconditional, it does not control for time, seasonality, closure severity, the 25% more closures in 2025, or systematic treated-vs-control differences. It also doesn't account for structural differences between OD-pairs as it compares very high ridership to very low ridership journeys. This is why a fixed-effects causal analysis is needed to find if any causality is present. Reassuringly, the 8+ day gap of roughly 4% is close to the effect estimated by the causal model, which lends validity to the main causal findings. This should not be read as evidence that the correlation comparison was sufficient. The following section turns to that analysis.

## 4.4 CAUSAL ANALYSIS

The causal effect of planned road closures on rail ridership is estimated using the restricted TWFE model specified in equation 2 in section 3.5.3.1 applied to a panel of 1,743 OD-pairs observed across the study period. The dependent variable is normalised ridership, defined as the ratio of observed to baseline passengers for a given OD-pair, part of day, and day of week, such that a value of 1.20 indicates ridership 20% above the typical level for that corridor and time slot. This normalisation ensures that treatment effects are estimated relative to each corridor's own demand pattern, and that $\beta$ coefficients are directly interpretable as proportional ridership changes. The key independent variable is treatment intensity, which measures the severity of the car travel time increase imposed by a closure as a proportion of the baseline travel time. Treatment intensity enters the model interacted with three duration bins, 2–3 days, 4–7 days, and 8 or more days, allowing the ridership response to vary both with the severity of the detour and the length of the closure.

The section is structured as follows: the main treatment effect is presented first (4.4.1), followed by a validation of the parallel trends assumption (4.4.2), heterogeneity by treatment intensity (4.4.3) and rail competitiveness (4.4.4), robustness checks (4.4.5), and passenger heterogeneity (4.4.7).

### 4.4.1 *Main treatment effect*

Results from the main causal effect analysis show significant increases in rail ridership during long road closures of 8 or more days, with a higher detour ratio leading to a higher predicted change in ridership (Figure 19). Recall that the reference bin covers 2-–3 day closures, corresponding predominantly to weekend maintenance windows. This reference bin shows a small negative but non-significant ridership response to road closures ($\beta = -0.024$, $p = 0.37$), indicating no detectable ridership response for weekend closures. The 4–7 day duration bin shows a similar response which is also negative and non-significant ($\beta = -0.007$, $p = 0.68$) meaning that week long closures also do not result in detectable changes in rail ridership. The 8+ day bin shows a different result. This duration bin has a positive and significant rail ridership effect ($\beta = 0.375$, $p < 0.001$), indicating that longer closures of more than a week lead to an increase in rail ridership on the affected OD-pairs. Specification 5 adds a continuous interaction between treatment intensity and the log travel time ratio to the duration bin model. The duration bin coefficients remain stable across both specifications, with the 8+ day bin yielding $\beta = 0.397$ ($p < 0.001$) in Specification 5, confirming that the main result is robust to the inclusion of the travel time ratio.

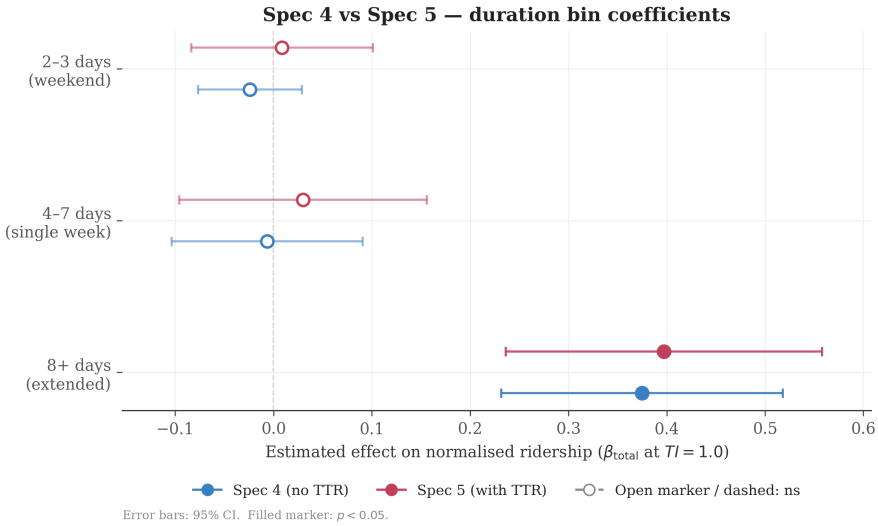


Figure 19: Estimated treatment effect on normalised rail ridership by closure duration bin, with 95% confidence intervals. Effects are reported as $\beta_{total}$ evaluated at $TI = 1.0$, representing a 100% increase in car travel time.

The non-significant results for shorter closures are worth mentioning. These road closures predominantly take place in the weekends on Saturday and Sunday sometimes including Friday or Monday. These days see very different commuting patterns from regular weekdays as they mainly attract leisure travellers, not work or school commuters (Yap and Cats, 2022). These trips are easier to reschedule and can therefore be less likely to push people to change modes (Yap et al., 2018; Van Exel and Rietveld, 2001).

For road closures under a week but longer than a weekend the result is also non-significant and near-zero. Multiple possible explanations exist for this, but are less clear than the explanation of non-significance for weekend closures. Firstly, a single working week might simply not be long enough to force travellers to meaningfully reassess their modal choice. Secondly, planned road works are communicated well in advance which could indicate that travellers respond by adjusting departure times, working from home or simply accepting the delay if the road closure is only for a single week. However, the 4–7 day bin contains relatively few observations compared to the 2–3 day reference. If genuine effects exist for week-long closures they may simply be too noisy to detect given the sample size in that bin. The 95% confidence interval spans from roughly -0.09 to +0.08 indicating that the effects on ridership are very near to zero, this is also a meaningful result and shows that travellers are very likely not to switch to the train during planned road works of 4–7 days but also not likely to walk away from the train service.

The predicted ridership increase scales linearly with treatment intensity within the 8+ day bin (Figure 21). This increase in rail ridership for road closure of longer durations is to be interpreted as follows. The $\beta$ resulting from the model is the increase in normalised rail ridership for the treated OD-pairs at $TI = 1$. A detour ratio of

$DR = 1 + TI$, so $TI = 1$ corresponds to a doubling of car travel time. Thus a journey which would usually take 60 minutes and now takes 78 minutes corresponds to a detour ratio of $78/60 = 1.3$ and therefore $TI = 0.3$, yielding a $0.375 \times (1.3 - 1) = 0.1125$ or 11.25% increase in rail ridership if the road closure takes longer than 7 days.

### 4.4.2 *Parallel trends assumption*

The parallel trends assumption is tested through a pre-trends event study as described in section 3.5.2.2. Under the baseline specification ($\tau_{min} = 1.05$, 836 isolated events, 2,580 OD-pairs), five of the thirteen pre-period coefficients ($k \in [-14, -2]$) are significant at $p < 0.05$. However, the mean absolute pre-period coefficient is small ($\overline{|\hat{\beta}|} = 0.023$), much smaller than the main 8+ day treatment effect of $\hat{\beta} = 0.375$. Following Roth et al. (2023), individual pre-trend significance is not a violation in itself. It is more a question of whether the magnitudes are large enough to threaten the assumption. At 0.023 normalised ridership, the pre-period differences are negligible in magnitude, and the $k = 0$ coefficient is not significant ($p = 0.527$), indicating that travellers are not showing anticipatory behaviour before the closure days themselves.

The results of the pre-trends test are reported in table 11 across four detour ratio thresholds, which were included as an additional sensitivity test. The test is stable across $\tau_{min} \in \{1.00, 1.05, 1.10\}$ where pre-trend magnitudes remain small and the parallel trends assumption is plausible. The $\tau_{min} = 1.20$ threshold is the exception, where a large and significant spike at $k = -14$ ($\hat{\beta} = 0.304$, $p = 0.004$) violates this assumption. This is likely due to the very thin sample at this restrictive detour ratio which reduces the sample to only 252 isolated events and 51,277 observations. This can lead to a single influential closure–OD-pair combination dominating the coefficient at the edge of the event window. Taken together, the event study provides evidence consistent with the parallel trends assumption across the baseline and main analysis specifications.

### 4.4.3 *Heterogeneity by treatment intensity*

Treatment intensity (TI) increases when the detour ratio due to the disruption increases. Results from specification 4 shown in figure 21 indicate that the response of a higher treatment intensity leads to higher rail ridership but only for closures of 8 or more days. For the 2–3 day and 4–7 day bins the total slope on TI is near zero and non-significant ($\beta = -0.024, p = 0.37$ and $\beta = -0.007, p = 0.68$ respectively), indicating that the severity of the detour does not meaningfully alter the, already absent, ridership response for shorter closures. For the 8+ day bin the total slope is $\beta = 0.375$ ($p < 0.001$), reflecting a base slope of $\hat{\beta}_1 = -0.024$ and an interaction term of $\hat{\beta}_3 = 0.399$ meaning that each additional percentage point of car travel time increase translates into a 0.4% increase in normalised ridership. At the

Table 11: Event study sensitivity analysis, parallel pre-trends test across minimum detour ratio thresholds

| | Minimum detour ratio ($\tau_{min}$) | | | |
|---|---|---|---|---|
| | 1.00 | 1.05 | 1.10 | 1.20 |
| Isolated events | 1,514 | 836 | 597 | 252 |
| Unique OD pairs | 3,512 | 2,580 | 1,996 | 960 |
| Observations | 226,571 | 139,218 | 107,374 | 51,277 |
| $R^2$ within | 0.0048 | 0.0055 | 0.0040 | 0.0090 |
| Sig. pre-trend coefs. ($p < 0.05$) | 1 | 5 | 3 | 4 |
| Mean $\|\hat{\beta}\|$ pre-period | 0.012 | 0.023 | 0.024 | 0.046 |
| $k = 0$ coefficient | −0.005 | 0.012 | 0.040* | 0.016 |
| $k = 0$ p-value | 0.726 | 0.527 | 0.098 | 0.617 |
| Parallel trends supported | yes | yes | yes | no |

* $p < 0.10$. Significant pre-trend coefficients are those individually significant at $p < 0.05$ in the pre-period $k \in [-14, -2]$. The $\tau_{min} = 1.20$ threshold is excluded from the main analysis due to sample size constraints and a large significant pre-trend spike at $k = -14$ ($\hat{\beta} = 0.304$, $p = 0.004$). The reference period is $k = -1$ in all specifications.

ridership-weighted median TI of 0.10, which corresponds to a 10% increase in car travel time, this yields a predicted ridership increase of 3.7%.

The modelled relationship is visible in the raw data. Figure 20 plots the observed ridership deviation against treatment intensity for each closure–OD-pair observation. The observed ridership deviation defined as the difference between the observed ridership and the baseline ridership for that OD-pair. For the 2–3 day and 4–7 day bins the scatter shows no systematic trend which is consistent with the non-significance finding in 4.4.1. Due to the high number of observations for these bins, as compared to the 8+ day bin, they were left out of the scatter plot due to cluttering. For the 8+ day bin a positive trend is visible which is confirmed by the trend line. The variance around the trend is substantial, reflecting the heterogeneity of individual OD-pair responses, especially at the lower level of disruption severity. This variance is expected given that the TWFE model controls for entity and time fixed effects whereas the raw deviation plotted here does not. The predicted effect from the fixed-effects mode, shown in figure 21, therefore represents a cleaner estimate of the average causal relationship than the raw scatter alone would suggest.

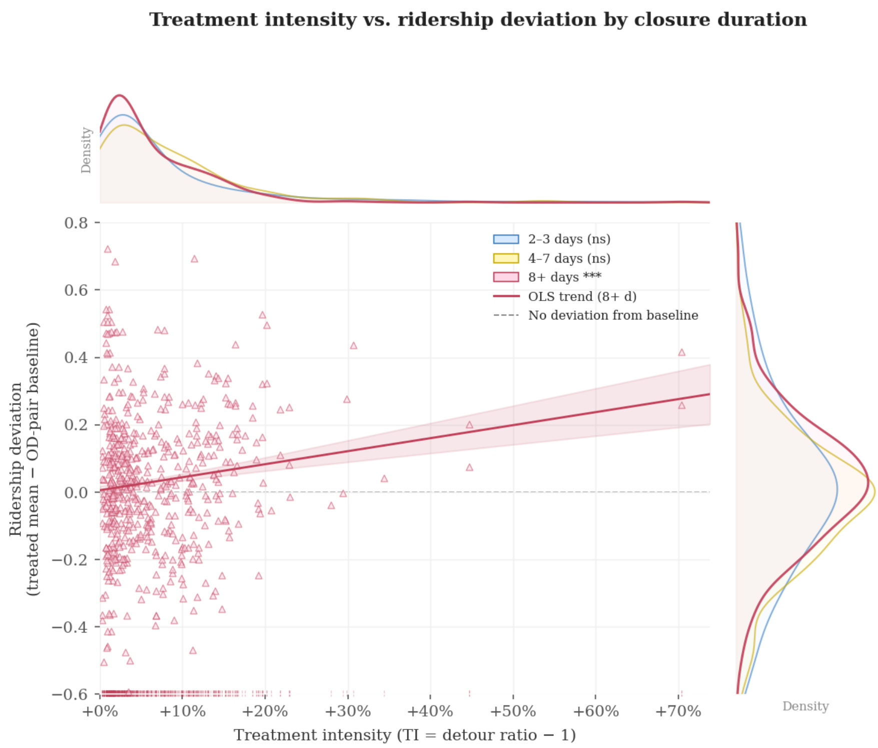


Figure 20: Marginal scatter plot of normalised ridership deviation for treated OD-pairs. Scatter only shows treatment with durations of 8+ days. The ridership deviation is the difference between the treated mean normalised ridership and the OD-pair baseline. The solid line shows the OLS trend for the 8+ day bin; the shaded band is the 95% confidence interval. Marginal density plots are shown on each axis. The positive slope for the 8+ day bin ($p < 0.001$) is visible in the raw data, while the 2–3 day and 4–7 day bins show no trend and are left out due to very high number of observations making graph too cluttered.

The predicted effect scales linearly across the full observed range of TI values, as shown in figure 21. At the 95th percentile TI of 0.287, corresponding to a 29% increase in car travel time, the predicted ridership increase reaches approximately 11.5%. However, this upper range is associated with a wider confidence interval, reflecting the fact that very high detour ratios occur less in the dataset and the estimates are therefore less precise at the extremes. The typical disruption in the dataset has a median TI of 0.044, at which the predicted effect for an 8+ day closure is modest at approximately 1.8%. This indicates that road closures produce relatively mild detours. Large ridership responses require both a long closure duration and a substantial increase in travel time.

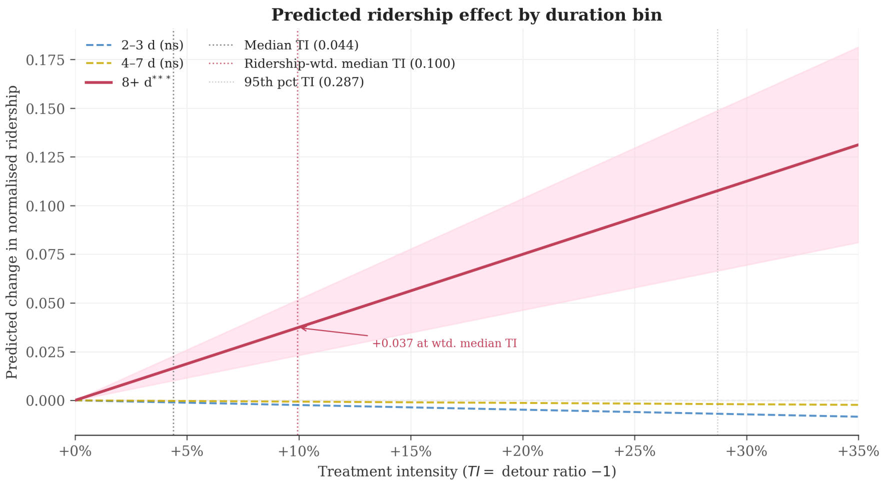


Figure 21: Predicted change in normalised rail ridership as a function of treatment intensity (TI), defined as the detour ratio - 1, by closure duration bin. Solid line indicates 8+ day bin, which is significant. Dashed lines indicate non-significant bins.

The positive and significant slope on treatment intensity for the 8+ day bin is consistent with the theory outlined in chapter 2. Under a random utility framework, a road closure increases the generalised cost of car travel due to the additional travel time imposed by the road closure and subsequent detour necessary. When this additional cost is larger, the relative attractiveness of car travel decreases further, making a mode switch to rail more likely for a greater share of travellers on that corridor. At the ridership-weighted median TI of 0.10, the predicted ridership increase of 3.7% is modest but still meaningful. At the 95th percentile TI of 0.287, this rises to approximately 11.5%, though estimates at this upper range carry greater uncertainty due to the relatively low frequency of high-detour closures in the 8+ day bin.

### 4.4.4 *Heterogeneity by rail competitiveness*

Travel time differences between the car and the train have a slight effect on expected rail ridership with a higher competitiveness for the train resulting in a larger increase in ridership as seen in figure 23. This output is based on specification 5 which uses Travel Time Ratio (TTR) as a continuous variable. Unless stated otherwise, all coefficients reported in this section refer to Specification 4.

Specification 4 creates four quartiles for the TTR which are added to the model as dummies with Q1 as reference each representing a quarter of the TTR for the exposed OD-pairs which gives every individual quarter 282 TTR values. The results from specification 4 are shown in figure 22. These results are very similar to the results from specification 5 indicating robustness of the model.

Road closures with a duration below 8 days do not show significant quartile interaction coefficients ($\delta_q$) which indicates that the influence of TTR on rail ridership for these, already non-significant, road closures is unpredictable. Only quartile 4 is marginally significant ($p < 0.1$). Besides, the value for these durations sits around 0 indicating that the effect is rather minimal. Longer road closures of 8+ days also do not show fully significant $\delta_q$ coefficients, with Q2 and Q3 being non-significant. Quartile 1 is significant ($p < 0.05$) and quartile 4 is marginally significant ($p < 0.1$).

