Dwell time delays for commuter trains

An analysis of the influence of passengers on dwell time delays

RUBEN ALARIC KUIPERS
DEPT. OF TECHNOLOGY AND SOCIETY | FACULTY OF ENGINEERING | LUND UNIVERSITY

Printed by Media-Tryck, Lund 2024 NORDIC SWAN ECOLABEL 3041 0903

Faculty of Engineering
Department of Technology and Society
Division of Transport and Roads

ISBN 978-91-8039-924-1
ISSN 1653-1930

LUND UNIVERSITY

Dwell time delays for commuter trains

An analysis of the influence of passengers on dwell time delays

Ruben Alaric Kuipers

DOCTORAL DISSERTATION

by due permission of the Faculty of Engineering, Lund University, Sweden.
To be defended at the Faculty of Engineering, John Ericssons väg 1, in auditorium V:A on March the 15th of March, 2024 at 10.00.

Faculty opponent
Professor John Preston, University of Southampton

Dwell time delays for commuter trains

An analysis of the influence of passengers on dwell time delays

Ruben Alaric Kuipers

LUND UNIVERSITY

Faculty of Engineering

Department of Technology and Society

ISBN 978-91-8039-924-1 (print), 978-91-8039-925-8 (electronic)

ISSN 1653-1930

Printed in Sweden by Media-Tryck, Lund University

Lund 2024

”Kedeng kedeng kedeng kedeng kedeng. Oe oe”

Guus meeuwis

Foreword and acknowledgements

The work presented here in this thesis before you is the culmination of several years of hard work and a learning process that took me on a different journey than I would have imagined at the start of my PhD. A strong emphasis is placed on data analyses in this PhD, for example. So much so that I have even read statistics books in an attempt to understand what a p-value actually means. This is indeed not the journey I envisioned. If you asked young me during my bachelor's if I would do a PhD on dwell time delays and argue about non-parametric versus parametric testing, the answer would have probably been “Who are you, and what is a PhD?”. Luckily it was not until halfway through my masters that the topic of maybe doing a PhD popped up, and by that time I did know what it was. It was actually after an internship at the Division of Transport and Roads in 2018 that I realized I could be a student for several additional years and do research at the same time. From that point on I decided that I wanted to pursue a PhD position.

I was still not excited about doing statistics, however, and I must admit that while interviewing for this PhD position I was most excited about the prospect of conducting observations on the behaviour of people on platforms. One pandemic making observations on platforms impossible, and many hours of data handling and statistical analyses later, I present you with my thesis titled “Dwell time delays for commuter trains - An analysis of the influence of passengers on dwell time delays”. The thesis I present you here includes much more statistics than I would have imagined (I must also admit that it got fun) and a discussion about non-parametric statistical testing and effect sizes. As I said, not the journey I expected at the start but still a good one.

To reflect on the last few years I started by asking myself the classic questions. Was it fun? On average, yes. Was it easy? No, not always. Would you do it again? For sure. A big reason for this PhD being a nice time to reflect back on, and the thing which makes it something I would happily do again, is the community of researchers and PhD students at the Division of Transport and Roads at the LTH and at K2, and the people I had the chance to meet along the way. So to all of you, I would like to say thanks for making this a fun part of my life and a good time to reflect back on.

A special thanks goes out to my supervisors. Carl-William, Lena, and Nils thank you for your guidance through this journey, being there during the supervision meetings to guide me through the PhD but also make sure that I feel good about what I am doing and stay happy. An additional thank you to Carl-William for creating a research environment within the railway research group where we could learn from each other. Thank you for putting us all in the same room, it made the time as a PhD student a lot less lonely. This being said, I would have probably finished my work in three years if I was not put in a room with Michelle who is just as easily distracted as I am. Although not always productive it did make the days much more fun! Many thanks to Kah Yong as well, for your moments of wisdom and discussions, and for justifying the purchase of noise-cancelling headphones with your excellent usage of your inside voice. Luckily we have Grace to bring some calmness into the room, much respect to her for not only travelling halfway across the world but also managing to balance being a parent and doing a PhD at the same time! Thank you to Daria for always being optimistic and happy, something we could all use a little more of sometimes! Thank you also to Frida for joining as a research assistant at just the right time, standing on a cold parking lot collecting data so that I could actually work on my thesis, good luck now that you are a PhD student again. Thank you to Joe for visiting us, I think it is fair to say that we consider you part of the railway research group by now. Thank you also to our extended railway group people over at KTH. Ingrid, Niloofar, Emil, and Elin thank you for providing valuable input during our manuscript discussions and bringing additional joy to this journey. A big thanks to Natchaya, Taku, and Neba as well for the nice time working on our paper and attending RailBelgrade together.

I must also extend a thank you to my parents for supporting me to make my own choices. In fact, if you had not made me feel comfortable enough that it was okay to stop with my first attempt at doing a bachelor's study, I would have never studied transportation science and this thesis would never have been written by me. Thank you for supporting me in my choices to move to Belgium and Sweden, and no the Netherlands is probably not the next stop but we will make sure to have a guest room for you this time around, wherever we end up.

Last but not least a very very special thanks to Raphaela for joining me on this adventure to Sweden even though you do not like being cold and your summers would ideally include sunshine. Thank you for taking care of all of our admin in a foreign place, making sure I actually have an insurance, and booking our holidays so that we got to enjoy trips through Scandinavia. Without you my life would probably be a mess. Hopefully, we still have many years together to reflect back on this period of our lives!

Abstract

The thesis presented here delves into why dwell time delays for commuter trains occur, with a specific focus on the impact of boarding and alighting passengers. The overarching aim is to develop knowledge of how time delays arise to identify and describe potential ways in which dwell time delays can be reduced. In addition to this, a secondary aim is to identify how dwell times can be studied on a network-wide level. Six research papers are included in this thesis, which all contribute to the aforementioned aims. The first paper presents a literature review on the influence of passengers on dwell times. The five subsequent papers present different data analyses on the impact of passengers on dwell times and make use of several years worth of automatic passenger count data collected on board commuter trains in Stockholm and the region of Skåne in Southern Sweden.

The findings from these studies indicate that although the volume of passengers is often stated as the main cause for dwell time delays, this is not necessarily the case. The results, instead, suggest that the volume of passengers acts as an accelerator for the negative impact of other aspects of the dwelling process such as the behaviour of passengers. With regards to studying dwell time delays it is important to make use of robust measures and to present dwell time delays in terms of frequency and size rather than just an average value. In addition to this, the value of having data on a level of seconds rather than minutes is highlighted. The latter is important since a majority of dwell time delays are smaller than one minute.

Several avenues to reduce the risk of dwell time delays are proposed, based on the findings from the included studies. The first avenue to be explored is that of adopting a more dynamic approach to dwell time scheduling. In practice, this means that, in contrast to what is common practice in Sweden, different dwell times should be used during peak and off-peak hours and between different stations. In addition to this, it is important to account for the behaviour of passengers during the boarding and alighting process. This can be done by making use of platform management measures. The third avenue that is identified states that dwell time scheduling should take on a more network-wide approach rather than treating stations as a single entity. This is important since there are interdependencies between stations that influence the behaviour of passengers, such as the way passengers spread out. Working on these points will help to reduce the risk of dwell time delays. Although this thesis has an emphasis on the Swedish context, given the origin of the data, the above-mentioned avenues are likely to be applicable in different geographical settings as well.

Popular science summary

Recent years have seen increased efforts to improve the punctuality of trains. One of the reasons for this is that punctuality is a key performance indicator concerning passenger satisfaction and punctuality can thus be seen as important to both retain current passengers as well as to attract new users. Attracting new passengers can help to induce a shift away from private motorized transport, reducing greenhouse gas emissions from the transport sector. In addition to this, punctual railways can help to increase the level of accessibility for those who do not have access to another mode of transport. Having longer travel times due to delays can result in reduced access to the job market and society. Making trains run on time thus not only makes it green but also a fair mode of transport. However, despite the efforts made to improve the punctuality of railways it is still below the desired target levels, both in Sweden and in other parts of Europe, and further improvements have to be made.

Punctuality is closely related to delays, indeed when a train is not delayed it will most likely be punctual. Trains can suffer from a delay for various reasons and one type of delays are the so-called dwell time delays. Dwell time delays arise when a train is stationary at a station for longer than scheduled. It is these dwell time delays that are the focus of the work presented here. Dwell time delays are relatively small, often in the range of several seconds to a minute, but can accumulate larger delays over an entire journey and cause other trains to have to wait outside of the station when the platform is occupied for too long. The impact of dwell time delays can be even larger from the perspective of passengers. This is especially the case when there is a need to make use of several modes of public transport during a single trip. In such cases, a small delay can snowball when a connection is missed.

The thesis presented here focuses on why dwell time delays occur, the impact of passengers on dwell times, how dwell time delays can be measured, and finally how the risk of dwell time delays can be reduced. Several years of passenger counts collected on board trains in Stockholm and the region of Skåne in Southern Sweden were used to study and analyse the impact of passengers on dwell times. The results of these analyses indicate that, although the volume of passengers is often seen as the main cause of dwell time delays, this is not necessarily the case. Instead, the findings from the studies show that the volume of passengers acts as an accelerator for other aspects such as an uneven spread of passengers or passengers queuing up in front of doors. A high volume of passengers on its own is not sufficient to increase the risk of dwell time delays, but this will happen when there is a high volume of boarding passengers that are unevenly spread between the doors.

To better understand dwell time delays it is important to better measure these delays. Currently, the punctuality of trains is measured at the final station in Sweden and other places around Europe, meaning that dwell times are not actively measured. Instead, dwell times should be measured based on their size and frequency, allowing to not only understand how large dwell time delays are but also how often they occur. A more novel way to measure dwell time delays was also explored in this thesis, using the relative dwell time performance of stations and railway services. By combining both how often a train is delayed and how often a dwell time delay occurs at a given station, it was possible to identify that there is no such thing as a station where all trains are delayed but that it is often a small number of trains that have a delay.

The findings presented in this thesis can assist planners to more accurately schedule dwell times, by having better insights into both the size and frequency of dwell time delays along with a better understanding of where and when dwell times are likely to occur. Although scheduling dwell times more accurately will be beneficial, the findings presented in this thesis indicate that these benefits will be watered down when the behaviour of passengers is not addressed. In addition to scheduling dwell times better, efforts should thus also be made to steer passenger behaviour such as how boarding passengers spread out. This can be done by making use of platform management measures. Furthermore, it is important to introduce a network-wide approach to dwell time scheduling, and not see stations as a single entity. Dwell time delays at one station can be caused by what happens at another station and it is important to not ignore these kinds of relations.

Working towards having fewer to no dwell time delays will help to increase the punctuality of both trains and passengers, making railways a more viable and enjoyable way to travel, and help to start a shift to rail. Although the points mentioned here are the result of studies within a Swedish context they are likely to be valid in a wider context as well.

Table of Contents

1 Introduction

Recent years have seen an increased pressure to change our daily habits in order to reduce greenhouse gas emissions and slow down climate change. One of these daily habits that requires a change is the way in which we travel. The transport sector accounted for around 26% of the total greenhouse gas emissions in Europe in 2019 (European Environment Agency, 2022), of which a large part was a result of car travel. Passenger cars emit 0.132 kilograms of CO_2 per passenger kilometre when travelling on the highway. In contrast to this, trains have an average emission of 0.027 kilograms of CO_2 per passenger kilometre travelled (Jones et al., 2017). The lower emissions from railways compared to private motorized modes of transport make that a modal shift towards public transport is one of the ways to reduce greenhouse gas emissions (Blayac & Stéphan, 2021). Shifting from private motorized transport to the use of railways can potentially lead to a reduction of 0.105 kilograms of CO_2 per passenger kilometre travelled (Weber et al., 2022).

Although a shift towards railways is good for the environment, it is annoying if taking the train means you do not always know if you will arrive on time or not, a metric captured and formalized in the term *punctuality*. The Cambridge Dictionary defines punctual as "*the fact of arriving, doing something, or happening at the expected or correct time and not late*" (Cambridge Dictionary, n.d.). Taking this definition and the common experience with railway travel it is easy to see why work towards improving the punctuality of trains is needed because "*at the expected or correct time*" and "*taking the train*" seem to fit together like a bowl of soup and a fork on some days.