Looking solely at duration bin 8+, it is evident that a more competitive train alternative leads to a higher rail ridership increase when comparing to the overall ridership increase for this duration bin. The effect increases modestly from Q1 ($\beta = 0.341$) to Q4 ($\beta = 0.437$), suggesting that ridership increases more during extended closures on corridors where rail is already a more viable alternative. The quartile interaction term for Q4 is marginally significant ($p = 0.062$); all other quartile interactions are non-significant. Table 12 shows the full values per quartile and duration bin.

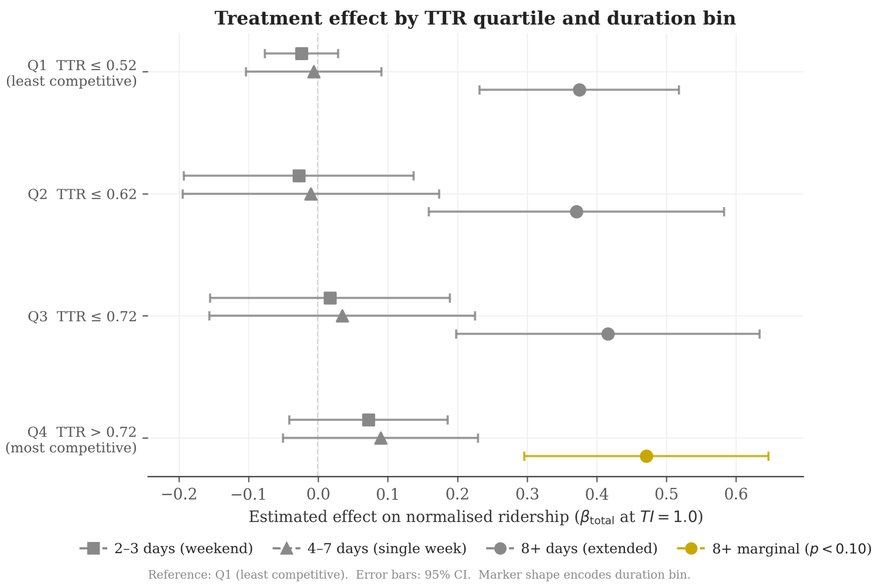


Figure 22: Estimated treatment effects on normalised ridership by TTR quartile and closure duration bin (95% CI). Standard errors are clustered at the entity level.

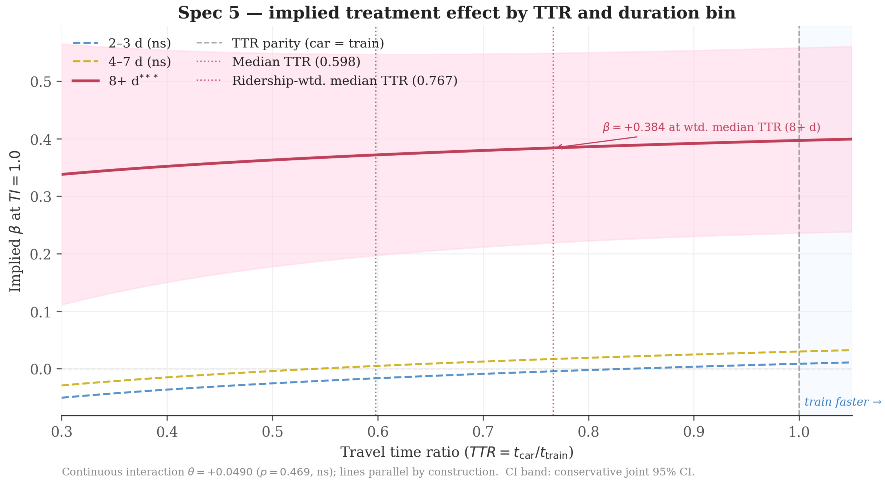


Figure 23: Continuous TTR interaction effect on normalised ridership (95% CI). Vertical lines show the unweighted median TTR (yellow) and the ridership-weighted median (pink). θ is non-significant.

Table 12: Estimated treatment effects on normalised ridership by travel time ratio (TTR) quartile and closure duration bin. The TTR quartile pooled model includes both duration bin interactions and TTR quartile interactions on treatment intensity, with Q1 as the reference quartile and the 2–3 day bin as the reference duration category. Standard errors are shown between brackets. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.001$.

| | | Duration bin | | |
|---|---|---|---|---|
| Quartile | TTR range | 2–3 days | 4–7 days | 8+ days |
| Q1 (ref.) | $\leqslant 0.52$ | −0.040 | −0.015 | 0.341*** |
| | | (0.032) | (0.055) | (0.075) |
| Q2 | $\leqslant 0.62$ | −0.044 | −0.019 | 0.337 |
| | | (0.086) | (0.097) | (0.110) |
| Q3 | $\leqslant 0.72$ | 0.001 | 0.026 | 0.382 |
| | | (0.090) | (0.100) | (0.112) |
| Q4 (most competitive) | $> 0.72$ | 0.056 | 0.082 | 0.437* |
| | | (0.061) | (0.075) | (0.091) |
| $\beta_Q$ (interaction) | | | | |
| Q2 vs Q1 | | −0.004 (0.080), $p = 0.960$ | | |
| Q3 vs Q1 | | 0.041 (0.084), $p = 0.624$ | | |
| Q4 vs Q1 | | 0.096 (0.051), $p = 0.062^{*}$ | | |
| $R^2$ within | | 0.0015 | | |
| Entities | | 4,504 | | |
| Observations | | 2,699,574 | | |

This finding is once again consistent with the findings in the literature which were discussed in the previous section on heterogeneity by treatment intensity (4.4.3). The ratio of travel time by train (TTR) and the relative travel time increase by car (TI) are two sides of the same coin when it comes to the relative generalised cost of travelling by car versus by train, but they play distinct roles. TI captures a *change*: the closure-induced increase in car travel time and thus in the generalised cost of the car alternative. TTR captures a *level*: the baseline ratio of car to train travel time on a corridor, which determines how attractive the rail alternative already is before any disruption occurs.

Under a random utility framework (McFadden, 1974; Ben-Akiva and Lerman, 1985), a traveller chooses the mode that maximises utility, and utility decreases with generalised cost. The share of travellers choosing rail is thus governed by the difference in generalised cost between the two modes: as the car becomes relatively more costly, a larger share crosses the point of indifference and switches to rail. A road closure raises the car's generalised cost by adding VoT $\times$ $\Delta t$ of disutility, where $\Delta t$ is the additional travel time imposed by the detour. This is precisely the mechanism behind Section 4.4.3: a higher TI means a larger $\Delta t$, therefore a larger cost which ultimately leads to a larger induced shift to rail, which is what the positive TI slope in the 8+ day bin reflects.

Rail competitiveness moderates the size of this shift. Mode choice is most sensitive to a change in car cost from TI when the generalised car cost is already close to that of the rail-based alternative. On corridors where rail is more competitive (high TTR), more travellers sit near this margin, so the same closure-induced increase in car time converts more of them than it would on a strongly car-dominated corridor (low TTR), where the car remains the lower-cost option even after the detour. The theory would therefore predict a positive interaction between TI and TTR where the ridership response should be the largest where the detour is severe and rail is already a viable alternative. This is consistent with the results across the TTR quartiles within the 8+ day bin, where the total effect rises from $\beta = 0.341$ in Q1 to $\beta = 0.437$ in Q4.

This finding is, however, statistically weak. The continuous TI $\times$ $\log(\mathrm{TTR})$ interaction ($\hat{\theta}_{\mathrm{TTR}} = 0.049$, $p = 0.469$) is not significant, and only the highest quartile is marginally so ($p = 0.062$). High-TTR corridors are also those where the highest share of travellers are already using the train as opposed to the car, leaving a car population that is rather captive. This smaller pool available to switch may work against the predicted interaction. Overall, the more competitive the train is, the higher the ridership increase will be, but the significance is low and the confidence intervals are large.

### 4.4.5 *Robustness checks*

The main finding is assessed against two robustness checks: sensitivity to alternative duration bin boundaries and the exclusion of the single longest closure event, a 62-day closure of the Coentunnel.

Table 13 reports results under four alternative bin specifications. Across all tested boundaries the long-duration bin is positive, large, and significant at $p < 0.001$, while all shorter bins remain non-significant. The 8–13 and 14+ day bins in the four-bin specification both produce significant effects of similar magnitude ($\beta = 0.340$ and $\beta = 0.343$ respectively), confirming that the effect is stable throughout the extended closure range rather than concentrated at a specific threshold. Merging the two short bins into a single 2–7 day reference does not alter the 8+ day estimate ($\beta = 0.375$). The results are therefore not sensitive to the precise location of the bin boundaries, which supports the interpretation that traveller behaviour shifts somewhere in the multi-week range rather than at an exact day count.

Table 13: Sensitivity of results to alternative duration bin specifications. All specifications use entity fixed effects and additive time dummies. $\beta$ reported at $TI = 1.0$. $^{***}p < 0.001$.

| Bin boundaries (days) | Short (ref) | Medium | Long | Extended |
|---|---|---|---|---|
| 2–3, 4–9, 10+ | −0.023 | +0.038 | +0.324*** | — |
| 2–3, 4–6, 8–13, 14+ | −0.024 | −0.006 | +0.340*** | +0.343*** |
| 2–3, 4–7, 8+ (chosen) | −0.024 | −0.007 | +0.375*** | — |
| 2–7, 8+ | −0.018 | — | +0.375*** | — |

Table 14 reports the chosen bin duration specification alongside one in which the Coentunnel closure is excluded. This closure could in principle drive the 8+ day result on its own due to the very long 62-day duration, for reference, the next longest closure is 26 days as can be seen in figure 11. Excluding all OD-pairs exposed to this closure leaves the coefficients unchanged to four decimal places ($\beta_{total} = 0.375$, $\beta_{bin} = 0.399$, $SE = 0.089$), confirming that the headline result is not driven by a single event.

Table 14: Preferred specification: treatment effects by duration bin (2–3 / 4–7 / 8+ days). Dependent variable: normalised rail ridership. Entity fixed effects and additive time dummies included. Clustered standard errors. $N = 3{,}624{,}153$ observations across 6,040 OD-pair entities.

| Duration bin | $\beta_{total}$ | $\beta_{bin}$ | SE | p |
|---|---|---|---|---|
| 2–3 days (reference) | −0.024 | — | 0.027 | 0.373 |
| 4–7 days | −0.007 | +0.017 | 0.042 | 0.677 |
| 8+ days | +0.375 | +0.399 | 0.089 | 0.000*** |
| *Robustness: Coentunnel excluded* | | | | |
| 8+ days | +0.375 | +0.399 | 0.089 | 0.000*** |

$\beta_{total} = \beta_{base} + \beta_{bin}$, evaluated at $TI = 1.0$ (i.e. a 100% increase in car travel time). $^{***}p < 0.001$.

Replacing the calendar time controls with a full set of date fixed effects leaves the qualitative findings unchanged. The 2–3 and 4–7 day bins remain indistinguishable from zero, and the 8+ day effect remains positive and highly significant ($\beta = 0.226$, $p < 0.001$, against $\beta = 0.375$ in the main specification, with overlapping confidence intervals). Table 15 reports the comparison. The smaller point estimate admits two readings that work in opposite directions. Date fixed effects absorb one-off national demand shocks that the calendar structure misses, which would imply the main estimate is somewhat inflated. At the same time, a date fixed effect is computed as the mean across all entities on that date, including treated ones. On days when large closures treat many OD-pairs simultaneously, part of the treatment effect is absorbed into the date effect, biasing the estimate downward. The two specifications therefore give the upper and lower bounds of the true effect. The out-of-sample validation in Section 4.5 is consistent with this reading: the main specification slightly overpredicted the A28 closure (+8.3% predicted against +6.7% observed), while the date fixed effects estimate would have underpredicted it at +5.0%.

Table 15: Robustness of the main specification to full date fixed effects. Total effects at TI = 1.0 per duration bin; entity-clustered standard errors.

| | Calendar controls | Date fixed effects |
|---|---:|---:|
| 2–3 days (ref.) | −0.024 | −0.007 |
| 4–7 days | −0.007 | +0.006 |
| 8+ days | +0.375*** | +0.226*** |

### 4.4.6 *Ridership effect in passenger kilometre*

To contextualise the estimated treatment effects into terms more suitable for operations, the point estimates obtained in the previous sections are translated into total extra passenger-kilometres. For each exposed OD-pair observation in the 8+ day bin, the model-attributed ridership uplift, $(\hat{\beta}_{TI} + \hat{\beta}_{bin})$ multiplied by the pair's treatment intensity and baseline ridership, is multiplied by the average rail distance between the origin and destination centroids. The reported confidence interval propagates the uncertainty of the coefficient estimates through this calculation.

Only 8+ day closures are included as the shorter closures show near-zero increases or decreases in ridership when aggregated. Across the 17 closure events of eight or more days in the 2024–2025 estimation sample, the model attributes a total of approximately 2.73 million extra passenger-kilometres to planned road closures, with a 95% confidence interval of [1.68M, 3.77M]. The corresponding mean increase per closure event is 160,000 pax-km (median: 112,000 pax-km), equivalent to a mean of 12,800 pax-km per closure day (median: 11,400 pax-km). Relative to the Dutch rail network's annual passenger-kilometre total of approximately 16.5 billion, the estimated annual contribution of the 8+ day road closures amounts to roughly 0.008% of total demand. Therefore the effect is negligible at the aggregate network level.

The aggregate figure includes several closures which dominate the aggregate significantly more than others. The A10 Amsterdam ring road alone accounts for approximately 1.08 million extra pax-km across four closure events, representing 40% of the study-wide total. The seven largest closure events jointly account for 87.4% of the total increase, all occurring on the high-traffic corridors of the A10, A12, and A1. The full concentration is shown in Table 16. This result implies that the relevant operational question for railway operators is not how to respond to road closures in general, but how to identify and prepare for the small number of high-impact events in advance.

Table 16: Concentration of passenger-kilometre uplift across the seven largest long-duration closure events. Combined, these events account for 87.4% of the study-wide total of 2.73 million extra pax-km.

| Rank | Road | Period | Days | Extra pax-km | Cum. share |
|---|---|---|---|---|---|
| 1 | A10 | 27 Jun – 11 Jul 2025 | 15 | 1,000,483 | 36.7% |
| 2 | A12 | 10 – 18 May 2025 | 9 | 287,188 | 47.2% |
| 3 | A1 | 10 – 24 May 2025 | 15 | 274,564 | 57.3% |
| 4 | A7 | 25 Oct – 3 Nov 2024 | 10 | 241,429 | 66.1% |
| 5 | A32 | 19 Sep – 8 Oct 2024 | 20 | 229,028 | 74.5% |
| 6 | A12 | 23 – 30 Aug 2025 | 8 | 225,880 | 82.8% |
| 7 | A27 | 2 – 12 Aug 2025 | 11 | 125,169 | 87.4% |
| *Study-wide total (all closures)* | | | | 2,726,238 | 100% |

### 4.4.7 *Passenger heterogeneity*

Passenger groups as determined in section 3.5.5 seem to respond to the road closures to different degrees. Table 17 shows the individual values per category. Incidental passengers have the highest response rate of 0.770 at a TI of 1 and a closure duration of 8+ days. Passengers travelling with an employer-issued card and students also respond more than the aggregated response with the other three categories (commute self-paid, social/recreation, and others) responding less than the aggregate (Figure 25). Significance is higher for the 8+ duration, however some significance is also present for the shorter duration bins especially with the social/recreation category. All categories have observations counts between $1{,}891{,}432 \leqslant N \leqslant 3{,}227{,}200$.

This heterogeneity between passenger groups is of interest and should be noted as this means that some groups are more likely to use the train during planned road works than others. Acknowledging this heterogeneity can help operators tailor their response to increase comfort for the passengers but also to increase revenue by targeting specific groups. This is most obvious when looking at the incidental passenger group who pay full fare for their journey which might in-

Table 17: Treatment effects by ticket category and closure duration

| | Closure duration bin | | | |
|---|---|---|---|---|
| Category | 2–3 days | 4–7 days | 8+ days | N |
| 1. Incidental | 0.0303 | 0.1827* | 0.7395*** | 3,227,200 |
| | (0.0287) | (0.0714) | (0.0905) | |
| 2. Employer-paid commute | −0.0698** | −0.0220 | 0.4602*** | 2,868,117 |
| | (0.0221) | (0.0662) | (0.0743) | |
| 3. Student | −0.0516 | −0.0360 | 0.5355*** | 3,053,641 |
| | (0.0267) | (0.0612) | (0.0898) | |
| 4. Personal commute | −0.0200 | −0.1017 | −0.0385 | 2,268,843 |
| | (0.0329) | (0.0708) | (0.0834) | |
| 5. Social/rec | 0.0856** | −0.1817* | 0.2003* | 2,750,175 |
| | (0.0281) | (0.0718) | (0.0869) | |
| 6. Other | −0.0427 | −0.2599* | −0.1707 | 1,891,432 |
| | (0.0278) | (0.1079) | (0.1120) | |

*Note:* Standard errors (clustered by entity) in parentheses.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
Dependent variable: normalised ridership (observed / baseline).
All models include entity FE and additive time dummies.
Reference bin: 2–3 days (weekend closures). Separate model per category.

dicate that they only take the train incidentally such as during road closures. This group could be incentivised to take the train as an alternative by pushing specific discounts on affected routes or other marketing strategies. Figure24 illustrates the full duration gradient across all categories

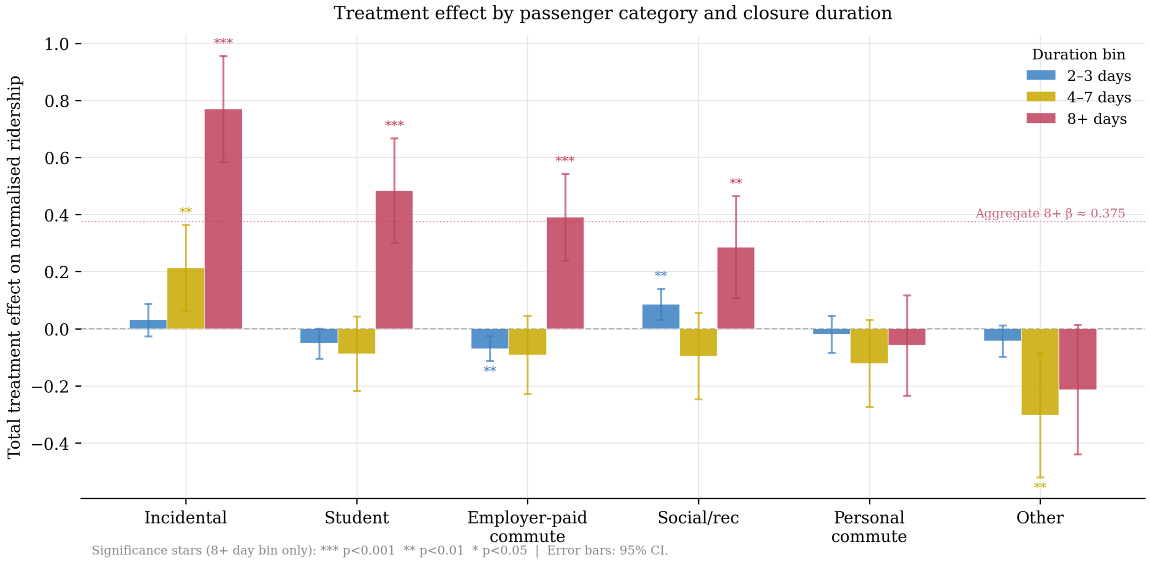


Figure 24: Total treatment effects by passenger category and closure duration bin, estimated from separate TWFE models. Total effect $= \beta_{base} + \beta_{bin}$. Significance stars: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Error bars: 95% CI. Dotted line: aggregate pooled 8+ day $\beta = 0.375$. Categories ordered by 8+ day total effect descending.