The on-time performance, or punctuality, of trains is one of the major determinants of the satisfaction of passengers (Brons & Rietveld, 2008; van Loon et al., 2011), with small delays already harming the perception of railways (Volovski et al., 2021). A study by van Lierop et al. (2018) found that the benefit of punctual services can even outweigh improvements in the on-board experience for passengers. The importance of travel time punctuality and reliability is also reflected in the value of travel time associated with delays. Parbo et al. (2016), for example, state that passengers place more importance on travel time certainty than on potential travel time reductions. One reason for the importance of punctuality over on-board comfort can be the type of travellers, commuters typically emphasise punctuality more since they travel during peak hours and their arrival or departure times are less flexible compared to leisure trips (Parbo et al., 2016).

Punctuality is closely related to the concept of delays, indeed when no delay is incurred one is expected to be punctual and vice versa. Train delays can be classified into two broad categories, one being run time delays which occur when a train is moving between two stations, and the other being dwell time delays which arise when a train is stationary at a station. The thesis presented here focuses on the latter of these and more specifically on dwell time delays for commuter trains.

Although dwell times make up a relatively small portion of the total travel time for commuter trains they are relevant to study given their strong relationship with the overall punctuality of trains, and thus also with passenger satisfaction. Despite the acknowledgement that dwell times are important, the underlying causes for dwell time delays are however not well understood (Harris et al., 2013). The overarching aim of this thesis is, therefore, to develop knowledge of how dwell time delays arise in relation to operational and passenger variables and describe potential ways in which dwell time delays can be reduced. The work presented here provides empirical evidence of different ways in which passengers and operational conditions influence dwell times. In addition to this overarching aim, a secondary aim is to identify how dwell times can be studied on a network-wide level.

1.1 Benefits of improved punctuality

A question that is often asked is why there is a need to strive for punctual trains, and what benefits are actually gained when a train arrives on time. Although the punctuality of trains is closely related to the operations of railways and improvements in the punctuality of trains will have benefits for railway operators, the potential benefits of improved punctuality will also transcend the world of railways into the wider society.

Given the importance of punctuality from the point of view of passenger satisfaction, it is feasible to expect that improving punctuality and reliability will result in increased ridership, and higher levels of ridership retention (Monsuur et al., 2021). The quality of public transport also has a direct effect on car ownership (Holmgren, 2020) and increasing punctuality can induce a modal shift away from private motorized transport in favour of railway travel, which helps to reduce greenhouse gas emissions. Increasing ridership has several societal benefits. First, it allows to further leverage the environmental benefits of railways as a mode of transport. Although seen as a green mode of transport, this benefit only really comes into effect when the passenger volumes are large enough (Givoni et al., 2009). In addition to making the air safer to breathe, train travel is also safer as a mode of transport compared to travelling by car by a factor of 50 (European Commission, 2021).

Increased ridership also affects the cost-effectiveness and profitability of the railway system. Operating a railway system more cost-effective will lead to less need for external funding, reducing the social costs of railway transport (Li & Preston, 2015). Furthermore, spreading the costs over a larger volume of passengers could allow for a reduction in ticket prices without a loss in revenue. This is important since the costs of public transport have been identified as a barrier to use by Mackett and Thoreau (2015). Although somewhat speculative, the examples above suggest that increasing ridership retention and attracting new passengers by providing punctual train services could potentially lead to a positive feedback loop in which the effect of increased ridership will lead to a further increase as a result of reduced ticket prices.

There are also more indirect societal benefits when the punctuality of railways improves. Accessibility to different opportunities, such as access to the job market, will be improved as the risk of long travel times due to delays decreases. This is especially relevant for those who do not have access to a private vehicle and are dependent on other modes of transport (Kawabata, 2003). This is especially the case for women, who make more use of public transport services for their mobility needs compared to men (Ng & Acker, 2018), and extended travel times have been shown to increase their distance to the labour market (Black et al., 2012). The latter becomes even more important when someone makes use of multiple modes of (public) transport since small delays can lead to larger delays when connections are missed (Rietveld et al., 2001; Vromans, 2005). A small delay of 30 seconds can lead to sitting in a bus station for 30 minutes due to a missed bus connection, for example. Ensuring that railways are a viable travel option can thus greatly benefit these groups of society and in turn make railways not only a green but also an equal form of transport.

1.2 The concept of punctuality in railways

As mentioned at the beginning of this introduction, the term *punctual* commonly refers to the act of being on time. The definition of punctuality used within the context of railways is, somewhat, different to this. Within the railway context, punctuality is a measure of adherence to the timetable based on a predefined level of acceptable deviation from the schedule (Olsson & Haugland, 2004). Similar to the term punctual, punctuality is a measure which says something about the difference between the scheduled time and the actual time. In this case, this is the arrival time of a train. Different to the term punctual is the use of an acceptable threshold within which a train is still considered to be punctual. This means that even if a train does not arrive exactly on time, it can still be punctual according to the definition used in the context of railways.

In broad terms, it is possible to measure punctuality based on the following three steps:

Step 1: define a threshold for a tolerable deviation from the agreed time.

Step 2: measure whether a train arrived at the specified time and within the tolerable deviation or not.

Step 3: divide the times that the train arrived at the specified time and within the threshold by the total number of train trips made.

An important aspect when measuring the punctuality of trains is to define the "acceptable deviation from the scheduled time" and to determine a suitable measuring point. In Sweden, this threshold is set at 5 minutes and 59 seconds and is measured at the final station of a train trip (Joborn & Ranjbar, 2022). In practice, this means that a train is only considered to be delayed when it arrives at the final station with a delay of 6 minutes or more.

The example above is valid for the Swedish railway network, within which the work done in this thesis took place. Different approaches to measuring punctuality are used across Europe. Both the threshold for what is considered to be a delayed train as well the measuring point are different between countries. Punctuality in Norway is commonly measured at the final stations and at some important stations such as Oslo. In addition to this, a distinction is made between local and regional trains and other types of services in Norway, where the former has a delay threshold of 3 minutes and the latter has a delay threshold of 5 minutes (Olsson & Haugland, 2004). Grechi and Maggi (2018) summarized how punctuality is measured differently for regional and long-distance trains in other countries across Europe, highlighting that the most common threshold for a delay is set at trains being more than 5 minutes late, slightly shorter than in Sweden. Some countries, such as Denmark, The Netherlands, and Spain use a much stricter threshold of around 3 minutes. Switzerland, known for its punctual trains, also utilizes a 3 minute delay threshold. Different to most other countries, the Swiss also measure the percentage of connections made. In this case, a connection is considered to have been made if the fixed interchange time between two trains at the transfer station is still provided to the passengers (SBB CFF FFS, n.d.). In contrast to other countries, the Swiss thus also provide a punctuality statistic from the point of view of passengers.

The examples above do not serve as a definitive list of punctuality measurements across all railway networks. Instead, these examples illustrate that although punctuality is often referred to as a measurement with a single meaning the way it is measured can be different between countries. As shown this difference can be found both in terms of the acceptable deviation, the type of trains considered, and the location where the measurement takes place.

It is worth noting that although terms such as punctuality, delays, and reliability are sometimes mentioned within the same context they indicate different things. For example, a train that is never punctual can still be reliable since one can rely on it not being punctual. Concerning the difference between punctuality and delays this difference is more technical. Delays are measured in time and punctuality is measured in percentages (Økland & Olsson, 2020). It is this percentage that is often formalized as a goal for the punctuality of railway services and acknowledged as a key performance indicator for railways (Olsson & Haugland, 2004). Punctuality is also often discussed in the media when talking about the service performance of railways (Joborn & Ranjbar, 2022) and included in contracts between infrastructure managers and operators (Noland & Polak, 2002).

In the case of Swedish railways, the goal is to have a punctuality of 95% (Järnvägsbranschens samverkansforum, 2021), meaning that 5% of the train trips are allowed to have a deviation from the scheduled time of more than 5 minutes and 59 seconds at the final station. This goal is, however, often not met in practice, and the punctuality of Swedish railways has been hovering around 90% between 2019 and 2021 in Sweden (Järnvägsbranschens samverkansforum, 2021).

Punctuality, or the lack thereof, is not just a Swedish issue. A 2021 report from the European Commission reports a steady decline in the punctuality of railways within the European Union. Here a maximum delay of 5 minutes is used to measure the punctuality of railways. The findings show that the overall punctuality of trains within the European Union dropped to 90% in 2018, compared to 93% in 2012 (European Commission, 2021). So even though the work that is presented in this thesis is performed within a Swedish context, the outcomes are also of interest to other (European) countries.

1.3 The concept of dwell times

The thesis presented here focuses on dwell times and it is, therefore, worth to dive deeper into the concept of dwell times. Within the context of commuter trains, dwell times refer to the time a train is stationary at a station allowing for passengers to board and alight. The actual dwell time is commonly measured as the difference between the arrival and departure time of a train (Li et al., 2014). Dwell times are, sometimes, referred to in terms of the *minimum dwell time*. The minimum dwell time corresponds with the minimum time needed to complete the boarding and alighting process and to depart from a station (Goverde, 2005; Pedersen et al., 2018). This minimum dwell time does not include any additional time such as a dwell time buffer or time to allow for fluctuations in the dwelling process, passenger flows and arrival delays (Goverde, 2005) and is thus not always a realistic dwell time.

One reason why the minimum dwell time is often not achievable in practice is due to the stochastic nature of the dwelling process. Although dwell time is commonly referred to as a single process, it consists of several different sub-processes (Buchmueller et al., 2008; Goverde, 2005; Heinz, 2003) with both static and dynamic time elements (Seriani et al., 2019b). A schematic overview of the dwell time process is shown in Figure 1, with static elements shown in white and the dynamic element of dwell time shown in orange.

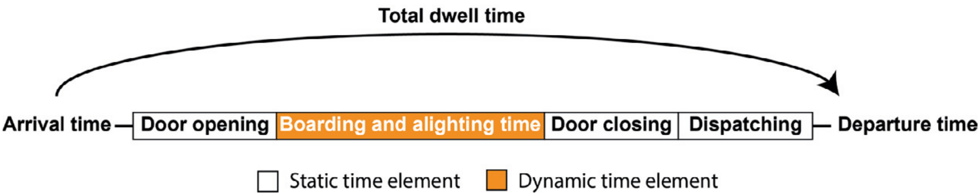

Figure 1: Schematic overview of the dwell time process. The orange colour indicates the dynamic elements and white boxes indicate which processes make up the static time elements of the total dwelling process.

The static time elements are governed by the technical aspects of the railway system and can be considered to be system constants (Heinz, 2003). The time needed to complete these static time elements can differ between train types. For example, the time it takes for the door to open depends on whether or not passengers need to request for the door to be opened, and the speed at which doors can open. Some trains also have sliding extensions to facilitate a level entry. When such a sliding extension is present this will extend the time needed to open the door since this process cannot begin before a train is stationary and must be completed before a door opens (Buchmueller et al., 2008). The time needed to close the doors is defined by the same technical features. The dispatching time refers to the time needed for the departure procedure after all the doors are closed during which the train driver prepares for departure or waits for permission to depart (Goverde, 2005).

Where the static time elements of dwell time are related to the technical features of the system, the dynamic time element of dwell times is governed by the boarding and alighting time (Goverde, 2005). The boarding and alighting time is defined by the time it takes for the boarding and alighting to be completed at the "slowest" door (Buchmueller et al., 2008), which is also known as the critical door. Several different reasons exist as to why the time needed for boarding and alighting is not a constant, which will be described in a later section of this thesis.

1.4 Terminology for dwell times used in this thesis

The terms *delay*, *deviation*, and *punctual* are used throughout this thesis and it is worth pointing towards the meaning of these terms within the context of the work presented here. The definition of the Cambridge Dictionary for punctual, which defines punctual as "*the fact of arriving, doing something, or happening at the expected or correct time and not late* ", is used when describing whether a dwell time is punctual or not meaning that a punctual dwell time is defined as "*the fact of the length of the actual dwell time corresponding with that of the scheduled dwell time*". Here the actual dwell time is the time measured by systems on board a train and the scheduled dwell time is defined in the timetable.

A further distinction is made between dwell time delays and dwell time deviations. In this context, the term *dwell time delays* exclusively refers to a situation where a train is stationary at a station for longer than scheduled, and departs after its scheduled departure time. The term *dwell time deviation*, on the other hand, refers to dwell processes that are either longer or shorter than scheduled and thus includes both dwell time delays as well as the recovery of delays.