Passengers travelling using an employer-issued card and student passengers are also a valuable category, although less for revenue and more for operational response. Employer-issued card passengers often travel in first class and students in second class. This could warrant a different response by planning rolling stock with alternative characteristics such as more room for either one of the travel classes.

For all categories except social/recreational and incidental, a significant positive effect on rail ridership only emerges for closures of 8 or more days. For employer-issued card travellers, this is consistent with a trip flexibility interpretation: shorter closures fall within the margin available to most professional travellers, who can reschedule meetings, work remotely, or suppress the trip entirely. Only once the closure extends beyond a week does at least some in-person travel become unavoidable, producing the observed ridership increase.

For students the mechanism is similar but the constraint is education rather than profession. Short closures may coincide with or be managed by studying from home, whereas multi-week closures eventually require physical attendance. Notably, however, students are among the most likely to already be rail users as many hold free passes meaning their 8+ day response likely stems from the group of students who travel to school by car. 35%-45% of trips by MBO and HBO-students and 20% of trips by WO-students are done by car despite free public transportation travel options (Kennisinstituut voor Mobiliteitsbeleid (KiM), 2026).

Incidental passengers show a somewhat different response as here the 4–7 day duration bin also shows a significant positive influence on rail ridership. This category likely captures a heterogeneous mix including occasional commuters, people providing care for family or friends, and others for whom the trip is not always a choice regardless of closure length.

Social and recreational passengers also make up a rather large share of the trips. They are the only category for which every value is significant at $p < 0.05$ including the short closures on the weekends. This category of traveller is likely to be travelling on the weekends which can be one of the reasons that it has a significant positive, albeit small, trend upwards.

Passengers travelling for work who are paying for their own subscription produce neither significant nor largely positive or negative results. A reason for this could be the catch-all nature of this category stemming from the inclusion of a large set of card types which might be used predominantly by work travellers but could also be used a lot by social-recreational or incidental passengers. For example the off-peak free card type is included in this category which is most often used by work passengers but could just as well be used by retirees travelling for recreational purposes. No clear trend can therefore be inferred for this category.

All groups except for the incidental travellers show at least some negative trends although only the 2–3 day value for employer-paid commutes and the 4–7 day bin for *other* passengers are significant. Nothing can be said about the *other* category as this is a catch-all with too many different card types including international tickets. The negative value for the employer-paid commute is very small and no valid reason for this trend can be found.

*Wald test*

The Wald test results show that every category is significantly different from the incidental passenger group, and all categories except "other" are significantly different from the personal commute group. This means that the difference in response from the incidental group ($\beta = 0.77$) is statistically significantly above all other groups and therefore incidental travellers respond more strongly to long road closures than every other passenger type. The same holds for the personal commute group but in the opposite direction as the corresponding response ($\beta = -0.06$) is lower than the rest of the groups except "other", for which the difference is only marginally significant ($p = 0.049$). This is not surprising as the "other" group is a catch-all category and therefore will show varying levels of response. The full results for the Wald test are in table 18.

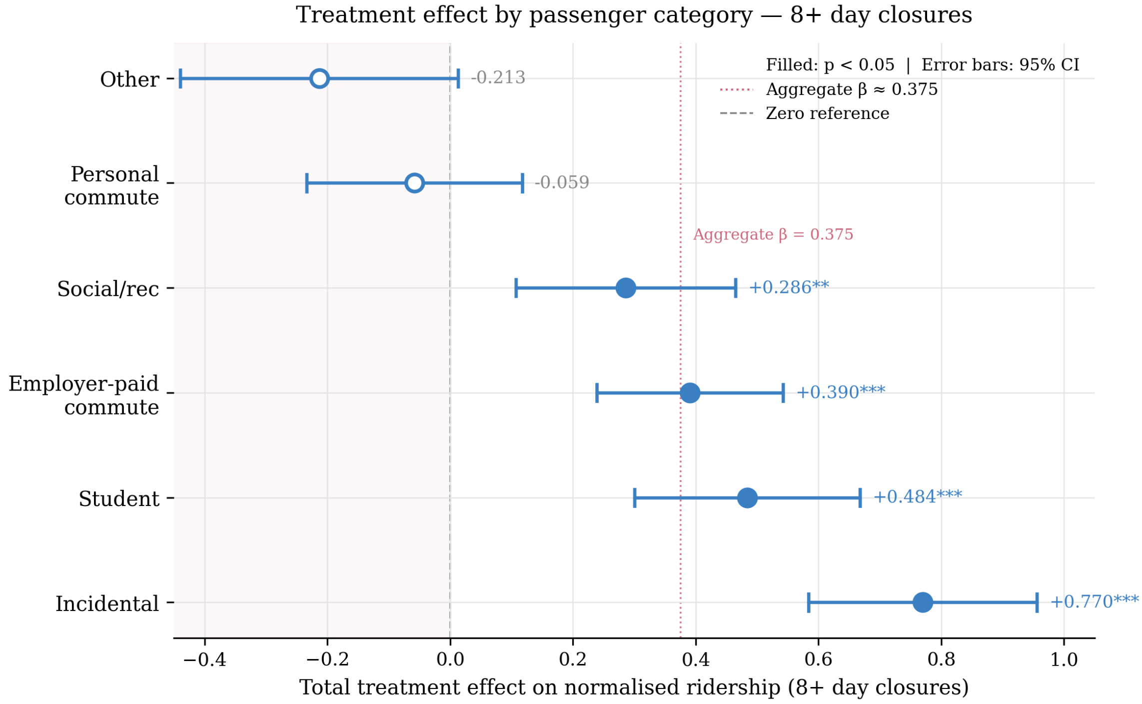


Figure 25: Total treatment effect on normalised ridership by passenger category for 8+ day closures, estimated from separate TWFE models. Total effect $= \beta_{\text{base}} + \beta_{\text{bin}}$. Filled markers: $p < 0.05$. Error bars: 95% CI. Dotted line: aggregate pooled $\beta = 0.375$. Categories ordered by total effect descending.

Table 18: Pairwise Wald tests — 8+ day total treatment effect

| Group A | Group B | $\Delta\beta$ (A–B) | SE | sig |
|---|---|---|---|---|
| *Panel A — Incidental vs all other categories* | | | | |
| Incidental | Employer-paid commute | 0.3795 | 0.1226 | ** |
| Incidental | Student | 0.2860 | 0.1334 | * |
| Incidental | Social/rec | 0.4839 | 0.1318 | *** |
| Incidental | Personal commute | 0.8284 | 0.1306 | *** |
| Incidental | Other | 0.9832 | 0.1495 | *** |
| *Panel B — Personal commute vs middle cluster* | | | | |
| Employer-paid commute | Personal commute | 0.4489 | 0.1185 | *** |
| Student | Personal commute | 0.5424 | 0.1297 | *** |
| Social/rec | Personal commute | 0.3445 | 0.1280 | ** |
| Other | Personal commute | −0.1549 | 0.1461 | ns |
| *Panel C — Within middle cluster (indistinguishable)* | | | | |
| Employer-paid commute | Student | −0.0935 | 0.1217 | ns |
| Employer-paid commute | Social/rec | 0.1044 | 0.1198 | ns |
| Student | Social/rec | 0.1979 | 0.1309 | ns |

*Note:* $H_0$: $\beta_{\text{total,A}} = \beta_{\text{total,B}}$ (two-sided). $\text{SE}_{\text{diff}} = \sqrt{\text{SE}_A^2 + \text{SE}_B^2}$ (conservative — ignores covariance between separate models).
$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, ns = not significant.

### 4.4.8 *Time-of-day heterogeneity*

Figure 26 shows the total treatment effects per time-of-day category for 8+ day closures, estimated on the OD-pair × time-of-day panel described in Section 3.5.6. Table 19 reports the underlying values. The pooled 8+ day effect of $\beta = 0.375$ is not distributed evenly across the day: it is composed of a weekday off-peak response that is statistically indistinguishable from zero ($\beta = -0.013$, $p = 0.91$) and strong, highly significant responses in the morning peak ($\beta = 0.433$, $p < 0.001$), the evening peak ($\beta = 0.615$, $p < 0.001$), and the weekend ($\beta = 0.537$, $p < 0.001$).

Table 19: Total treatment effects by time-of-day category, 8+ day closures. Effects are total effects at TI = 1.0; standard errors clustered at the entity level.

| Category | $\beta_{\text{total}}$ | 95% CI | p-value | Sig. |
|---|---|---|---|---|
| Weekday off-peak (ref.) | −0.013 | [−0.230, 0.204] | 0.907 | |
| Morning peak | 0.433 | [0.207, 0.659] | < 0.001 | *** |
| Evening peak | 0.615 | [0.370, 0.859] | < 0.001 | *** |
| Weekend | 0.537 | [0.275, 0.799] | < 0.001 | *** |

The near-symmetry of the morning and evening peak effects is expected rather than coincidental as a commuter who switches to the train in the morning returns by train in the evening, so the same behavioural decision is expected in both periods. The evening point estimate exceeds the morning one, but the confidence intervals overlap substantially, and the three significant categories cannot be statistically distinguished from one another except from the off-peak reference. The results therefore support the conclusion that the response is concentrated outside the weekday off-peak, not a ranking among the peak and weekend periods.

The flat off-peak response is consistent with the mechanism developed in the previous subsections. Weekday off-peak travel is dominated by discretionary trips with flexible timing; travellers making such trips during a closure can reschedule, reroute, or cancel the journey without switching modes. The travellers with the least temporal flexibility, commuters bound to peak-hour schedules, are precisely those for whom the detour is unavoidable, and their response concentrates in the peaks. The significant weekend effect may appear to contradict the below-average response of social and recreational cardholders found in Section 4.4.7, but the two results measure different things. Passenger categories partition travellers by card product, while time-of-day categories partition trips by when they occur. Weekend trains also carry incidental, full-fare travellers, the group with the strongest estimated response, and a long closure affects their weekend detour just as it affects a weekday commute. The weekend effect is therefore most likely carried by incidental travellers rather than by habitual social and recreational cardholders.

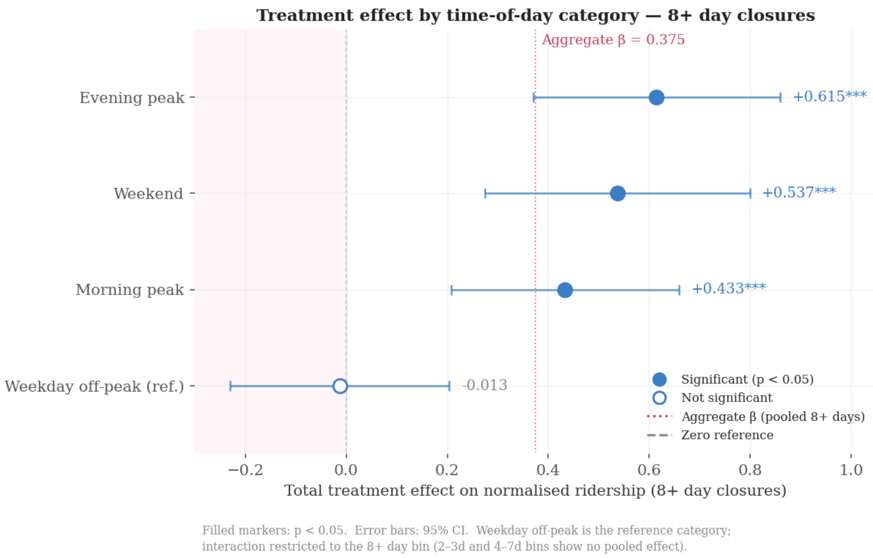


Figure 26: Coefficients of the time-of-day heterogeneity analysis

The robustness specification that additionally interacts the 4–7 day bin confirms that the restriction to 8+ day closures discards no information: all four time-of-day effects in the 4–7 day bin are near zero and far from significance ($p \geqslant 0.73$), as expected when an effect that is null in the aggregate is subdivided.

For operational purposes, the timing of the additional demand matters as much as its size. The extra passengers do not arrive in the off-peak, where spare seating capacity is typically available, but in the peaks, where the train load factors are already at their highest. A given aggregate ridership increase during a long closure therefore translates into a larger capacity challenge than its size alone suggests, which strengthens the case made in Section 4.4.6 for identifying high-impact closures in advance.

## 4.5 PREDICTIVE MODELLING

This section presents the results of the predictive application of the causal estimates described in Section 3.6. The predictive modelling pipeline is tested on three closure cases from 2026, outside of the estimation sample, each run under prospective conditions. The cases jointly answer the fifth sub-question: the causal estimates transfer to unseen closures within the empirical support of the estimation sample. When effects are small, ridership noise dominates and creates conditions which can not be estimated.

Table 20 summarises the validation results for all ten closure cases. For each closure, the predicted uplift is compared with the control-adjusted observed uplift: the observed relative deviation from baseline on treated OD-pairs.

**INTERPOLATION: THE A28 CLOSURE.** The A28 closure between Fluitenberg and Hoogeveen (25 April–7 June, 44 days) is the case most comparable to the estimation sample: an 8+ day closure with

Table 20: Out-of-sample validation of predicted ridership uplift against control-adjusted observed uplift, 2026 closures. Predicted uplift is zero for closures in the 2–3 and 4–7 day bins, for which the estimated total effects are statistically indistinguishable from zero.

| Closure | Period (2026) | Bin | Predicted | 95% CI | Observed[a] |
|---|---|---|---|---|---|
| A12 Gouwe–Bodegraven | 21–22 Mar | 2–3d | -0.5% | [−1.6, 0.6] | +20.5% |
| A12 Gouwe–Bodegraven | 28–29 Mar | 2–3d | -0.5% | [−1.6, 0.6] | +11.7% |
| A12 Bodegraven–Gouwe | 11–12 Apr | 2–3d | -0.5% | [−1.6, 0.6] | +31.0% |
| A12 Bodegraven–Gouwe | 18–19 Apr | 2–3d | -0.5% | [−1.6, 0.6] | +3.4% |
| A12 Oudenrijn–Bodegraven | 2–3 May | 2–3d | -0.5% | [−2.3, 0.9] | −5.2% |
| A12 Oudenrijn–Bodegraven | 9–10 May | 2–3d | -0.5% | [−2.3, 0.9] | +1.1% |
| A12 Gouwe–Oudenrijn | 14–17 May | 4–7d | -0.5% | [−2.6, 2.3] | −12.5% |
| A28 Fluitenberg–Hoogeveen | 25 Apr–7 Jun | 8+d | +8.3% | [+5.1, +11.4] | +6.7% |
| A58 Vlaketunnel (eastbound) | 10–19 Apr | 8+d | +7.3% | [+4.5, +10.1] | +23.3% |
| A58 Vlaketunnel (westbound) | 10–19 Apr | 8+d | +7.8% | [+4.8, +10.8] | +16.9% |

[a] Control-adjusted: observed deviation on treated pairs minus deviation on control pairs.

treatment intensities within empirical support and multiple road alternatives available. The pipeline predicted an uplift of +8.3% [+5.1, +11.4]; the control-adjusted observed uplift was +6.7%, well within the prediction interval. The unadjusted observed deviation (+11.4%) also falls inside the interval, but part of it is accounted for by a +4.6% contemporaneous rise on the control pairs, underlining the importance of the control adjustment. Under sample-like conditions, the causal estimates thus transfer to a closure the model has never seen: the prediction is accurate not only in direction but in magnitude.

THE NOISE FLOOR: THE A12 CLOSURES. The seven A12 closures which consist of six weekend closures and one four-day closure on the Gouwe–Oudenrijn stretch all fall in duration bins for which the estimated total effects are statistically indistinguishable from zero, so the pipeline predicts no uplift, with confidence intervals of roughly ±2 percentage points. The control-adjusted observed deviations instead range from −12.5% to +31.0%, varying in sign and magnitude without any pattern, with a mean of +7.1% across the seven cases. This scatter is not attributable to small samples, baseline volumes on the affected pairs range from roughly 36,000 to 99,000 trips per closure window, but to the shortness of the observation windows. A two-day closure offers two days of data, so ordinary day-to-day variation in ridership (weather, events, incidents) dominates any effect of the size the model predicts. This confirms the expectation formulated in Section 3.6.3 and carries a practical implication: for short closures with low treatment intensity, predicted effects lie below the noise floor of daily ridership, and observed deviations during such closures should not be read as evidence for or against a closure effect.