A graphical representation of different "types" of dwell time is shown in Figure 2. As stated before, dwell times are considered *punctual* when the actual dwell time is equal to the scheduled dwell time. This happens when a train arrives and departs according to the scheduled time, but it can also happen when a train arrives with a delay and departs with a delay of an equal size. In that case, the actual dwelling process is still completed within the scheduled time despite the train departing with a delay, and there is thus no dwell time delay as such.

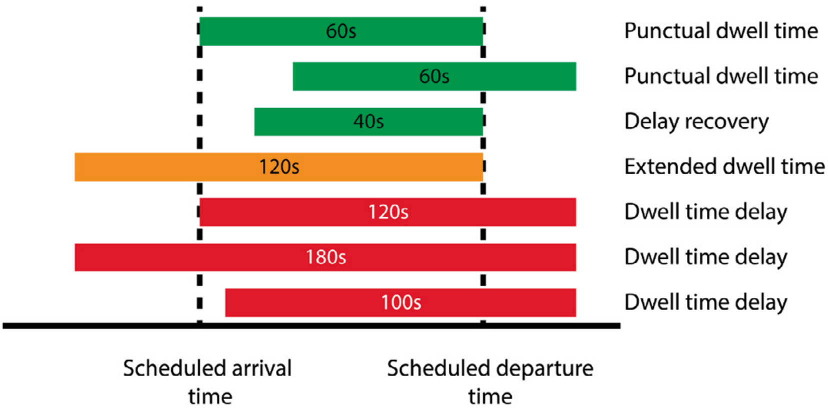

Figure 2: Graphical representation of dwell time deviations, with actual dwell times indicated.

Dwell time deviations can either be classified as *delay recovery*, as *extended dwell time*, or as a *dwell time delay*. Delay recovery takes place when a train arrives with a delay and departs without a delay. In this case, the dwell time deviation is negative, indicating that the dwelling process took less time than scheduled and the train was able to recover some of its delayed time.

Extended dwell times arise when a train arrives early and can be attributed to trains having to wait longer than scheduled since it is not allowed to depart ahead of the scheduled departure time (Coulaud et al., 2023; Kecman & Goverde, 2015). These extended delays are, however, not necessarily a delay as such when the train still departs at the scheduled time. This does not mean that such cases are positive since reoccurring instances of extended dwell times suggest that the available capacity cannot be optimally utilized. The third type of dwell time deviation is that of dwell time delays where both the dwelling process takes longer than scheduled and the train departs after the scheduled departure time.

1.5 Dwell time delays and the impact on punctuality

Even though dwell times make up a small part of the total travel time, commonly scheduled somewhere between thirty seconds and two minutes, the impact of dwell time delays on the punctuality and robustness of passenger train services is rather large (Harris, 2005; Palmqvist, 2019; van den Heuvel, 2016). Small dwell time delays can accumulate over a journey and result in a larger delay (Christoforou et al., 2020) meaning that a train with several small dwell time delays along its journey can end up with a sizeable delay at the final station. This delay can then exceed the tolerable deviation from the scheduled arrival time, even though no major events occurred to cause the train to be delayed. Furthermore, a train departing with a small delay of only a few seconds can go unnoticed but this small delay reduces the probability of an on-time arrival at the subsequent station (Vromans, 2005). In line with this, Luethi et al. (2005) found that dwell time delays were a major cause of delays in the Zurich area, stating that the buffer time between stations was not sufficient to absorb these delays. This notion is, somewhat, contrasted by Denti and Burroni (2023) who state that a departure delay can be recovered between stations but provide no analysis to which extent this occurs.

It is worth pointing out that dwell time delays do not only affect the punctuality of the train incurring the delay but can also cause so-called *knock-on delays*. This can happen when a delayed train blocks access to a station or platform for a subsequent train (Yamamura et al., 2012). These knock-on delays are more likely to in dense networks and when services are run with short headways (Goverde, 2005). The dwell time can become a major constraint for operations in such cases, and may prevent the ease of delay recovery (Daamen et al., 2008; Harris et al., 2022).

Despite the potentially large impact of dwell times on punctuality, only a few studies quantified this relationship. Palmqvist and Kristoffersson (2022), studying how both run and dwell time delays influence punctuality, identified a strong link between the frequency of dwell time delays and the punctuality of railways. The authors state that reducing the frequency of dwell time delays is one way to achieve an improvement in the overall punctuality of railways. Describing the punctuality of trains in Norway, Olsson and Haugland (2004) state that increasing dwell times by a few minutes at stations along a single-track line can have large effects on overall punctuality. In the example, the authors describe, the punctuality of trains increased from 68.5% in 2002 to 85.7% during the same period in 2003 after dwell times were extended on some stations along a long-distance line.

1.6 Structure of this thesis

The remainder of the thesis is structured as follows. Section 2 provides an overview of the relevant background information. Here the concept of timetable planning is briefly explained and an explanation of how dwell time for commuter trains fits within the timetabling process is provided. This section also explains how dwell times can be modelled, and different causes for dwell time delays. This leads to an overview of the different ways in which passengers influence the duration of dwell times and how passenger flows are influenced by the design of rolling stock. The background section is rounded off with an overview of the different ways in which passenger count data can be collected on board of commuter trains.

The background section is followed up by an overview of the scope of this thesis, including the problem outline, research gap, and research questions in Section 3. The list of papers included in the thesis is presented in Section 4, which also includes the declaration of contributions and the relationship between the papers themselves and between the papers and the different research questions presented in Section 3.

The methods that were used in the included papers are described in Section 5, and the available data and data processing is presented in Section 6 and Section 7 respectively. A summary of the included papers is provided in Section 8, with the answers to the research questions being presented in Section 9. The thesis is rounded off with a reflection on the data and methods used in Section 10, the contributions of the thesis to research and policy in Section 11, the limitations in Section 12, an overview of future research in Section 13, and the conclusion in Section 14.

2 Background

The following chapter provides an overview of timetable planning and dwell time scheduling practices, as well as a short overview of dwell time models proposed in the past and why the results of such models are not always valid. This is followed by a description of different causes for dwell time delays, a description of how passengers influence dwell times, and a description of how passenger flow data can be collected within the context of commuter trains.