EXTRAPOLATION: THE A58 VLAKETUNNEL CLOSURE. The ten-day Vlaketunnel closure (10–19 April) is the designed stress test: nearly all affected OD-pairs with high ridership carry the extrapolation flag, with treatment intensities up to TI $= 0.885$, roughly three times the 95th percentile of the estimation sample. The pipeline predicted uplifts of $+7.3\%$ (eastbound) and $+7.8\%$ (westbound); the control-adjusted observed uplifts were $+23.3\%$ and $+16.9\%$ equal to a factor 2 to 3 above the prediction, and far outside the prediction intervals. The direction of the error is consistent with the interpretation developed in Section 4.4: the linear-in-TI effect is estimated predominantly on corridors where alternative routes distribute the disruption, and understates the response in a single bottleneck corridor where no road alternative exists. Two further observations qualify the comparison. First, the OD-pairs with the highest treatment intensity (Goes and Vlissingen towards Breda and Roosendaal) also carry a disproportionate share of demand, so the aggregate error is dominated by exactly the pairs where extrapolation is most severe. Second, the control pairs themselves rose by $+7$–$8\%$ during the closure window, a substantially larger control movement than in the other cases, which may reflect a concurrent NS discount campaign; to the extent that campaign also raised ridership on treated pairs, part of the observed uplift is not attributable to the closure. The Vlaketunnel case therefore does not indicate model failure but heterogeneous treatment effects outside empirical support. In single-bottleneck corridors the pipeline's prediction should be read as a lower bound.

The three cases jointly indicate the domain of validity of the prediction pipeline, and thereby answer the fifth sub-question. Within the empirical support of the estimation sample, the causal estimates transfer out-of-sample with quantitative accuracy (A28). Where predicted effects are small relative to day-to-day ridership variation, predictions are formally correct but operationally uninformative, and short observation windows preclude meaningful validation (A12). Far outside empirical support, in corridors whose structure is unrepresented in the estimation sample, predictions remain directionally correct but should be treated as lower bounds (A58). The estimated causal effects can thus be used to predict ridership changes during planned closures, provided each prediction is accompanied by its position relative to the estimation sample's support which is precisely what the extrapolation flag in the pipeline provides.

For operational application to upcoming closures the pipeline requires only the closure specification and planned duration, both available from Rijkswaterstaat planning well in advance. If the road closures fall within the empirically supported boundaries, the output serves as a reasonable estimate. Elsewhere they serve as lower bounds, especially when the closed road leads to very significant detours. Shorter closures remain difficult to estimate and fall outside of the empirical scope of the predictive modelling.

# 5 DISCUSSION

This chapter interprets the results of this thesis and places them in a broader context. Section 5.1 discusses the findings against the literature and the behavioural mechanisms that drive them, from the duration threshold to the heterogeneity across corridors, passengers, and times of day, and closes with a high-level synthesis of what the results mean beyond the individual estimates. Section 5.2 translates the findings into concrete recommendations for NS. Section 5.4 then discusses the limitations and simplifications of the study and their expected effect on the estimates, and Section 5.5 identifies the research directions that follow from them.

## 5.1 INTERPRETATION OF THE FINDINGS

The main finding of this study shows an increase in rail ridership during road closures with a consecutive duration of longer than one week. This is broadly consistent with prior evidence that long road capacity reductions can induce a modal shift towards public transportation. Kattan et al. (2013) found that a 14-month closure produced a 6.6% public transportation ridership increase which is higher than the median-TI ridership increase ($1.8 - 3.7\%$) estimated in this study. This difference is expected as a 14-month closure forces travellers to fundamentally overhaul their travel routine, whereas the 8+ day closure studied here represents the lower boundary of the duration range at which travellers will begin to reasses their travel habits. The positive direction of the effect is consistent across both studies, reinforcing the theoretical prediction under a random utility framework that a sufficiently large and sustained increase in generalised car travel cost will shift a share of travellers across the threshold from car to train.

The methodological differences between these studies are, however, large. Kattan et al. (2013) relied on aggregate before-after ridership counts without a control group, making it difficult to separate the effect of the closure from trends in travel demand occurring concurrently. Tennøy and Hagen (2021) reported a modal share increase from 40% to 49% during a 14-month tunnel closure in Oslo, but this statistic was derived from stated preference surveys, which are known to overstate socially desirable behaviour such as switching to sustainable public transport (Ben-Akiva and Lerman, 1985). This study addresses both limitations by combining revealed preference AFC data with a two-way fixed effects identification strategy to find the causal effect of the closure by isolating each closure event from time-invariant corridor characteristics and seasonal time trends. It should be noted, however, that while a causal increase in rail ridership is identified, the AFC data does not allow direct observation of which mode the additional passengers switched from. The increase is consistent with a car-to-rail modal shift but this cannot be confirmed without complementary data on road traffic volumes during the same closures or stated preference surveys asking more detailed modal shift-oriented questions. The more conservative effects found here are therefore not a weakness but a direct consequence of a more rigorous and credible identification. Beyond identification, this study also has a different scope. Where previous work relied on single-event case studies (Zhu and Levinson, 2008; Kattan et al., 2013; Tennøy and Hagen, 2021), this study estimates effects for 215 closure events across more than 1,500 exposed OD-pairs for a total of 882 closure days. This allows for conclusions to be drawn about the conditions under which modal shift occurs rather than simply whether it occurs on a specific corridor.

Only road closures with a duration longer than a week show a significant upward trend in rail ridership, a boundary which can be understood by looking at habitual travel behaviour. Mode choice is a habit that is resistant to change even when alternatives are objectively competitive (Gärling and Axhausen, 2003). For disruptions contained within a single working week, travellers have a range of within-mode coping strategies available to them. They could accept the delay as it is only for less than a week, they could adjust their departure time to avoid peak congestion, or work from home for the duration of the road closure (Hillcoat, 2018; Zhu and Levinson, 2012). These strategies allow travellers to temporarily accept the disruption without questioning their modal choice. The robustness analysis supports this interpretation as the specific eight-day boundary carries no special status, as the 8–13 and 14+ day bins in the four-bin specification produce nearly identical effects ($\beta = 0.340$ and $\beta = 0.343$ respectively). What matters is not the precise day count but likely that the disruption extends into a second working week cycle. A traveller who has made use of their coping strategies in the first week might not see them as options any more in the second week. Working from home is a coping strategy which is possible for a short amount of time in most jobs, but not for weeks at a time. This is consistent with Eltved et al. (2021), who showed that sustained exposure across multiple weeks is a precondition for meaningful behavioural adaptation, albeit in the case of rail disruptions. Long road closures therefore operate as what Marsden and Docherty (2013) describe as an opportunity for policy-relevant behaviour change, and the 8+ duration window found in this study defines where the opportunity begins.

The duration threshold carries one further subtlety introduced in the conceptual framework as closure duration does not reach travellers in isolation. Longer closures are announced earlier and communicated more intensively by road and rail authorities. Travellers can only adapt to disruptions in advance if they know about them which means that part of the effect attributed to duration actually stems from the communication. The estimated duration effect is therefore the joint effect of duration and the communication that accompanies it. The practical implication of this is that if announcement practice changes, the threshold itself may move. Communication intensity is not observed in this study, so duration and communication cannot be separated empirically, disentangling them is taken up in the future research section.

The choice of treatment intensity (TI) as the independent variable comes from the mechanism through which road closures affect mode choice. Under a random utility framework, travel time is weighted by the passenger's inherent value of time which translates to the disutility of that journey option (Ben-Akiva and Lerman, 1985; McFadden, 1974). A road closure adds $\Delta_t$ additional travel time due to the detour necessary, increasing the generalised cost of travel by car and reducing its utility relative to rail. TI comes from the detour ratio which is scale-invariant as it indicates how much longer the journey has gotten relative to the original journey time. This is a more accurate representation of the detour time than the raw time increase as a 20 minute detour on a 30 minute journey has relatively more impact than that same 20 minute detour has on a 2 hour journey.

The non-significant results of the 2–3 and 4–7 day durations bins are not an absence of evidence, they tell the story of how people respond to closure of a shorter duration. The 95% confidence interval for both of these bins spans approximately $[-0.09, +0.08]$ providing results which indicate that the response for closures of this duration is low and neither very positive nor negative in nature. Both bins have differing explanations for this finding. The 2–3 day bin is mostly weekend closures during which mainly social and recreational travel takes place. These journeys are often optional and much less likely to get someone to switch modes. Leisure travel in general are more easily rescheduled or cancelled than work or school trips, creating no pressure for modal switching (Yap and Cats, 2022; Van Exel and Rietveld, 2001). For closures of 4–7 days a different mechanism applies. Planned road works are communicated well in advance, giving travellers enough time to adapt by rescheduling meetings, allowing for a different departure time or cancelling trips. For work and school trips, which predominantly take place during the weekdays which are covered by the 4–7 day duration bin, this flexibility is acceptable for a short period of time. Mode switching is the last adaptation people are willing to take being preceded by the aforementioned adaptations (Zhu and Levinson, 2012). Chorus et al. (2006) reinforce this by saying that mode choice is the most habitual component of travel behaviour and the hardest to change.

The detour ratio, which translates to TI, has a significant positive result for closures of 8+ days. By using TI as the independent variable in this study instead of a binary treatment indicator, the results have shown that the ridership increase is not solely dependent on a road being closed, but also on how significant of an effect this has on the traveller's journey, something that a binary specification would have been unable to detect. The positive slope implies that OD-pairs which experience a relatively high detour will see more people take the train than those where the detour is minimal. This is once again consistent with the theory of value of time. At the ridership-weighted median TI of 0.10, the predicted ridership increase is modest at 3.7%, rising to approximately 11.5% at the 95th percentile TI of 0.287. This upper range, however, is associated with wider confidence intervals reflecting the relatively scarce observation count of high-detour closures in the dataset. The typical planned highway closure in the Netherlands therefore produces a real but modest shift in rail ridership, and large responses require the coincidence of both long duration and substantial detour severity.

This modest increase is consistent with what the distribution of TI values in the dataset implies about Dutch highway closures in general. Most of the closures produce relatively mild detours as alternative routes on the highway network are often available. A road closure which forces a genuinely severe detour is rare in practice, meaning that the large $\beta$ values reported at $\text{TI} = 1.0$ serve more as a theoretical anchor than a realistic real-life scenario. Higher TIs also most often happen on short-distance OD-pairs where the original travel time was very short and the detour time is therefore relatively high. Another nuance is the exclusion of local and rural roads in the routing model. This forces journeys between OD-pairs of relatively short distances to use highways even if the most likely alternative is a local road. Take for example the A12 closure between Driebergen and Utrecht. If someone were to take a detour due to this closure they would not take the 50 minute detour over the nearest highway, but rather the 10 minute detour using local roads. Since the original journey time is only 20 minutes, this has a severe effect of overestimating the detour ratio for specific OD-pairs.

It is also worth noting that TI captures only the expected increase in car travel time due to the detour. Road closures additionally introduce variability in travel time reliability, and travellers place significant value on reliability independent of mean travel time (Small et al., 2005; Jamous and Balijepalli, 2018). The true increase in generalised car travel cost during a closure is therefore likely larger than TI alone implies, suggesting that the estimated ridership responses represent a lower bound on the full effect of the road closure. This is a limitation that works in favour of the main finding.

The heterogeneity analysis by rail competitiveness produces a result that is directionally consistent with theory but statistically weak. Under a random utility framework, corridors where rail is already close in relative cost with car, it should produce larger ridership response to the same TI as more travellers sit near the point of modal indifference and are therefore sensitive to a marginal change in generalised cost. The quartile results are consistent with this, the total effect rises from $\beta = 0.341$ in Q1 (least competitive) to $\beta = 0.437$ in Q4 (most competitive) within the 8+ day bin. However, only the Q4 interaction is marginally significant ($p = 0.062$) and the continuous specification is not significant at any level ($p = 0.469$). The evidence is therefore not conclusive although the direction of the trend is in line with expectations from the literature.

Two factors might jointly be pulling the results in opposite directions which results in the large confidence intervals and low significance. Firstly, on corridors where rail is relatively competitive, a large share of car-using travellers sit near the cost-parity and are responsive to the relative cost increase, this would increase ridership. On the other side, highly competitive corridors might already have a large share of travellers making the journey by train and therefore the people travelling by car might be captive or have such a negative view of public transportation that they will likely never switch no matter the circumstances. This self-selection means that the pool of travellers available to switch on highly competitive corridors, think Amsterdam to Utrecht, is smaller than the aggregate TTR value would suggest. Lastly, the TTR ratio is computed in an unfair way by comparing free-flow car time with level-of-service train travel time. This means that the car time will reflect the most optimal condition, which is not realistic given the likely increase in congestion on detour roads due to the road closure. The level-of-service value also contains the time to transfer, frequency of service, and the experienced time when waiting rather than the actual clock-face time. This negatively impacts the LoS compared to the free-flow car time, although as it is used consistently in every comparison it doesn't have an effect on the validity of the results, it just means that the ratio can not be used to accurately determine the real travel time ratio between OD-pairs. The directionality consistency with the literature reassures that the results are valid, but stronger conclusions on rail competitiveness as a moderator require richer input data.

The variation in ridership response across passenger categories suggests two different moderators drive rail ridership increase. For incidental passengers, they might not have a choice to travel or not, which forces them to buy a full-fare ticket as they usually travel by car. Their mode choice is therefore more responsive to a change in relative cost than that of car users whose behaviour is anchored by routine. For employer-paid and student subscriptions, a different mechanism applies because their journey is free at the point of use which means that the financial barrier to switching is structurally absent. When the car alternative becomes worse due to a closure, the train becomes the dominant option without requiring any out-of-pocket expenses. The near-zero response among personally paid commuter subscriptions may not reflect an unwillingness to switch but rather that this group already uses the train regularly on the affected corridor, leaving little additional ridership to be generated by the closure. It could also stem from the methodology used to classify cards. The personal-subscription category houses a lot of card types that could also be used by social and recreational travellers as well as more incidental passengers which could make this category more of a catch-all than it was intended to be.

The out-of-sample validation gives meaning to the causal estimates in a different way than the estimation results alone can achieve. The correct estimation of the A28 closure within a few percentage points (+8.3% predicted against +6.7% observed after control adjustment) demonstrates that the estimated effects are not artefacts of the 2024–2025 estimation sample but capture regular behaviour that transfers to new closures. This is a stronger form of validation than the within-sample robustness checks, as no amount of specification testing can rule out that a model has merely fitted its own sample. At the same time, the validation indicates where the applicability ends. The A12 cases show that, for short closures, the predicted effects lie below the noise floor of day-to-day ridership variation so that neither the prediction nor the validation carries meaning that can be translated into operational response. The Vlaketunnel example shows that far outside the empirical support of the estimation sample, in a corridor where no same-hierarchy road alternatives exist, the linear TI effect underestimates the observed response. This is consistent with the earlier observation that severe detours are rare in the Dutch highway network: the coefficients are estimated predominantly on closures where alternative routes distribute the disruption, and a single-bottleneck corridor such as Zeeland is structurally unrepresented in the estimation sample. The underprediction should therefore be read as heterogeneity of the treatment effect across corridor structures rather than as a failure of the causal estimates.

Overall, the findings of this thesis say something more general than that road closures increase rail ridership. Planned closures function as recurring natural experiments in the disruption of travel habits. They cause, at known times and places, exactly the kind of context change that the behavioural literature identifies as a precondition for reconsidering mode choice following normal habits (Verplanken et al., 2008). The results delineate how much disruption that reconsideration requires. A closure must extend beyond a full working-week cycle (8+ days) before any measurable response appears, and even then the response scales with the degradation of the car alternative. Mode choice is, in other words, robust to substantial temporary friction evident from the non-significant response closures of a short duration. This finding puts the optimism of studies documenting behaviour change after forced disruptions into perspective. Larcom et al. (2017), for example, found lasting change after a metro strike, but that change concerned rerouting within the same network rather than a switch to another mode. The results here confirm the underlying mechanism, as travellers do try the alternative, but show that a change of mode requires a far longer and more severe disruption.

Who responds, and when, sharpens this picture into an imbalance for operations. The additional ridership is generated predominantly by travellers without an established rail habit, the incidental full-fare passengers, and it arrives predominantly in the peaks, when the rail system has the least spare capacity. The closure effect is therefore inconvenient in operational terms. It recruits the passengers that are hardest to predict from existing travel patterns, at the moments in which they are the hardest to accommodate. At the same time, this is precisely what makes the effect policy-relevant rather than merely academic. The travellers in the window are new to rail, and whether their first experience occurs on an overcrowded peak train or a lengthened one may influence whether any of the behaviour persists beyond the closure. This study cannot answer that question, but its results make it an interesting future study.

The groups most responsive to the closure are also those with the least to lose from trying the train namely occasional travellers and those whose journey is already subsidised. Marsden and Docherty (2013) argue that long closures represent rare windows in which travel behaviour is genuinely open to change. This passenger heterogeneity result identifies exactly which travellers are in that window which suggests that route-specific promotions targeting occasional travellers would be the most efficient use of resources during 8+ day closures. Coordination with employers and educational institutions could also play a large role.

## 5.2 PRACTICAL RECOMMENDATIONS

The findings of this study translate into recommendations in three areas, namely capacity planning around closures, commercial targeting during closures, and the organisational embedding of the prediction capability.

For capacity planning, the main recommendation is to treat the maintenance calendar as a demand forecast input. The prediction pipeline developed in this study can be run against the published closure planning each planning cycle, and the validation results define how its output should be used. Within the empirical support of the estimation sample, the forecasts are accurate enough to inform capacity decisions for individual closures. For short and low-intensity closures, the correct operational conclusion is that no measurable ridership response should be expected and no additional capacity is warranted, which is itself valuable as it prevents costly reinforcement of services where no demand increase will occur. For closures in single-bottleneck corridors, the forecast serves as a flagged lower bound, signalling that the corridor deserves extra attention and a capacity buffer precisely because the model cannot be trusted to capture the full response there. The passenger-km concentration results underline this further, as a small number of long closures on the busiest corridors accounts for the vast majority of the total effect. Capacity preparation should therefore concentrate on the handful of 8+ day closures with substantial detour severity on high-volume corridors, rather than being spread across all planned works.