2.1 Timetable planning

Timetable planning refers to the task of scheduling when trains are where, matching the trains to the available infrastructure in both space and time (Goverde, 2005). This thesis will not delve too deep into the actual process of timetable planning itself since the focus is placed on dwell times. It is nevertheless worth briefly explaining the timetable planning process in Sweden, given the location where the research presented in this thesis took place. This description is based on an explanation provided by Palmqvist et al. (2018). In Sweden, it is the Transport Administration that acts as the infrastructure manager and supplies the capacity on the railway and it is the train operating companies that demand the use of these tracks.

In general terms, the train operating companies send a request for capacity on the tracks to the Swedish Transport Administration where planners combine these requests and define a draft of the timetable. This draft timetable can still change depending on potential conflicts and disputes that arise, such as cases where two train operating companies want to make use of the same timeslot for a given section of track. Once these problematic cases are taken care of, a final operational timetable is produced. Going from a timetable design to a final timetable it thus an iterative process in which many things needs to be balanced. The final timetable is important as it forms the backbone for the successful operation of trains, becoming the means of communication between different actors and providing an indication to railway passengers regarding when and where trains will run (Goverde, 2005; Vromans, 2005).

2.1.1 Run and dwell times

Although somewhat of an oversimplification, it is possible to say that there are two main processes that make up the timetable and these are the run and dwell time of trains. Both these elements can be shown by making of use a *graphical timetable*, also known as a *time-distance diagram* (Goverde, 2005), of which an example is shown in Figure 3. The stations along the line are shown on the y-axis and the time is shown on the x-axis. The example shown here includes two directions, one being trains going from station A to station G, which is shown in red, and in the opposite direction which is shown in green. The vertical lines depict the run times, where the slope indicates the speed, and dwell times are depicted by the horizontal lines where the length of the line indicates the length of the dwell time.

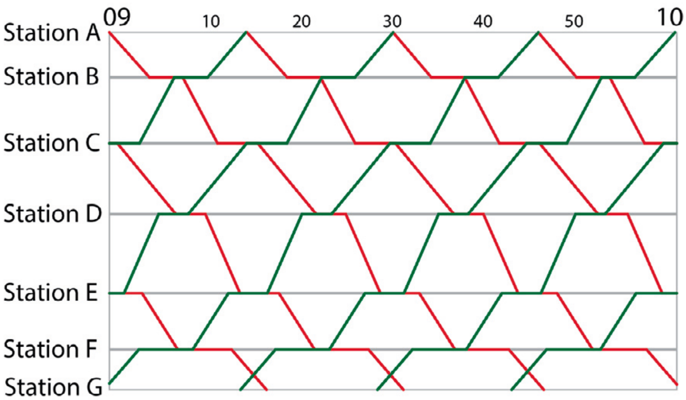

Figure 3: Simplified example of a graphical timetable.

The run time dictates the time it takes for a train to travel between two subsequent stations. The technical minimum running time is defined based on the track and train characteristics which determine the maximum running speed and the spacing between two stations (Goverde, 2005). This minimum time is often not realistic, however, due to driver behaviour for example, and margins in the form of running time supplements are added to the minimum technical running time (Vromans, 2005). These supplements are useful for the robustness of timetables since they allow for the absorption of small disturbances.

To have a timetable that performs well it is important to include both precise and realistic running and dwell times (Buchmueller et al., 2008; Hansen, 2010). This is a balancing act between allowing enough time for each task set out within a timetable and planning a dense enough timetable to allow for a sufficient frequency of trains (Goverde, 2005). The latter is not only important from a passenger perspective but also from the perspective of network utilization, which can be measured as the number of realized train kilometres for each kilometre of rail in a network (Vromans, 2005). Being able to run trains at a higher frequency results in a better use of the available capacity for railway operators.

2.1.2 Stability and robustness

Timetable planners often aim for a stable and robust timetable. Stability, in the context of a railway system, is a measure of both the time and effort required for the railway system to return to its normal state after a disturbance (Vromans, 2005). Two examples highlighting a different degree of stability are shown in Figure 4. A stable system, shown in Figure 4a, requires little time to return to its original state, whereas the system performance will be reduced for a longer period in an unstable system, as shown in Figure 4b.

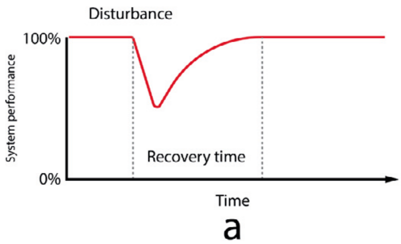

Figure 4: Example of different degrees of stability for railway systems, with a stable system in “a” and an unstable system shown in “b”.

A similar notion can be made regarding robustness, where the effect of disturbances on the operation of a system will be limited when said system is robust, as shown in Figure 5a, and the system performance will deteriorate more when the timetable is less robust, as shown in Figure 5b. A robust railway system can also be defined as a system in which (small) external influences do not cause large delays and delays do not propagate through the network, and the opposite is true when a system is not robust (Vromans, 2005).

Figure 5: Example of different degrees of robustness for railway systems, with a robust system in “a” and an unstable system shown in “b”.

Small fluctuations in the actual run time are inevitable, and margins or supplements are often added to the run time to ensure that this does not lead to delays. These margins provide additional run time that can be used to absorb small fluctuations. The size of these margins is formalized in national guidelines in Sweden (Palmqvist, 2019), where run time supplements are set at a minimum of three percent of the running time across all cases, with headway between two trains set to be at least two to seven minutes. This headway is most commonly in the range of three to five minutes. Dwell times can include some buffer times as well (Goverde, 2005), which serve as a way to ensure robustness by scheduling slightly longer dwell times which allow a delayed train to recover some of its delay.

2.1.3 Cyclicity and symmetry of timetables

The next concept to describe is the cyclicity and symmetry of timetables. It is common to make use of cyclic timetables in many countries, including Sweden, in which all trains are operated with some fixed time intervals (Vromans, 2005). Cyclical timetables are usually also symmetrical. Symmetry in timetables indicates that the travel time from station A to station B and the travel time from station B back to station A is more or less the same. The transfer times between trains are kept the same in both directions as well when making use of a symmetrical timetable (Vromans, 2005).