The timing results add an important qualification to any reinforcement decision. The additional demand does not arrive in the off-peak, where spare seating is typically available, but in the morning peak and evening peak when load factors are already at their highest. Train reinforcement should therefore be an increase in peak capacity, for example through train lengthening on the affected corridors, rather than a uniform service increase across the day.

NS is often at capacity for rolling stock especially during peak hours. Therefore it might be necessary to dampen the ridership increase by targetting specific groups identified by this study to be more likely to take the train during road closures. Students and business-card travellers can be influenced directly by their educational institutions and employers. The recommendation to NS is therefore to talk to these authorities to spread the demand during the period of the closure or at least on the absolute peak moments.

For the road and rail authorities jointly, the practical recommendation is that the national maintenance calendar is not only an engineering schedule but a behavioural instrument. The small number of long, high-detour closures that dominate the aggregate effect are known months in advance. The prediction pipeline developed here identifies them, and the heterogeneity results identify whom to target and when to reinforce.

## 5.3 GENERALISABILITY

The method and the findings of this study generalise differently and it is worth separating both. The method is not tied to the Dutch context. The conceptual framework, the closure detection approach, the treatment intensity measure, the fixed-effects design, and the prediction pipeline are all defined independently of any specific dataset. The data used is described in section 3.3.2 but follows a set of requirements for the data which can be followed to select what dataset to use should the method be applied to a different case. Together with the parameter overview in in Table 2, this forms a complete recipe for applying the analysis in another country or region. The Dutch implementation presented here is one case of that recipe.

The estimates themselves transfer with more caution. The effect sizes are conditional on the Dutch setting including a dense highway network with relatively frequent full-lane closures for maintenance. The rail-based alternatives are also not always transferable to other settings where the train might offer a better or worse alternative to the car. Each of these shifts the response. A sparser highway network would produce higher treatment intensities, a weaker rail alternative would mute the behavioural response to the same intensity, and a different road works regime could change the way road works are communicated and therefore perceived by the public. The Vlaketunnel validation case makes the same point within the Netherlands itself, as the effect estimated on corridors with distributed road alternatives already underestimates the response in a single-bottleneck corridor. If transferability breaks between Dutch corridor types, it should not be assumed across countries. The direction of the effect is well grounded in theory and is expected to hold generally, but the magnitudes reported here should be treated as specific to the network and service context in which they were estimated.

## 5.4 LIMITATIONS AND SIMPLIFICATIONS

The most fundamental limitation of this study is that a causal increase in rail ridership is identified, but a modal shift from car to rail cannot be directly confirmed. The AFC data records only rail trips, meaning nothing can be observed about the travel behaviour of the additional passengers before the road closure. The ridership increase is consistent with a car-to-rail modal shift, but this interpretation rests on inference rather than direct observation. Using the, already extracted, SHiVi induction loop data during the same closure events, but on the likely detour routes, would be needed to confirm that the rail increase coincides with a measurable reduction in car trips. Even then, it can not be completely ruled out that travellers simply do not travel during road works. Similarly, this study did not observe whether any behavioural change is sustained after the closure ends. The literature on habitual travel suggests that prolonged disruptions can produce lasting modal change (Gärling and Axhausen, 2003; Marsden and Docherty, 2013), but this research only looks at ridership during closures, not after them. Finally, trip purpose is not directly observable from AFC data and passenger card category is used as a proxy. The personal commute category in particular is a catch-all that likely includes social, recreational, and incidental travellers, which introduces noise into the passenger heterogeneity analysis and makes the null results for that group harder to interpret with confidence.

Methodologically a limitation holds in terms of the model applied. The magnitude of the 8+ day effect carries specification uncertainty: depending on how common time shocks are modelled, the total effect at $TI = 1.0$ ranges from 0.23 to 0.38 (Section 4.4.5). The passenger-kilometre aggregation and the prediction pipeline are computed under the main specification and would scale proportionally.

The treatment intensity variable is an approximation of the generalised cost increase imposed by a closure. The routing model used to compute detour ratios does not account for congestion on detour routes. In reality, a road closure diverts traffic onto alternative roads which increases congestion and therefore travel times beyond the free-flow travel time which TI reflects. This means that TI systematically underestimates the true car travel time increase, and the estimated ridership responses are therefore a lower bound on the full effect of the generalised cost increase. Working in the opposite direction, the routing model excludes local and regional roads, forcing short distance OD-pairs onto highway detours that are longer than the realistic alternative. For a closure such as the A12 between Driebergen and Utrecht, the model assigns a lengthy highway detour where a traveller would in practice use local roads, increase the TI for those OD-pairs beyond reasonable levels. These two biases partially offset each other, but their magnitude varies by OD-pair, introducing measurement errors that get larger at the extremes of the TI distribution. A third measurement limitation is that TI captures only the expected travel time increase and not the unreliability that comes with the detour. Road closures introduce variability into car journey times which penalise the generalised cost independently of the actual travel time (Small et al., 2005), meaning the true generalised cost increase is larger than the TI alone. If you take all of these measurement limitations together, the effect is likely conservative relative to the true response.

The travel time ratio used to measure rail competitiveness rests on an unfair comparison. Free-flow car travel time is divided by level-of-service train travel time. Free-flow car time reflects the most optimistic driving condition, while the level-of-service includes waiting time, transfer penalties, and frequency. This inflates the train travel time relative to the car. Because this comparison is applied consistently across every OD-pair, the asymmetry does not threaten the validity of the heterogeneity results. What the asymmetry does mean is that the TTR values themselves cannot be interpreted as a realistic representation of actual travel time differences between modes. Additionally, the TTR is computed using undisrupted car travel conditions and does not change during a closure, even though car travel times increase and reliability falls when a closure is in effect (Jamous and Balijepalli, 2018). The actual competitive position of rail during a closure is therefore more favourable than the static TTR implies.

The scope of this study is restricted to full highway closures, leaving partial road works outside the detection algorithm. Partial closures are common in Dutch highway maintenance practice and likely also affect car travel times. Although the response is fundamentally different as partial closures still leave the preferred route open for those who really want to take it as opposed to full closures which completely prevent the usage of that roadway during the closure. While the Coentunnel exclusion robustness check confirms that no single event drives the result, the limited geographic spread reduces confidence that the finding is evenly applicable to all locations around the Netherlands. The two-year study period further constrains the number of long duration events available for estimation. The analysis is also restricted to interurban OD-pairs and only looks at AFC-data from NS trains meaning that the response of local and regional public transportation is not captured. The total public transport response is therefore likely larger than what this study measures. Finally, the results are based on the Dutch context which includes a dense highway network paired with a competitive intercity rail service. This may not generalise well across other networks where rail is a weaker alternative or where highway density provides fewer detour options.

The predictive application inherits every limitation of the causal estimates and adds three of its own. First, the linear-in-TI functional form is only validated within the empirical support of the estimation sample. The Vlaketunnel case demonstrates that outside this support the response can be substantially non-linear in ways the predictive model cannot anticipate. Second, the validation rests on ten closure events on three corridors, of which only one, the A28, tests the model in which the pipeline is claimed to be quantitatively accurate. A single accurate case is positive but not conclusive, and continued validation as new long-duration closures occur is necessary but was not possible due to the timing of this research and the low number of long closures which had taken place in the first half of 2026. Third, the observed uplifts are control-adjusted, and the adjustment assumes that network-wide influences affect treated and control pairs equally. During the Vlaketunnel closure the control pairs themselves rose by seven to eight percent, which may reflect a simultaneous NS promotional campaign.

Table 21 provides an overview of the simplifications and assumptions made in both the exposure modelling and causal analysis. The justifications and the directional effect on the estimates are also included.

Table 21: Simplifications and assumptions made in the exposure modelling and causal analysis, with justification and directional effect on estimates

| Simplification | Justification | Effect |
|---|---|---|
| *Exposure modelling* | | |
| Free flow on detour routes | Modelling congestion on detour routes requires dynamic traffic assignment, this is computationally hard and infeasible for this scope. A free-flow assumption gives a reproducible baseline assumption. | TI understates true cost increase |
| Deterministic route choice and limited alternatives | Stochastic route choice requires individual-level data unavailable in this scope. A deterministic fastest-path approach is transparent and reproducible. | Exposure may be over- or underestimated for individual OD-pairs |
| Single centroid per municipality | To route every possible OD-pair when using detailed origins and destinations, e.g. postal codes, would require massive computational power. A single centroid ensures consistent and comparable travel time estimation across all OD-pairs. | Travel times and route assignments may deviate from those experienced by travellers, especially in large municipalities |
| Highway network only, local roads excluded | The study targets interurban car-to-rail competition, for this the highway network is most relevant. Local roads are not covered by the induction loop network. | TI overestimated for short-distance OD-pairs where the realistic detour uses local roads |
| Closure identification via intensity drops | No comprehensive historical dataset of confirmed closures with severity attributes exists. Traffic intensity provides a practical and observable proxy. | Treatment set is conservative, some closures may be missed or misclassified |
| Daily aggregation of traffic conditions | Sub-daily data introduces noise events other than full road closures. Daily aggregation improves comparability and reduces incidental disturbances. | Peak-period impacts may be understated; within-day variation is not captured |

*Continued on next page*

| Simplification | Justification | Effect |
|---|---|---|
| No within-mode behavioural adaptation | The exposure model isolates relative travel condition changes between car and rail rather than simulating all behavioural responses. | Traffic remaining on the road network may be overstated in the exposure calculation |
| *Causal analysis* | | |
| Linear TI response within bins | A linear specification is the natural first-order approximation consistent with RUM theory. Non-linear specifications require stronger parametric assumptions and a larger per-bin sample. | Diminishing returns at high TI values would not be detected |
| Additive fixed effects | Fully interacted entity $\times$ time fixed effects are computationally infeasible at this panel size and would consume most identifying variation. | Time-varying unobservables that differ by entity are not controlled |
| Parallel trends as maintained assumption | The identifying assumption cannot be directly tested but is supported by the pre-trends event study across three detour ratio thresholds. | If violated, estimated effects may be biased in either direction |
| No spatial spillovers between corridors | Modelling spatial interdependence between OD-pairs requires a spatial econometric framework that substantially increases complexity. | If traffic diverts through adjacent corridors, some control OD-pairs may be indirectly treated, reducing the strength of estimates |

## 5.5 FUTURE RESEARCH

The most pressing future research is to confirm whether the identified ridership increases reflect an actual modal shift from car to rail. Two approaches could address this. First, a stated preference survey administered to travellers during a major 8+ day closure would allow direct surveying about prior mode use and the reasons for switching, if switching at all. This would provide more context that a revealed preference study alone cannot give. While SP surveys do potentially have a bias as discussed earlier, the combination of SP with RP would allow for cross-validation of both sources. Second, traffic intensity data from vehicle induction loops, also used during this study, could be used to test whether measured rail ridership increases coincide with measurable increase in road congestion on alternative routes. Together these two approaches would allow the ridership increase identified in this study to be verified against both stated intentions and observed road behaviour.

This analysis is restricted to full highway closures of at least two consecutive days, leaving two categories of disruptions outside of the scope. Firstly, partial road closures are not identified which can still reduce capacity without fully closing the road. These types of capacity reductions are common in Dutch highway maintenance and still produce longer travel times due to congestion. This has an effect on the reliability and travel time by car and therefore the generalised cost of travel by car. Extending the detection algorithm to also flag partial road closures, for example by setting higher thresholds or looking at lane-specific induction loops, would increase the number of treatable events and allow the severity response to be studied across a wider range. Closures of only a single day are excluded due to the minimum duration filter. The expectation is that a one-day closure falls well below the behavioural threshold identified in this study, confirming this would strengthen the duration response results and rule out the possibility that very severe single-day closures produce ridership effects.

Some improvements to the independent variables used in this study would increase the reliability of both the causal estimates and the heterogeneity analyses. The treatment intensity variable can be expanded to include real-time or historical speed data on detour routes to model congestion more precisely. Weather is another omitted moderator as rain may increase the ridership response for closures that coincide with poor weather. Making the TTR measure more dynamic could be done by recomputing rail competitiveness under disrupted car conditions rather than free-flow optimal conditions. Adding communication as another variable would allow the estimation of this mediator.

Understanding whether the ridership increases observed during closures translate into lasting behavioural change would answer a policy-relevant question that this study did not touch. Access to individual anonymous card-level AFC data would allow a before, during, and after comparison at the individual passenger level, directly testing whether 8+ day closures can change habits. This would require a more detailed dataset than the aggregated data used here but it is technically feasible.

The validation results also point to a specific gap in the estimation sample: closures in single-bottleneck corridors with treatment intensities far above the sample's 95th percentile. Similar closures with high detour ratios might occur more often in the future and could have a significant effect on ridership. Therefore a study looking at these edge-cases can increase the boundaries in which this model can be applied.

The scope of the study could be broadened which would increase its generalisability. The analysis covers only interurban rail operated by NS, but a road closure also affects demand for regional trains, buses, and metro services operated by other operators. To analyse the full public transport response to highway closures would require access to ridership data from these operators. In the Netherlands however, data sharing between transport operators is constrained by confidentiality which limits the feasibility of such analysis. This is an unfortunate restriction given the density of the Dutch network and the way that the railways complement each other. It points to a broader need for open data across Dutch public transport operators. Beyond the Netherlands, replicating the analysis in countries with different rail network densities, highway maintenance operations, and traveller profiles would test the generalisability of the findings. Countries with weaker rail alternatives would expect to produce smaller or no effects, while those with denser and more competitive rail network may show stronger responses.

# 6 CONCLUSION

The problem that road closures pose to rail operators is the uncertainty of how many extra passengers to expect during that closure and how different closure characteristics affect that response. That uncertainty can lead to inadequate operational responses as a result of which trains can become overcrowded if the ridership is underestimated or too many trains and personnel can be planned which leads to higher expenses and wear on the rolling stock. In order to more accurately predict the ridership during these road closures, a causal effect needs to be established between road closures and rail ridership. This led to the research question: *What is the causal effect of planned road closures on rail ridership in the Netherlands?*

To identify causality, first it had to be determined which road closures actually took place. No open archived dataset is available to be used for this analysis and therefore one had to be constructed. By using a dataset that is open to the public, the SHiVi road intensity data, full-day road closures were inferred from intensity drops in the observed vehicle counting data. 215 road closure events were identified which, when combined, account for 882 closure days in 2024 and 2025. 1,533 of the 2,070 (74.1%) OD-pairs were exposed to at least one road closure. This exposure is widespread with a concentration of road closures around Amsterdam and only the sparsely populated North left largely untouched. This constructed dataset was cross-validated using a private dataset from RWS and forms a useful dataset for future research that requires a comprehensive list of full road closures in the Dutch highway network. The methodology can also be extended to cover future road closures once that data becomes available.

With a complete road closure dataset constructed, the causal analysis could be performed. The duration of the closure is the most important moderator of ridership effect which was found by splitting the duration of the closure up into three bins of 2–3 days, 4–7 days, and 8+ days with the last bin of 8+ days showing a positive and significant ridership effect during road closures. Shorter closures between 2–7 days do not lead to any ridership effect, neither positive nor negative, which comes from the ability to adapt travel patterns when the effects are short in nature. The longer road closures of 8+ days spur an increase in ridership of $\beta = 0.375$ at $\mathrm{TI} = 1.0$ which means the travel time by car has increased by 100% relative to the undisrupted condition. A more realistic value can be found at the ridership-weighted median of $\mathrm{TI} = 0.1$ which results in a $\beta$ of 0.0375 and thus a ridership increase of 3.7%.

The duration of the road closure is therefore the governing characteristic steering the resulting ridership. The detour ratio of 8+ day closures also has a significant positive effect on the ridership which indicates that road closures in denser parts of the country, where relative distances between OD-pairs are short and therefore the detours are relatively large, will lead to higher ridership responses. Road closures in the peripheries of the country will therefore likely induce less of a ridership response compared to those taking place in the heart of the population. The relative attractiveness of the rail-based alternative, measured through travel time ratio, matters only weakly and non-significantly although the positive trend is in line with expectations and literature.

Different types of passengers respond heterogeneously with incidental, full-fare passengers responding most strongly at $\beta = 0.77$. This indicates that many of the extra passengers travelling by train during the road closures are those who do not travel often and are likely travelling incidentally due to the closure. Passengers who are not paying for their journey themselves, students ($\beta = 0.484$) and business travellers with employer-paid cards ($\beta = 0.390$), also see a significant increase in ridership during the closure, larger than the aggregate. This can be explained by the low cost of switching for these passengers which can tip them over the point of indifference during the road closure. Social and recreational travellers show less of a response except during shorter, weekend closures. Commuters paying for their own subscription and other passengers show no significant response.

The causal estimates were subsequently turned into the prediction capability that motivated this research. A pipeline was constructed that translates a planned closure specification into a predicted ridership uplift per OD-pair, and it was validated under prospective conditions against ten closure events from 2026 that the model had never seen. The validation delineates three regimes. Under conditions comparable to the estimation sample, the pipeline is quantitatively accurate. The 44-day closure of the A28 was predicted at +8.3% against an observed, control-adjusted uplift of +6.7%, within the confidence interval. For short closures the predicted effects are smaller than the day-to-day variation in ridership, and the correct operational expectation is no measurable response. For closures far outside the empirical support of the estimation sample, such as the ten-day Vlaketunnel closure where the observed response exceeded the prediction by a factor of two to three, the forecasts serve as lower bounds.

To conclude, planned highway closures cause a measurable increase in rail ridership but only under certain circumstances. The closure needs to be longer than a week for any aggregate effect to take place and the detour caused by the closure plays a meaningful secondary role in the magnitude of the response. Incidental, discretionary passengers and passengers not personally paying for their journey respond most strongly, which can be used to tailor operational response as effectively as possible. These estimates address the uncertainty with which this chapter opened: NS can now anticipate, before a planned closure begins, whether a ridership response should be expected at all, how large it is likely to be, and on which OD-pairs it will concentrate. This is done by using the validated prediction pipeline indicating, for every forecast, whether it falls within the domain in which its accuracy has been demonstrated. Beyond the Dutch findings, the thesis leaves behind a documented and validated procedure, from closure detection through prediction, that future research can apply wherever traffic intensity, road network, and fare collection data are available.