Cyclic and symmetric timetables have some benefits. Cyclic timetables are, for example, easier to remember for passengers since trains leave at the same time within each cycle. This means that a train that leaves at 8:10 will also leave at 9:10 and leave again at 10:10. This repetitiveness increases the ease of use of trains as a travel mode (Robenek et al., 2016; Vromans, 2005) since increased clarity lowers travel resistance (Annema, 2009). A cyclic and symmetric timetable has benefits for the operator as well since it only requires the operator to plan for one cycle and this plan can be repeated throughout the remainder of the operating window (Vromans, 2005). Symmetric timetables also allow to better schedule the transfer of passengers at large nodes, known as symmetric nodes, where all trains from all directions arrive at the same time and depart at the same time (Vromans, 2005). This again with the aim to increase the ease of use for passengers.

However, as the late great Johan Cruijff famously said "Every advantage has its disadvantage" (Cruijf, n.d.), and this also holds true for cyclic timetables. As Vromans (2005) mentions, for example, cyclic timetables are not flexible and cannot be adapted to fluctuations in passenger demand or specific wishes for operators. This can result in trains being overloaded during peak hours and empty during off-peak hours, as well as dwell times being too short during peak hours or too long during off-peak hours. This problem can be overcome by using a cyclic timetable for different periods of the day, making it possible to schedule extra trains or time during peak hours and reducing this during off-peak hours (Goverde, 2005).

2.2 Dwell time scheduling

An important part of the timetable is the scheduled dwell time (Buchmueller et al., 2008). Dwell times often follow the same cyclicity as the timetable and this means that the same dwell time is scheduled for both peak and off-peak hours and between weekdays and weekends. Even though scheduling the same dwell time during both peak and off-peak hours is beneficial to ensure cyclicity and symmetry in the timetable, this static approach can lead to unrealistic timetables and can cause the actual dwell times to exceed the scheduled dwell times on a regular basis (Goverde et al., 2001; Nash et al., 2006; Palmqvist et al., 2020). For example, Pedersen et al. (2018) mention that dwell times in Sweden are scheduled based on the time needed during off-peak hours. As a result of this approach, there is little to no slack in the dwell time during the peak hours, where longer dwell times are more likely to occur, and the likelihood of delays during peak hours increases as a result of this. A similar notion concerning making use of the off-peak hours as the normative time to base the scheduled dwell time on is made by Olsson and Haugland (2004) regarding dwell times in Norway. When scheduling dwell times in a static manner one effectively inserts delays into the timetable by not accounting for fluctuations in the time needed to complete the dwelling process at different times of the day and by ignoring station-specific characteristics.

Despite the importance of dwell times for the operation of railways, the actual dwell time scheduling process is sparsely described. It is often stated that dwell times are generally scheduled based on general assumptions and guidelines, rules of thumb, and experience from the past (Christoforou et al., 2020; Wiggenraad, 2001). This approach based on general rules and guidelines is also present in Sweden (Palmqvist, 2019), where 2 minutes is the standard and a dwell time of 1 minute is scheduled when the number of passengers is small (Palmqvist et al., 2018). Volovski et al. (2021) describe three methods to forecast dwell times proposed in the "Transit Capacity and Quality of Service Manual", used in the US. Two of these methods estimate the upper bound of the scheduled dwell time for a given station based on the mean dwell time that was previously experienced at the station of interest or at similar stations. An operational margin of either 15 to 20 seconds or one to two standard deviations is subsequently added to this time. The third method described by the authors involves modelling dwell time as a function of passenger flow and operational variables. The authors do not mention if these methods are used in practice or not.

2.3 Modelling dwell times

Scheduling dwell times is a non-trivial task since it includes several different factors related to passenger demand and different modelling approaches have been developed in the past in an attempt to assist planners when scheduling dwell times. This section provides a brief overview of some dwell time models that have been proposed. For a more in-depth overview, see the work done by Yang et al. (2019).

In their review of dwell time models, Yang et al. (2019) classify the models as being either statistical models or simulation-based models. The examples provided on the latter mostly focus on the simulation of passenger movement, either on a platform (Ahn et al., 2016) or inside a train (Baee et al., 2012). Such models do not provide insights into the necessary dwell time and are, therefore, not discussed in the following overview which is limited to the use of statistical models to predict the necessary dwell times.

The most notable example of a statistical dwell time model, often referred to in dwell time literature, is the model proposed by Weston (1989) and is shown in Equation 1.

$$SS = t_0 + \left[1.4 * \left(1 + \frac{F}{35} * \left(\frac{(T-S)}{F}\right)\right)\right] * \left[\left(F * \frac{B}{D}\right)^{0.7} + \left(F \frac{A}{D}\right)^{0.7} + \left(0.027 * \left(F * \frac{B}{D}\right)\left(F * \frac{A}{D}\right)\right)\right] \qquad (1)$$

This model, designed for the London underground, can be used to calculate dwell time here referred to as service time (SS) based on the volume of boarding (B) and alighting (A) passengers and the ratio between the busiest and average door (F). The model also takes the number of through passengers (T), referring to the number of passengers that stay on the train during the dwelling process, into account. The values used in this model are aggregated on a train level. The model also requires inputs on the number of seats (S) and doors on a train (D), as well as a constant for the technical time needed to open and close the doors of a train (t_0) which is set at 15 seconds in the case of the London Underground. Although developed in the late eighties, the model has been tested more recently by Harris and Anderson (2007) who found the model to have validity around the world, although some of the values in the model had to be varied slightly.

Another model that is often referred to is the model proposed by Puong (2000) who developed a linear regression model which includes fewer variables compared to the model proposed by Weston (1989). The model proposed by Puong (2000) is shown in Equation 2:

$$DT = A_d + B_d + TS_d \qquad (2)$$

Where the dwell time (DT) is a function of the alighting passengers per door (A_d), the boarding passengers per door (B_d) and the number of through standees (TS_d).

Both of these models require detailed information on passenger flows which is not always available. Some models have recently been proposed that do not make use of the number of passengers to predict dwell times to overcome this issue. Examples of these so-called passenger-disregarded dwell time models can be found in the models proposed by Kecman and Goverde (2015) and Li et al. (2016). Passenger-disregarded dwell time models do come with their own drawbacks, however. For example, Li et al. (2016) conclude that their models have potential but only during peak hours. Furthermore, both Kecman and Goverde (2015) and Li et al. (2016) state that including passenger-related variables will improve the accuracy of their models.