# BIBLIOGRAPHY


Bates, John (2003). 'Departure Time Shifts, Waiting and Headway.' In: *Public Transport Quality Literature Review*. Appendix B. Prepared for the UK Department for Transport. Faber Maunsell.

Ben-Akiva, Moshe and Steven R. Lerman (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (Feb. 2004). 'How much should we trust Differences-in-Differences estimates?' In: *The Quarterly Journal of Economics* Vol. 119, No. 1, pp. 249–275.

Callaway, Brantly and Pedro H.C. Sant'Anna (Dec. 2021). 'Difference-in-Differences with multiple time periods.' In: *Journal of Econometrics* 225.2, pp. 200–230. ISSN: 0304-4076. DOI: 10.1016/J.JECONOM.2020.12.001.

Cameron, A Colin and Douglas L Miller (2015). 'A Practitioner's Guide to Cluster-Robust Inference.' In.

Cats, O. (Feb. 2016). 'The robustness value of public transport development plans.' In: *Journal of Transport Geography* 51, pp. 236–246. ISSN: 0966-6923. DOI: 10.1016/J.JTRANGEO.2016.01.011.

Cats, Oded and Erik Jenelius (Apr. 2018). 'Beyond a complete failure: the impact of partial capacity degradation on public transport network vulnerability.' In: *Transportmetrica B: Transport Dynamics* 6.2, pp. 77–96. ISSN: 21680582. DOI: 10.1080/21680566.2016.1267596. URL: https://www.tandfonline.com/doi/abs/10.1080/21680566.2016.1267596.

Cats, Oded, Gert Jaap Koppenol, and Martijn Warnier (Nov. 2017). 'Robustness assessment of link capacity reduction for complex networks: Application for public transport systems.' In: *Reliability Engineering & System Safety* 167, pp. 544–553. ISSN: 0951-8320. DOI: 10.1016/J.RESS.2017.07.009.

Chorus, Caspar G., Eric J.E. Molin, and Bert van Wee (Sept. 2006). 'Travel information as an instrument to change cardrivers' travel choices: a literature review.' In: *European Journal of Transport and Infrastructure Research* 6.4, pp. 335–364. ISSN: 1567-7141. DOI: 10.18757/ejtir.2006.6.4.3456. URL: https://journals.open.tudelft.nl/ejtir/article/view/3456.

De Keizer, Bart, Marco Kouwenhoven, and Freek Hofker (2015). 'New insights in resistance to interchange.' In: *Transportation Research Procedia*. Vol. 8. Elsevier, pp. 72–79. DOI: 10.1016/j.trpro.2015.06.043.

Eltved, Morten, Nils Breyer, Jesper Bláfoss Ingvardson, and Otto Anker Nielsen (2021). 'Impacts of long-term service disruptions on passenger travel behaviour: A smart card analysis from the Greater Copenhagen area.' In: DOI: 10.1016/j.trc.2021.103198. URL: https://doi.org/10.1016/j.trc.2021.103198.

Gärling, Tommy and Kay W. Axhausen (2003). 'Introduction: Habitual travel choice.' In: *Transportation* 30.1, pp. 1–11. ISSN: 00494488. DOI: 10.1023/A:1021230223001/METRICS. URL: https://link.springer.com/article/10.1023/A:1021230223001.

Hillcoat, Chris (Mar. 2018). 'Reducing congestion during road works through travel demand management.' In: *Proceedings of the Institution of Civil Engineers - Municipal Engineer* 171.1, pp. 51–60. ISSN: 0965-0903. DOI: 10.1680/JMUEN.16.00064. URL: https://dx.doi.org/10.1680/jmuen.16.00064.

Jacobs, Jane (1961). *The Death and Life of Great American Cities*. New York: Random House.

Jamous, Wael and Chandra Balijepalli (Aug. 2018). 'Assessing travel time reliability implications due to roadworks on private vehicles and public transport services in urban road networks.' In: *Journal of Traffic and Transportation Engineering (English Edition)* 5.4, pp. 296–308. ISSN: 2095-7564. DOI: 10.1016/J.JTTE.2017.09.003.

Kattan, Lina, Alexandre G. De Barros, and Hina Saleemi (Nov. 2013). 'Travel behavior changes and responses to advanced traveler information in prolonged and large-scale network disruptions: A case study of west LRT line construction in the city of Calgary.' In: *Transportation Research Part F: Traffic Psychology and Behaviour* 21, pp. 90–102. ISSN: 1369-8478. DOI: 10.1016/J.TRF.2013.08.005.

Kennisinstituut voor Mobiliteitsbeleid (KiM) (2026). *De student als reiziger: over het reisgedrag, bijbehorende uitdagingen en mogelijke handelingsperspectieven ter verbetering*. Tech. rep. Den Haag: Ministerie van Infrastructuur en Waterstaat. DOI: 10.82230/KiM.MB2508. URL: https://doi.org/10.82230/KiM.MB2508.

Krijgsman van Spangenberg, Thijs Marinus Willem, Niels van Oort, Alexandra Gavriilidou, Freek Verhoof, and Panagiotis Spartalis (2026). 'Long-term impact of rail maintenance works on passenger patterns: A longitudinal analysis of individual passenger patterns in the Dutch main rail network.' Master's thesis, Delft University of Technology. Unpublished; presented 15 July 2026.

Larcom, Shaun, Ferdinand Rauch, and Tim Willems (Nov. 2017). 'The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network.' In: *The Quarterly Journal of Economics* 132.4, pp. 2019–2055. ISSN: 0033-5533. DOI: 10.1093/QJE/QJX020. URL: https://dx.doi.org/10.1093/qje/qjx020.

Lo, Shih Che and Randolph W. Hall (Dec. 2006). 'Effects of the Los Angeles transit strike on highway congestion.' In: *Transportation Research Part A: Policy and Practice* 40.10, pp. 903–917. ISSN: 0965-8564. DOI: 10.1016/J.TRA.2006.03.001.

Marsden, Greg and Iain Docherty (May 2013). 'Insights on disruptions as opportunities for transport policy change.' In: *Transportation Research Part A: Policy and Practice* 51, pp. 46–55. ISSN: 0965-8564. DOI: 10.1016/J.TRA.2013.03.004.

McFadden, Daniel (1974). 'Conditional Logit Analysis of Qualitative Choice Behavior.' In: *Frontiers in Econometrics*. Ed. by Paul Zarembka. New York: Academic Press, pp. 105–142.

Nationaal Dataportaal Wegverkeer (2024). *Handleiding Nationaal WegenBestand (NWB)*. Version 1.5. Version 1.5, last updated 12 May 2025. Accessed: April 2026. URL: https://docs.ndw.nu/handleidingen/nwb/.

Nationaal Dataportaal Wegverkeer and Rijkswaterstaat (2019). *SHiVi: Samenwerking Historische Verkeersinformatie*. Dataset available from 3 January 2019. Accessed: April 2026. URL: https://docs.ndw.nu/handleidingen/DEXTER/shivi/.

Rijkswaterstaat (2024). *Staat van de Infrastructuur*. Rapportage. Utrecht: Rijkswaterstaat. URL: https://www.rijksoverheid.nl/documenten/2025/12/08/bijlage-2-staat-van-de-infrastructuur-rijkswaterstaat-2024.

Roth, Jonathan (2022). 'Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends †.' In: *AER: Insights* 2022.3, pp. 305–322. DOI: 10.1257/aeri.20210236. URL: https://doi.org/10.1257/aeri.20210236.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe (Aug. 2023). 'What's trending in difference-in-differences? A synthesis of the recent econometrics literature.' In: *Journal of Econometrics* 235.2, pp. 2218–2244. ISSN: 0304-4076. DOI: 10.1016/J.JECONOM.2023.03.008.

Small, Kenneth A. and Erik T. Verhoef (2007). *The Economics of Urban Transportation*. London: Routledge.

Small, Kenneth A., Clifford Winston, and Jia Yan (July 2005). 'Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability.' In: *Econometrica* 73.4, pp. 1367–1382. ISSN: 1468-0262. DOI: 10.1111/J.1468-0262.2005.00619.X. URL: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1468-0262.2005.00619.xhttps://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0262.2005.00619.xhttps://onlinelibrary.wiley.com/doi/10.1111/j.1468-0262.2005.00619.x.

Spyropoulou, Ioanna (Feb. 2020). 'Impact of public transport strikes on the road network: The case of Athens.' In: *Transportation Research Part A: Policy and Practice* 132, pp. 651–665. ISSN: 0965-8564. DOI: 10.1016/J.TRA.2019.12.022.

Sullivan, J. L., D. C. Novak, L. Aultman-Hall, and D. M. Scott (June 2010). 'Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach.' In: *Transportation Research Part A: Policy and Practice* 44.5, pp. 323–336. ISSN: 0965-8564. DOI: 10.1016/J.TRA.2010.02.003.

Tennøy, Aud and Oddrun Helen Hagen (July 2021). 'Urban main road capacity reduction: Adaptations, effects and consequences.' In: *Transportation Research Part D: Transport and Environment* 96, p. 102848. ISSN: 1361-9209. DOI: 10.1016/J.TRD.2021.102848.

Van Exel, N. Job A. and Piet Rietveld (Oct. 2001). 'Public transport strikes and traveller behaviour.' In: *Transport Policy* 8.4, pp. 237–246. ISSN: 0967-070X. DOI: 10.1016/S0967-070X(01)00022-1.

Verplanken, Bas, Ian Walker, Adrian Davis, and Michaela Jurasek (June 2008). 'Context change and travel mode choice: Combining the habit discontinuity and self-activation hypotheses.' In: *Journal of Environmental Psychology* 28.2, pp. 121–127. ISSN: 0272-4944. DOI: 10.1016/J.JENVP.2007.10.005.

Wardman, Mark (Oct. 2004). 'Public transport values of time.' In: *Transport Policy* 11.4, pp. 363–377. ISSN: 0967-070X. DOI: 10.1016/J.TRANPOL.2004.05.001.

Yap, M. D., S. Nijënstein, and N. van Oort (Jan. 2018). 'Improving predictions of public transport usage during disturbances based on smart card data.' In: *Transport Policy* 61, pp. 84–95. ISSN: 0967-070X. DOI: 10.1016/J.TRANPOL.2017.10.010.

Yap, Menno and Oded Cats (Jan. 2022). 'Analysis and prediction of ridership impacts during planned public transport disruptions.' In: *Journal of Public Transportation* 24, p. 100036. ISSN: 1077-291X. DOI: 10.1016/J.JPUBTR.2022.100036.

Zhu, Shanjiang and David M. Levinson (2012). 'Disruptions to Transportation Networks: A Review.' In: pp. 5–20. DOI: 10.1007/978-1-4614-0947-2{_}2. URL: https://link.springer.com/chapter/10.1007/978-1-4614-0947-2_2.

Zhu, Shanjiang and David Levinson (2008). 'Planned and Unplanned Disruptions to Transportation Networks.' In: URL: www.research.dot.state.mn.us.

# A

# LITERATURE MATRIX

| Paper | Authors | Phenomenon | Trigger | Modes | Data | Purpose |
|---|---|---|---|---|---|---|
| | | What is being studied? | What causes the behavioural change? | Car, rail, bus, active… | RP, SP, Smartcard, Detection loops, GPS | Explanation, Prediction, Evaluation… |
| (2021), Impacts of long-term service disruptions on passenger travel behaviour: A smart card analysis from the Greater Copenhagen area | Eltved, M; Breyer, N; Bláfoss J; Nielsen, O | Behaviour of passengers before and after a long-term disruption on a railway line | A long-term closure (3 months) of a suburban railway line | Rail | Smartcard, RP | Prediction |
| (2008) Planned and Unplanned Disruptions to Transportation Networks | Zhu S; Levinson D | Behaviour during major unplanned disruptions(route choice, departure time, mode choice, trip suppression, long-term habit formation) | A bridge collapse or other major unplanned disruptions | Road | Aggregate traffic counts, short-term surveys (SP) | Review empirical findings |
| (2013) Insights on disruptions as opportunities for transport policy change | Marsden G; Docherty I | Discussion how travel disruptions provide opportunities for behaviour change | Various disruptions, planned and unplanned | Road, Rail, PT | Existing studies | policy opportunity could unlock more radical transport behaviour change than conventional incremental policy; not an empirical study of behavioural |
| (2018) Reducing congestion during road works through travel demand management | Hillcoat C | How communication to drivers about planned road works affects their travel behaviour compared to non-communicated road works | Transport disruptions in general (planned and unplanned), treated as management and communication challenges | Road | ANPR | To argue that how disruption is communicated and managed strongly shapes user response and system performance |
| (2001) Public transport strikes and traveller behaviour | Van Exel N; Rietveld P | How passengers respond to planned and unplanned public transportation strikes in the short and the long term | Public transportation strikes | PT, Road | Literature | Literature review |
| (2014) Estimation of a route choice model for urban public transport using smart card data | Jánošíkova, L; Slavík J, Koháni M | Individual public transport route choice behaviour | No disruption. Normal day-to-day operations | PT | Large-scale revealed preference smart card data, linked to network attributes | To demonstrate how individual PT choices can be inferred and modelled using smart card data rather than surveys |
| (2020) Crowding valuation in urban tram and bus transportation based on smart card data | Yap M; Cats O; van Arem B | How service quality attributes (specifically crowding) affect public transport choice behaviour | No disruption. Variation comes from operational conditions (crowding levels) | PT | Large-scale revealed preference smart card data, combined with vehicle occupancy and service attributes | To quantify disutility of crowding and demonstrate how smart card data can support behavioural valuation of PT attributes |
| (2013) Travel behavior changes and responses to advanced traveler information in prolonged and large-scale network disruptions: A case study of west LRT line construction in the city | Kattan L; De Barros A; Saleemi H | Travel behaviour during constructions works of a new LRT line in Calgary. Mainly looking at private vehicle travellers changing to other modes or other routes | Construction of new public transportation route which impedes road traffic both private and public | PT, Road | Stated Preference survey | Evaluation |
| (2022) Analysis and prediction of ridership impacts during planned public transport disruptions | Yap M; Cats O | Demand response due to planned PT disruptions | Planned PT disruptions | PT | RP based on smart card data | Evaluation and prediction using model |
| (2016) The robustness value of public transport development plans | O. Cats | Network robustness | Planned PT disruptions | PT | Modeling | Modeling |
| (2006) Effects of the Los Angeles transit strike on highway congestion | Shih-Che Lo; Randolph W. Hall | Traffic behaviour due to PT strike | Semi-planned PT disruption (strike) | PT, Road | RP vehicle counters | Evaluation and prediction |
| (2020) Impact of public transport strikes on the road network: The case of Athens | Ioanna Spyropoulou | Traffic behaviour due to PT strike | Semi-planned PT disruption (strike) | PT, Road | RP vehicle counters | Evaluation and prediction |
| (2021) Urban main road capacity reduction: Adaptations, effects and consequences | Aud Tennøy; Oddrun Helen Hagen | Travel behaviour due to road capacity reduction. Modal shift included | Planned road capacity reduction | PT, Road | AFC, SP | Evaluation and prediction |

| | | | | | | |
|---|---|---|---|---|---|---|
| (2017) Robustness assessment of link capacity reduction for complex networks: Application for public transport systems | Oded Cats; Gert-Jaap Koppenol; Martijn Warnier | Network robustness under planned capacity reduction | Planned capacity reduction PT | PT | AFC | Evaluation and prediction |
| (2018) Beyond a complete failure: the impact of partial capacity degradation on public transport network vulnerability | Oded Cats; Erik Jenelius | Societal costs due to partial capacity reduction in PT network | Partial capacity reduction PT | PT | | |
| (2010) Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach | J.L. Sullivan; D.C. Novak; L. Aultman-Hall; D.M. Scott | Capacity reduction | Road capacity reduction | Road | Modeling | Modeling |
| (2006) Travel information as an instrument to change car drivers' travel choices: A literature review | Chorus, Caspar G.; Molin, Eric J. E.; van Wee, Bert | Travel information presented during capacity reduction | Road capacity reduction | Road | Literature | Literature review |

# B ROAD CLOSURES

Table 22: Full list of identified road closures ($n = 2$ days or longer) used in the exposure and causal analysis, grouped by road.

| Road | Location (km) | Start | End | Days | Baseline intensity |
|---|---|---|---|---|---|
| A1 Re | 89.669 - 94.801 | 08-03-2024 | 09-03-2024 | 2 | 1761 |
| A1 Li | 6.623 - 5.851 | 21-07-2024 | 22-07-2024 | 2 | 1465 |
| A1 Li | 6.623 - 5.851 | 24-07-2024 | 04-08-2024 | 12 | 1535 |
| A1 Li | 6.623 - 5.851 | 08-08-2024 | 11-08-2024 | 4 | 1459 |
| A1 Li | 94.670 - 90.072 | 26-10-2024 | 27-10-2024 | 2 | 1680 |
| A1 Re | 5.602 - 6.630 | 10-05-2025 | 24-05-2025 | 15 | 3284 |
| A1 Li | 22.565 - 18.654 | 31-10-2025 | 01-11-2025 | 2 | 2411 |
| A10 Re | 18.271 - 19.576 | 13-07-2024 | 14-07-2024 | 2 | 3230 |
| A10 Li | 9.595 - 7.375 | 21-07-2024 | 14-08-2024 | 25 | 2330 |
| A10 Re | 7.358 - 9.601 | 15-08-2024 | 08-09-2024 | 25 | 1963 |
| A10 Re | 2.035 - 3.639 | 31-08-2024 | 01-09-2024 | 2 | 1527 |
| A10 n | 30.902 - 28.938 | 26-10-2024 | 27-10-2024 | 2 | 1539 |
| A10 Li | 19.669 - 17.577 | 14-06-2025 | 15-06-2025 | 2 | 4146 |
| A10 Li | 19.669 - 17.577 | 27-06-2025 | 11-07-2025 | 15 | 4239 |
| A10 Li | 14.601 - 14.059 | 27-06-2025 | 11-07-2025 | 15 | 4073 |
| A10 Re | 28.954 - 30.927 | 13-07-2025 | 12-09-2025 | 62 | 2844 |
| A10 Re | 14.053 - 14.594 | 20-09-2025 | 04-10-2025 | 15 | 3685 |
| A10 Li | 19.669 - 17.577 | 18-10-2025 | 19-10-2025 | 2 | 4552 |
| A12 Re | 46.069 - 50.178 | 08-06-2024 | 09-06-2024 | 2 | 1784 |
| A12 Re | 46.069 - 50.178 | 15-06-2024 | 16-06-2024 | 2 | 1784 |
| A12 Re | 46.069 - 50.178 | 22-06-2024 | 23-06-2024 | 2 | 1784 |
| A12 Re | 46.069 - 50.178 | 29-06-2024 | 30-06-2024 | 2 | 1784 |
| A12 Li | 77.120 - 73.103 | 10-05-2025 | 18-05-2025 | 9 | 1962 |
| A12 Re | 109.324 - 119.237 | 14-06-2025 | 15-06-2025 | 2 | 1862 |
| A12 Re | 8.704 - 12.313 | 16-08-2025 | 17-08-2025 | 2 | 1915 |
| A12 Li | 77.120 - 73.103 | 23-08-2025 | 30-08-2025 | 8 | 1811 |
| A12 Re | 19.095 - 22.821 | 23-08-2025 | 24-08-2025 | 2 | 1206 |
| A12 Re | 8.704 - 12.313 | 23-08-2025 | 24-08-2025 | 2 | 1915 |
| A12 Re | 19.095 - 22.821 | 30-08-2025 | 31-08-2025 | 2 | 1206 |
| A12 Re | 8.704 - 12.313 | 30-08-2025 | 31-08-2025 | 2 | 1915 |
| A12 Li | 22.764 - 18.697 | 15-11-2025 | 16-11-2025 | 2 | 1976 |
| A12 Li | 22.764 - 18.697 | 22-11-2025 | 23-11-2025 | 2 | 1976 |
| A12 Li | 22.764 - 18.697 | 13-12-2025 | 14-12-2025 | 2 | 1529 |
| A13 Re | 6.176 - 6.844 | 13-04-2024 | 14-04-2024 | 2 | 2605 |
| A13 Re | 6.176 - 6.844 | 20-04-2024 | 21-04-2024 | 2 | 2605 |
| A13 Re | 6.176 - 6.844 | 12-10-2024 | 13-10-2024 | 2 | 2617 |

*continued on next page*

*Table 22 – continued from previous page*

| Road | Location (km) | Start | End | Days | Baseline intensity |
|---|---|---|---|---|---|
| A13 Re | 11.786 - 16.405 | 30-11-2024 | 01-12-2024 | 2 | 2100 |
| A13 Re | 16.805 - 18.548 | 30-11-2024 | 01-12-2024 | 2 | 1946 |
| A13 Re | 16.805 - 18.548 | 05-04-2025 | 06-04-2025 | 2 | 1921 |
| A13 Re | 16.805 - 18.548 | 04-10-2025 | 05-10-2025 | 2 | 1019 |
| A13 Li | 18.554 - 17.363 | 25-10-2025 | 26-10-2025 | 2 | 2062 |
| A15 Re | 88.285 - 94.490 | 10-08-2024 | 15-08-2024 | 6 | 1491 |
| A15 Li | 94.460 - 88.342 | 17-08-2024 | 21-08-2024 | 5 | 1327 |
| A15 Li | 77.928 - 76.652 | 22-08-2024 | 25-08-2024 | 4 | 1452 |
| A15 Re | 119.006 - 121.134 | 27-09-2024 | 29-09-2024 | 3 | 1318 |
| A15 Re | 64.747 - 65.852 | 28-09-2024 | 29-09-2024 | 2 | 2046 |
| A15 n | 61.996 - 61.311 | 05-10-2024 | 06-10-2024 | 2 | 1164 |
| A15 Re | 75.965 - 77.961 | 19-10-2024 | 20-10-2024 | 2 | 1490 |
| A15 Re | 119.006 - 121.134 | 20-12-2024 | 29-12-2024 | 3 | 1389 |
| A15 Re | 119.006 - 121.134 | 04-04-2025 | 06-04-2025 | 3 | 1254 |
| A15 Re | 119.006 - 121.134 | 11-04-2025 | 13-04-2025 | 3 | 1254 |
| A15 Li | 114.222 - 105.923 | 07-07-2025 | 10-07-2025 | 4 | 1093 |
| A16 Li | 26.936 - 24.680 | 10-05-2024 | 12-05-2024 | 3 | 2415 |
| A16 Li | 17.415 - 16.511 | 10-05-2024 | 12-05-2024 | 3 | 1966 |
| A16 Li | 26.936 - 24.680 | 07-12-2024 | 08-12-2024 | 2 | 2190 |
| A16 Li | 26.936 - 24.680 | 23-08-2025 | 24-08-2025 | 2 | 1863 |
| A16 Li | 26.936 - 24.680 | 25-10-2025 | 26-10-2025 | 2 | 2324 |
| A16 Li | 26.936 - 24.680 | 01-11-2025 | 02-11-2025 | 2 | 2379 |
| A17 Li | 23.488 - 21.663 | 08-07-2025 | 15-07-2025 | 8 | 951 |
| A2 Li | 106.888 - 103.167 | 25-05-2024 | 26-05-2024 | 2 | 2275 |
| A2 Re | 69.056 - 70.592 | 16-08-2024 | 18-08-2024 | 3 | 2558 |
| A2 Re | 69.056 - 70.592 | 23-08-2024 | 25-08-2024 | 3 | 2558 |
| A2 Re | 69.056 - 70.592 | 30-08-2024 | 01-09-2024 | 3 | 2433 |
| A2 Re | 69.056 - 70.592 | 21-09-2024 | 22-09-2024 | 2 | 2246 |
| A2 Li | 247.840 - 246.235 | 28-09-2024 | 29-09-2024 | 2 | 1298 |
| A2 Re | 206.246 - 212.716 | 26-10-2024 | 27-10-2024 | 2 | 917 |
| A2 Li | 138.225 - 132.373 | 02-11-2024 | 03-11-2024 | 2 | 1569 |
| A2 Re | 132.379 - 138.169 | 09-11-2024 | 10-11-2024 | 2 | 1602 |
| A2 Re | 206.246 - 212.716 | 25-10-2025 | 26-10-2025 | 2 | 1058 |
| A2 Re | 229.404 - 232.122 | 15-11-2025 | 16-11-2025 | 2 | 1676 |
| A20 Re | 46.534 - 48.974 | 08-06-2024 | 09-06-2024 | 2 | 1171 |
| A20 Re | 46.534 - 48.974 | 15-06-2024 | 16-06-2024 | 2 | 1171 |
| A20 Re | 46.534 - 48.974 | 22-06-2024 | 23-06-2024 | 2 | 1171 |
| A12 h | 27.244 - A20 Li (HR) 46.704 | 02-11-2024 | 03-11-2024 | 2 | 1382 |
| A20 Re | 27.912 - 28.552 | 09-08-2025 | 10-08-2025 | 2 | 1676 |
| A20 Li | 28.438 - 27.594 | 16-08-2025 | 17-08-2025 | 2 | 1365 |
| N200 Li | 5.616 - 2.576 | 13-12-2024 | 15-12-2024 | 3 | 528 |
| N200 Li | 5.616 - 2.576 | 04-01-2025 | 05-01-2025 | 2 | 432 |

*continued on next page*

*Table 22 – continued from previous page*

| Road | Location (km) | Start | End | Days | Baseline intensity |
|---|---|---|---|---|---|
| A22 Li | 13.010 - 11.103 | 27-04-2024 | 27-04-2024 | 1 | 1000 |
| A22 Re | 11.061 - 12.939 | 27-04-2024 | 27-04-2024 | 1 | 1104 |
| A27 Li | 28.301 - 24.373 | 16-03-2024 | 17-03-2024 | 2 | 1390 |
| A27 Re | 24.567 - 27.964 | 06-07-2024 | 07-07-2024 | 2 | 1251 |
| A27 Li | 65.059 - 58.712 | 13-07-2024 | 14-07-2024 | 2 | 1772 |
| A27 Li | 10.762 - 6.794 | 14-09-2024 | 15-09-2024 | 2 | 1083 |
| A27 Li | 28.301 - 24.373 | 05-10-2024 | 06-10-2024 | 2 | 1431 |
| A27 Re | 58.690 - 65.160 | 05-10-2024 | 06-10-2024 | 2 | 2064 |
| A27 Re | 43.261 - 49.836 | 09-11-2024 | 10-11-2024 | 2 | 1610 |
| A27 Li | 49.882 - 43.229 | 23-11-2024 | 24-11-2024 | 2 | 1614 |
| A27 Re | 43.261 - 49.836 | 23-11-2024 | 24-11-2024 | 2 | 1610 |
| A27 Li | 49.882 - 43.229 | 29-03-2025 | 30-03-2025 | 2 | 1456 |
| A27 Re | 58.690 - 65.160 | 10-05-2025 | 11-05-2025 | 2 | 1638 |
| A27 Re | 43.261 - 49.836 | 24-05-2025 | 25-05-2025 | 2 | 1528 |
| A27 Li | 65.059 - 58.712 | 30-05-2025 | 01-06-2025 | 3 | 1854 |
| A27 Li | 28.301 - 24.373 | 14-06-2025 | 15-06-2025 | 2 | 1420 |
| A27 Re | 24.567 - 27.964 | 28-06-2025 | 29-06-2025 | 2 | 1312 |
| A27 Li | 10.762 - 6.794 | 19-07-2025 | 20-07-2025 | 2 | 1424 |
| A27 Re | 43.261 - 49.836 | 26-07-2025 | 03-08-2025 | 9 | 1739 |
| A27 Li | 49.882 - 43.229 | 02-08-2025 | 12-08-2025 | 11 | 1548 |
| A27 Li | 65.059 - 58.712 | 13-09-2025 | 14-09-2025 | 2 | 2154 |
| A27 Li | 49.882 - 43.229 | 13-09-2025 | 14-09-2025 | 2 | 1914 |
| A27 Re | 43.261 - 49.836 | 13-09-2025 | 14-09-2025 | 2 | 1702 |
| A27 Re | 6.834 - 10.753 | 04-10-2025 | 05-10-2025 | 2 | 1340 |
| A27 Li | 65.059 - 58.712 | 18-10-2025 | 19-10-2025 | 2 | 2006 |
| A27 Re | 58.690 - 65.160 | 18-10-2025 | 19-10-2025 | 2 | 2294 |
| A28 Re | 120.818 - 128.189 | 22-03-2024 | 23-03-2024 | 2 | 1138 |
| A29 Re | 12.444 - 15.517 | 17-02-2024 | 18-02-2024 | 2 | 1387 |
| A29 Re | 12.444 - 15.517 | 24-02-2024 | 25-02-2024 | 2 | 1387 |
| A29 Re | 12.444 - 15.517 | 02-03-2024 | 03-03-2024 | 2 | 1392 |
| A29 Re | 12.444 - 15.517 | 09-03-2024 | 10-03-2024 | 2 | 1392 |
| A29 Re | 12.444 - 15.517 | 06-04-2024 | 07-04-2024 | 2 | 1501 |
| A29 Re | 12.444 - 15.517 | 27-07-2024 | 08-08-2024 | 13 | 1730 |
| A29 Re | 12.444 - 15.517 | 14-09-2024 | 15-09-2024 | 2 | 1404 |
| A29 Re | 12.444 - 15.517 | 26-10-2024 | 27-10-2024 | 2 | 1407 |
| A29 Li | 97.843 - 95.023 | 05-07-2025 | 06-07-2025 | 2 | 920 |
| A29 Re | 93.330 - 97.824 | 05-07-2025 | 06-07-2025 | 2 | 912 |
| A30 Re | 10.553 - 14.644 | 09-07-2024 | 26-08-2024 | 2 | 31433 |
| A30 Li | 14.541 - 10.504 | 28-08-2024 | 10-10-2024 | 2 | 36187 |
| A30 Li | 14.541 - 10.504 | 02-05-2025 | 15-10-2025 | 4 | 35235 |
| A31 Li | 34.654 - 31.401 | 08-07-2024 | 08-07-2024 | 1 | 598 |
| N31 Li | 52.953 - 50.897 | 28-09-2024 | 28-09-2024 | 1 | 544 |
| N31 Li | 52.953 - 50.897 | 29-12-2024 | 29-12-2024 | 1 | 448 |

*continued on next page*

*Table 22 – continued from previous page*

| Road | Location (km) | Start | End | Days | Baseline intensity |
|---|---|---|---|---|---|
| N31 Li | 52.953 - 50.897 | 08-07-2025 | 08-07-2025 | 1 | 757 |
| A32 Li | 57.635 - 48.984 | 19-09-2024 | 08-10-2024 | 20 | 608 |
| A32 Li | 69.062 - 66.105 | 10-09-2025 | 11-09-2025 | 2 | 920 |
| A35 Li | 47.456 - 44.020 | 15-12-2024 | 15-12-2024 | 1 | 431 |
| A4 Li | 75.543 - 74.311 | 16-03-2024 | 17-03-2024 | 2 | 1447 |
| A4 Re | 74.576 - 75.262 | 23-03-2024 | 24-03-2024 | 2 | 1514 |
| A4 Li | 75.543 - 74.311 | 29-06-2024 | 30-06-2024 | 2 | 1570 |
| A4 Li | 223.513 - 215.358 | 27-07-2024 | 04-08-2024 | 9 | 671 |
| A4 Re | 215.329 - 223.043 | 20-09-2024 | 22-09-2024 | 3 | 496 |
| A44 Re | 1.194 - 2.020 | 23-06-2025 | 25-06-2025 | 3 | 1283 |
| A44 Re | 19.638 - 21.645 | 23-06-2025 | 26-06-2025 | 4 | 795 |
| A44 Re | 8.311 - 10.722 | 23-06-2025 | 25-06-2025 | 3 | 1361 |
| A44 Li | 21.642 - 20.057 | 24-06-2025 | 25-06-2025 | 2 | 955 |
| A44 Li | 2.005 - .088 | 13-09-2025 | 14-09-2025 | 2 | 876 |
| A44 Re | 1.194 - 2.020 | 13-09-2025 | 14-09-2025 | 2 | 908 |
| A5 Re | 1.456 - 7.411 | 16-11-2024 | 17-11-2024 | 2 | 1390 |
| A5 Li | 7.261 - 1.422 | 23-06-2025 | 25-06-2025 | 3 | 1643 |
| A5 Re | 1.456 - 7.411 | 23-06-2025 | 25-06-2025 | 3 | 1758 |
| A5 Re | 14.951 - 18.854 | 13-07-2025 | 12-09-2025 | 62 | 1281 |
| A5 Li | 7.261 - 1.422 | 25-07-2025 | 17-08-2025 | 24 | 1420 |
| A50 Re | 113.472 - 117.519 | 28-06-2025 | 29-06-2025 | 2 | 1074 |
| A58 Li | 62.323 - 59.265 | 06-04-2024 | 07-04-2024 | 2 | 1717 |
| A58 Li | 53.907 - 47.746 | 07-09-2024 | 08-09-2024 | 2 | 1374 |
| A58 Li | 29.469 - 19.912 | 21-09-2024 | 22-09-2024 | 2 | 1260 |
| A58 Li | 53.907 - 47.746 | 16-11-2024 | 17-11-2024 | 2 | 1598 |
| A58 Li | 46.845 - 40.270 | 16-11-2024 | 17-11-2024 | 2 | 1503 |
| A58 Re | 59.266 - 62.184 | 29-03-2025 | 30-03-2025 | 2 | 1590 |
| A58 Re | 59.266 - 62.184 | 12-04-2025 | 13-04-2025 | 2 | 1602 |
| A58 Re | 40.621 - 46.833 | 10-05-2025 | 11-05-2025 | 2 | 1374 |
| A58 Re | 40.621 - 46.833 | 17-05-2025 | 18-05-2025 | 2 | 1374 |
| A58 Re | 40.621 - 46.833 | 24-05-2025 | 25-05-2025 | 2 | 1374 |
| A58 Re | 47.783 - 54.515 | 21-06-2025 | 22-06-2025 | 2 | 1446 |
| A58 Re | 40.621 - 46.833 | 21-06-2025 | 22-06-2025 | 2 | 1450 |
| A58 Li | 80.300 - 76.496 | 26-07-2025 | 29-07-2025 | 4 | 1270 |
| A58 Li | 163.751 - 155.533 | 13-09-2025 | 14-09-2025 | 2 | 688 |
| A58 Re | 100.753 - 103.356 | 31-10-2025 | 02-11-2025 | 3 | 1317 |
| A58 Re | 19.910 - 29.570 | 01-11-2025 | 02-11-2025 | 2 | 1427 |
| A59 Li | 151.253 - 146.510 | 24-08-2024 | 25-08-2024 | 2 | 1031 |
| A59 Re | 114.919 - 116.614 | 31-08-2024 | 01-09-2024 | 2 | 1114 |
| A59 Re | 60.489 - 65.221 | 10-03-2025 | 13-03-2025 | 4 | 514 |
| A59 Re | 114.919 - 116.616 | 12-07-2025 | 13-07-2025 | 2 | 967 |
| A59 Li | 110.048 - 107.902 | 26-07-2025 | 27-07-2025 | 2 | 936 |

*continued on next page*

*Table 22 – continued from previous page*

| Road | Location (km) | Start | End | Days | Baseline intensity |
|---|---|---|---|---|---|
| A59 Re | 96.950 - 100.002 | 18-10-2025 | 19-10-2025 | 2 | 1122 |
| A59 Re | 96.950 - 100.002 | 01-11-2025 | 02-11-2025 | 2 | 907 |
| A6 Re | 61.826 - 72.787 | 05-07-2025 | 06-07-2025 | 2 | 1374 |
| N65 Li | 6.957 - 5.940 | 31-07-2024 | 03-08-2024 | 4 | 812 |
| N65 Li | 6.957 - 5.940 | 28-12-2024 | 31-12-2024 | 4 | 796 |
| N65 Li | 6.957 - 5.940 | 03-01-2025 | 05-01-2025 | 3 | 740 |
| A65 Re | 18.236 - 19.196 | 06-09-2025 | 07-09-2025 | 2 | 836 |
| A67 Re | 24.075 - 27.650 | 22-06-2024 | 23-06-2024 | 2 | 1236 |
| A67 Re | 24.075 - 27.650 | 29-06-2024 | 30-06-2024 | 2 | 1236 |
| A67 Re | 24.075 - 27.650 | 29-03-2025 | 30-03-2025 | 2 | 1117 |
| A67 Re | 24.075 - 27.650 | 10-05-2025 | 11-05-2025 | 2 | 1200 |
| A67 Re | 24.075 - 27.650 | 17-05-2025 | 18-05-2025 | 2 | 1200 |
| A67 Re | 50.151 - 57.849 | 17-05-2025 | 18-05-2025 | 2 | 896 |
| A67 Re | 24.075 - 27.650 | 14-06-2025 | 15-06-2025 | 2 | 1461 |
| A67 Re | 50.151 - 57.849 | 14-06-2025 | 15-06-2025 | 2 | 1166 |
| A67 Re | 37.833 - 41.474 | 14-06-2025 | 15-06-2025 | 2 | 1176 |
| A67 Li | 57.904 - 50.098 | 13-09-2025 | 14-09-2025 | 2 | 888 |
| A7 Re | 140.217 - 142.616 | 25-10-2024 | 03-11-2024 | 10 | 606 |
| A7 Li | 132.512 - N7 Li (HR) 123.606 | 02-11-2024 | 03-11-2024 | 2 | 376 |
| N7 Re | 123.655 - A7 Re (HR) 132.620 | 02-11-2024 | 03-11-2024 | 2 | 388 |
| A7 Re | 140.217 - 142.616 | 05-11-2024 | 12-11-2024 | 7 | 64 |
| A7 Li | 132.512 - N7 Li (HR) 123.593 | 12-07-2025 | 13-07-2025 | 2 | 418 |
| N7 Re | 123.655 - A7 Re (HR) 132.620 | 12-07-2025 | 13-07-2025 | 2 | 480 |
| N7 Re | 123.655 - A7 Re (HR) 132.620 | 29-08-2025 | 31-08-2025 | 3 | 515 |
| A7 Li | 132.512 - N7 Li (HR) 123.593 | 30-08-2025 | 31-08-2025 | 2 | 377 |
| A7 Li | 132.512 - N7 Li (HR) 123.593 | 18-10-2025 | 19-10-2025 | 2 | 434 |
| A76 Re | 5.303 - 6.515 | 15-06-2024 | 16-06-2024 | 2 | 1372 |
| A76 Re | 5.303 - 6.515 | 29-06-2024 | 30-06-2024 | 2 | 1372 |
| A76 Li | 18.696 - 16.232 | 05-07-2025 | 10-07-2025 | 6 | 625 |
| A79 Re | 10.411 - 13.560 | 20-12-2024 | 29-12-2024 | 3 | 659 |
| A79 Re | 7.332 - 9.096 | 24-02-2025 | 25-02-2025 | 2 | 742 |
| A79 Re | 10.411 - 13.560 | 17-12-2025 | 24-12-2025 | 2 | 826 |
| A79 Re | 10.411 - 13.560 | 28-12-2025 | 31-12-2025 | 4 | 750 |
| A8 Li | 7.676 - 5.709 | 27-04-2024 | 27-04-2024 | 1 | 1392 |
| A8 Li | 4.689 - 2.783 | 27-04-2024 | 27-04-2024 | 1 | 2498 |
| A8 Re | 5.720 - 7.690 | 27-04-2024 | 27-04-2024 | 1 | 1410 |
| A8 Re | 2.783 - 4.792 | 27-04-2024 | 27-04-2024 | 1 | 2542 |
| A9 Li | 31.520 - 30.212 | 03-02-2024 | 04-02-2024 | 2 | 2051 |
| A9 Li | 31.520 - 30.212 | 01-06-2024 | 02-06-2024 | 2 | 2977 |
| A9 Re | 28.378 - 31.394 | 01-06-2024 | 02-06-2024 | 2 | 2538 |
| A9 Li | 31.520 - 30.212 | 15-06-2024 | 16-06-2024 | 2 | 2977 |
| A9 Re | 28.378 - 31.394 | 15-06-2024 | 16-06-2024 | 2 | 2538 |

*continued on next page*

*Table 22 – continued from previous page*

| Road | Location (km) | Start | End | Days | Baseline intensity |
|---|---|---|---|---|---|
| A9 Li | 31.520 - 30.212 | 10-08-2024 | 15-08-2024 | 6 | 2284 |
| A9 Li | 31.520 - 30.212 | 05-10-2024 | 06-10-2024 | 2 | 2028 |
| A9 Re | 28.378 - 31.394 | 26-10-2024 | 27-10-2024 | 2 | 2090 |
| A9 Li | 31.520 - 30.212 | 08-03-2025 | 09-03-2025 | 2 | 3416 |
| A9 Re | 28.378 - 31.394 | 08-03-2025 | 09-03-2025 | 2 | 1976 |
| A9 Li | 31.520 - 30.212 | 29-03-2025 | 30-03-2025 | 2 | 3416 |
| A9 Re | 28.378 - 31.394 | 29-03-2025 | 30-03-2025 | 2 | 1976 |
| A9 Re | 28.378 - 31.394 | 12-04-2025 | 13-04-2025 | 2 | 2014 |
| A9 Li | 31.520 - 30.212 | 03-05-2025 | 04-05-2025 | 2 | 1377 |
| A9 Re | 28.378 - 31.394 | 03-05-2025 | 04-05-2025 | 2 | 2232 |
| A9 y | 10.375 - 9.414 | 08-05-2025 | 26-05-2025 | 19 | 201 |
| A9 Li | 31.520 - 30.212 | 11-10-2025 | 12-10-2025 | 2 | 2235 |
| A9 Re | 28.378 - 31.394 | 11-10-2025 | 12-10-2025 | 2 | 2154 |
| A9 Li | 31.520 - 30.212 | 25-10-2025 | 26-10-2025 | 2 | 2235 |
| A9 Li | 10.499 - 7.000 | 08-11-2025 | 09-11-2025 | 2 | 604 |
| A9 Re | 7.065 - 10.502 | 08-11-2025 | 09-11-2025 | 2 | 724 |
| A9 x | 9.415 - 10.246 | 08-11-2025 | 09-11-2025 | 2 | 480 |
| A9 y | 10.375 - 9.414 | 08-11-2025 | 09-11-2025 | 2 | 215 |
| A9 Li | 31.520 - 30.212 | 15-11-2025 | 16-11-2025 | 2 | 2084 |
| A9 Li | 31.520 - 30.212 | 06-12-2025 | 07-12-2025 | 2 | 1747 |
| A9 Re | 28.378 - 31.394 | 06-12-2025 | 07-12-2025 | 2 | 1911 |
| N9 Li | 103.888 - 101.089 | 08-07-2025 | 02-08-2025 | 26 | 383 |
| N9 Re | 101.658 - 103.888 | 08-07-2025 | 02-08-2025 | 26 | 385 |

# C

# CENTROIDS

Table 23: Selected centroids, their coordinates, reason for inclusion and abbreviation. Where the study area name differs from the formal municipality name, the municipality is given in parentheses.

| CITY | ABBR. | REASON | LATITUDE | LONGITUDE | SOURCE |
|---|---|---|---|---|---|
| Alkmaar | ALK | ns32 | 52.6346 | 4.7615 | centroid |
| Almere | ALM | ns32 | 52.3725 | 5.2343 | centroid |
| Almelo | ALO | network | 52.3545 | 6.6613 | centroid |
| Alphen aan den Rijn | ALP | network | 52.1217 | 4.6490 | centroid |
| Amersfoort | AME | ns32 | 52.1735 | 5.3937 | centroid |
| Amsterdam | AMS | ns32 | 52.3599 | 4.8952 | centroid |
| Apeldoorn | APE | ns32 | 52.2063 | 5.9716 | centroid |
| Arnhem | ARN | ns32 | 51.9728 | 5.9022 | centroid |
| Assen | ASS | ns32 | 53.0027 | 6.5558 | centroid |
| Den Bosch ('s-Hertogenbosch) | BOS | ns32 | 51.7086 | 5.3235 | centroid |
| Breda | BRE | ns32 | 51.5892 | 4.7735 | centroid |
| Delft | DEL | ns32 | 52.0017 | 4.3539 | centroid |
| Deventer | DEV | passengers | 52.2619 | 6.1875 | centroid |
| Dordrecht | DOR | ns32 | 51.7981 | 4.6830 | centroid |
| Driebergen-Zeist | DRI | passengers | 52.0652 | 5.2611 | manual |
| Ede-Wageningen (Ede) | EDE | passengers | 52.0449 | 5.6566 | centroid |
| Eindhoven | EIN | ns32 | 51.4495 | 5.4712 | centroid |
| Enschede | ENS | ns32 | 52.2152 | 6.8924 | centroid |
| Geldermalsen (West Betuwe) | GEL | network | 51.8632 | 5.2269 | centroid |
| Goes | GOE | ns32 | 51.5059 | 3.8866 | centroid |
| Gouda | GOU | passengers | 52.0168 | 4.7131 | centroid |
| Groningen | GRO | ns32 | 53.2182 | 6.5740 | centroid |
| Den Haag ('s-Gravenhage) | GVC | ns32 | 52.0666 | 4.2974 | centroid |
| Haarlem | HAA | ns32 | 52.3823 | 4.6424 | centroid |
| Heerlen | HEE | terminus | 50.8991 | 5.9665 | centroid |
| Den Helder | HEL | terminus | 52.9318 | 4.7543 | centroid |
| Hilversum | HIL | passengers | 52.2251 | 5.1734 | centroid |
| Hoofddorp (Haarlemmermeer) | HOO | ns32 | 52.3076 | 4.6723 | centroid |
| Hoorn | HRN | network | 52.6548 | 5.0707 | centroid |
| Leeuwarden | LEE | ns32 | 53.1948 | 5.8034 | centroid |
| Leiden | LEI | ns32 | 52.1587 | 4.4875 | centroid |
| Lelystad | LEL | ns32 | 52.5072 | 5.4741 | centroid |
| Maastricht | MAA | ns32 | 50.8485 | 5.6951 | centroid |

*continued on next page*

Table 23 – *continued from previous page*

| CITY | ABBR. | REASON | LATITUDE | LONGITUDE | SOURCE |
|---|---|---|---|---|---|
| Meppel | MEP | network | 52.6983 | 6.1918 | centroid |
| Nijmegen | NIJ | ns32 | 51.8344 | 5.8413 | centroid |
| Roermond | ROE | network | 51.1940 | 6.0035 | centroid |
| Roosendaal | ROO | network | 51.5248 | 4.4533 | centroid |
| Rotterdam | ROT | ns32 | 51.9164 | 4.4785 | centroid |
| Schiphol | SCH | ns32 | 52.3094 | 4.7619 | manual |
| Tiel | TIE | terminus | 51.8845 | 5.4185 | centroid |
| Tilburg | TIL | ns32 | 51.5702 | 5.0676 | centroid |
| Utrecht | UTR | ns32 | 52.0930 | 5.0925 | centroid |
| Venlo | VEN | terminus | 51.3666 | 6.1571 | centroid |
| Vlissingen | VLI | terminus | 51.4584 | 3.5794 | centroid |
| Zaanstad | ZAA | ns32 | 52.4637 | 4.8006 | centroid |
| Zwolle | ZWO | ns32 | 52.5149 | 6.0917 | centroid |

# D

# CARD TYPE PASSENGER CATEGORIES

Table 24: OV-Chipkaart product types and their passenger category classification. Columns indicate whether the card type incurs a fare during dal (off-peak), spits (peak), and weekend hours respectively.

| DESCRIPTION | DAL | SPITS | WEEKEND |
|---|---|---|---|
| *1. Incidental* | | | |
| Consumenten RoS | y | y | y |
| Consumenten Enkele Reis | y | y | y |
| Consumenten Retour | y | y | y |
| Consumenten Dagkaart | y | y | y |
| Prijs Tijd Deals | y | y | y |
| Enkele Reis Daluren | y | y | y |
| Consumenten NS Flex Vol Tarief (BB) | y | y | y |
| Flex | y | y | y |
| *2. Business (employer-paid)* | | | |
| Zakelijk ERET | n | n | n |
| BC Basis | n | n | n |
| BC Dal | n | y | n |
| BC Traject Vrij | n | n | n |
| BC Trein Vrij | n | n | n |
| Resellers OVCP/OV-pay | n | n | n |
| BKA NS Resell oud | n | n | n |
| Resellers e-ticket | n | n | n |
| Zakelijk bij overige vervoerders | n | n | n |
| *3. Student* | | | |
| Studenten Week Vrij | n | n | y |
| Studenten Weekend Vrij | y | y | n |
| Studenten Week Vrij Saldo | y | y | y |
| Studenten Weekend Vrij Saldo | y | y | n |
| *4. Personal commute* | | | |
| Prepaid Traject Vrij | n | n | n |
| Prepaid Altijd Vrij | n | n | n |
| ROS Altijd Vrij Maand | n | n | n |
| Consumenten NS Flex Dal Onbeperkt (BB) | n | y | n |
| Consumenten NS Flex Spits Onbeperkt (BB) | n | n | n |
| Consumenten NS Flex Spits Korting (BB) | y | y | y |

*continued on next page*

Table 24 – *continued from previous page*

| DESCRIPTION | DAL | SPITS | WEEKEND |
|---|---|---|---|
| Consumenten NS Flex Traject Vrij | n | n | n |
| ROS Altijd Voordeel | y | y | y |
| ROS Dal Vrij | n | y | n |
| Consumenten Traject Vrij | n | n | n |
| ROS Altijd Vrij | n | n | n |
| *5. Social / Recreation* | | | |
| Consumenten Railrunner | n | n | n |
| Consumenten Samenreiskorting | y | y | y |
| Spoordeelwinkel Combideals | y | y | y |
| Groepsticket | y | y | y |
| Samenreisticket Enkele Reis Daluren | y | y | y |
| Toeristenkaarten (ATT/ARTT/HTT) | n | n | n |
| Jongerendagkaart | n | n | n |
| Dagretour Daluren | y | y | y |
| Meereisretour Daluren | n | y | y |
| Consumenten NS Flex Dal Korting (BB) | y | y | y |
| Consumenten NS Flex Weekend Onbeperkt (BB) | y | y | n |
| Consumenten NS Flex Weekend Korting (BB) | y | y | y |
| ROS Dal Voordeel | y | y | y |
| ROS VDU | y | y | y |
| ROS Keuzedag | n | n | n |
| ROS Weekend Vrij | y | y | n |
| ROS Kids Vrij | n | n | n |
| *6. Other* | | | |
| ⟨Niet gevonden⟩ | — | — | — |
| Consumenten Overig | y | y | y |
| Internationaal Vervoerbewijs | n | n | n |
| Thalys vervoerbewijs | y | y | y |
| Zitplaatsreserveringen | y | y | y |
| Overige vervoerbewijzen | y | y | y |

# E

# AI STATEMENT

This thesis was written during the dawn of artificial intelligence (AI) at a time where most students and teachers, but also companies and other individuals, are still finding the most efficient, secure, and fair way to use AI in daily use and, in my case, for writing academic research. For this thesis I have used AI. I set myself some guidelines and rules at the beginning which I have tried to adhere to throughout the process. First and foremost, AI is a tool. A tool which, when used carefully and efficiently, can help in making this research easier and allowing me to spend more time on actual research and less on the repetitive tasks. It also allows me to have a daily brainstorming "partner" at my fingertips. These two tasks is what I used AI for mainly.

**Rules** I set for myself:

1. Check everything AI outputs, don't take anything at face value
2. Always read the sources cited
3. Don't input sensitive data into public AI agents
4. Critically read and review any text that goes into the report, even if the AI-generated text was only a draft

**Guidelines**

1. When asking for references, always ask for the DOI or a link and read it
2. Spend time writing prompts so that the output is more complete
3. Try to do everything yourself, only if you get stuck, use AI

I used AI extensively for writing code. Whenever I knew what I wanted I drafted an extensive prompt which I then input into my preferred AI agent and asked it to write the code. These prompts were long and mimicked "pseudo-code", this way I had full control over what was coded. Afterwards I always scanned all the code and in each code I always asked the AI to output certain values so that I could manually verify that every step was being taken. Writing code often took multiple iterations. Using AI for this task saved me considerable amounts of time, to the point that I think I would not have been able to complete this research in time had I not used AI in this way.

I also used AI for most of the tasks that involved translating results into LaTeXreadable code, mainly for tables but also for some figures. When I had an internal block of what to write, or if I just couldn't find a good way to paraphrase something, I also asked AI to help me out. Usually this was in the form of asking it to write a few examples and I then used that as a starting point for what I wrote in the report. For spelling and grammar checks I put my entire thesis through an AI scan and it summarised my mistakes for me to correct. Literature is also easily found through AI and is, in my opinion, partially the future of literature research. However, it can not be ruled out that AI misses literature and therefore a good literature matrix with thorough search terms is still necessary. I mainly used AI for defining my search terms in the literature chapter.

During the 5-month period of this thesis I switched my AI agents multiple times after finding the old ones not to be sufficient enough, this highlights the rapidly changing environment. I started off using ChatGPT but this quickly became too slow and did not handle academic research well, a lot of the sources were "hallucinated" and I could therefore no longer trust what it output. I then switched to Anthropic's "Claude" AI which is lauded as an academic, high-level, AI-agent which always cites its sources and writes solid code that it tests itself (agentic AI). This did require a subscription which highlights that resources like this are becoming more and more a thing that only those with spare money can use. Personally I find this starting to be a moral problem. When some people who have the money to spend on AI agents they can produce better or more work, raising expectations for all students regardless of access. I therefore think that universities should provide free subscriptions to their students so that everyone is on the same level.

# F

# CODE

All code is on GitHub.

GitHub repository link: https://github.com/HugovdT/Hugo-Thesis-Git