Whilst the above-mentioned models can be helpful when scheduling dwell times, the lack of research into these models is identified in the Transit Capacity and Quality of Service Manual as a major impedance to the acceptance and implementation of dwell time models by transit agencies (Volovski et al., 2021). Caution is also needed when interpreting the outcome of dwell time models. Although not often clearly stated, the objective of most dwell time models is to determine the minimum dwell time. This means that the results of the models indicate the shortest possible dwell time given the technical time and the shortest possible boarding and alighting time as a function of the number of passengers. A major assumption made in these models is that the boarding and alighting rate stays more or less stable throughout the boarding and alighting process and is the same across all stops for a given train type. Both of these aspects have been found to not be true, however. A study in Switzerland found that dwell times can be different across platforms at a single station served by the same type of trains (Gysin, 2018). Furthermore, several studies have highlighted that boarding and alighting rates can differ throughout the boarding and alighting process (Fernández et al., 2015; Thoreau et al., 2016). Harris et al. (2014) state that the range of alighting rates is as large as 0.4 and 2.6 passengers per second and boarding rates are in the range of between 0.3 and 2.1 passengers per second. The fluctuations in the flow rate of passengers is one possible explanation as to why the minimum dwell time is often not the same as the time that is actually needed to complete the boarding and alighting process. This also means that scheduling based on the minimum dwell time is thus likely to result in dwell time delays.

2.4 Causes for dwell time delays

The next section describes some of the possible reasons for dwell time delays. Based on the literature the following general reasons for dwell time delays can be identified (Harris et al., 2013; Pritchard et al., 2021):

- Design of the infrastructure
- Rolling stock design and operation
- Timetabling/train scheduling
- Train control and operation
- External factors (e.g. weather)
- Station operation
- Station design
- Behaviour of passengers
- Behaviour of train staff
- Passenger characteristics

Not all of the above-mentioned reasons have received similar attention in the literature and the following descriptions are, therefore, limited to the behaviour of train staff, the state of the rolling stock, the behaviour of passengers, and the design of rolling stock with the latter two being described more in-depth in *Section 2.5* and *Section 2.6* respectively.

With regard to the effect of train staff on the duration of dwell times, Harris et al. (2013) state two ways in which they can negatively influence the duration of dwell times, this being their ability to control the flow of passengers as well as staff members holding open train doors for late-arriving passengers. The authors state that if the train staff is not correctly positioned near entrance points, their ability to control the flow of passengers is reduced. Having reduced control of the flow of passengers can lead to late-arriving passengers being able to board the train just before it departs, which leads to extended dwell times since the doors cannot be closed on time. Better staff placement was found to reduce the average delay by four seconds. Another aspect highlighted is that of excessive service by staff members where staff members, with the best intention, hold the train to allow late-arriving passengers to still board. This behaviour was found to extend dwell times by up to twenty-five seconds.

Harris (2015) also identified poor communication between train attendants and the driver to be a possible cause of dwell time delays, although the effect is only small, with it being around one or two seconds. Staff lateness was mentioned by Carey and Carville (2000) as a possible cause for a delay. In such a case a train cannot depart due to the crew not being present to operate the train. It is worth noting that staff lateness is not necessarily caused by staff members themselves, as this can also be the result of major disruptions resulting in cases where the train staff displaced, for example (Pritchard et al., 2021).

The state of the rolling stock has been identified as a potential reason for dwell times to be delayed. An example of this is a situation where one or more doors of the train are broken. A reduced number of doors can be used in such a situation and passengers have to reposition themselves upon finding out the door they planned to use is broken, putting extra pressure on the nearest doors to the broken one. Research on the impact of this is limited, however, a study by Dinmohammadi et al. (2016), using data from 38 trains in the UK, found a total delay time of 518 minutes as a result of broken doors. When the train doors do work, the way in which they are opened affects the technical time. Studies found that if unbeknownst to the passengers a door does not open automatically the dwell time can be extended as passengers have to actively open the train door (Douglas, 2012; Harris, 2015).

2.5 Passengers and dwell time

Dwell time is sometimes also referred to as passenger service time (Fernández et al., 2008; Heinz, 2003), a fitting name since dwell time is most of the times explicitly scheduled to allow for the exchange of passengers. Furthermore, the time it takes for the exchange of passengers can be seen as the main determinant for the length of dwell times. The effect of the boarding and alighting time on dwell times has been illustrated by Buchmueller et al. (2008) and an adapted version of their illustration is shown in Figure 6.

Since trains cannot depart before all doors are locked, it is the door where the boarding and alighting process takes the longest time that defines the length of dwell times. This door is also known as the *critical door*. The critical door, Door 2, is indicated in orange in Figure 6. Comparing the length of the different sub-processes it can be seen that it is indeed the time it takes for the boarding and alighting to be completed that makes this door the critical door. In fact, door 3 opens at the same time as door 2 but can be closed earlier due to a shorter boarding and alighting time.

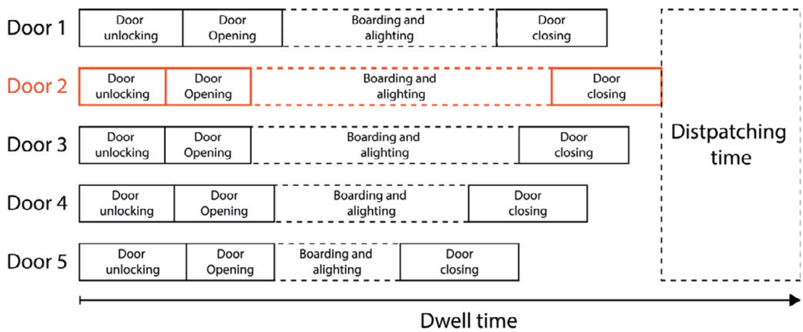

Figure 6: Influence of boarding and alighting time on dwell times, the critical door is indicated in orange (Image adapted from Buchmueller et al. (2008)).

In their conceptual model of dwell times, Li et al. (2016) state that the boarding and alighting time is influenced by the number of passengers and their flow rates. Studying flow rates for both boarding and alighting passengers, Harris and Anderson (2007) found that these are typically in the range of 1 passenger per door per second but that these flow rates are situation-specific. A later study by Harris et al. (2014) found that the possible range of alighting rates is as large as 0.4 and 2.6 passengers per second and boarding rates were found to fall between 0.3 and 2.1 passengers per second. In addition to this, studies have found that passenger flow rates can differ between platforms of the same station served by the same trains (Gysin, 2018), as well as throughout the boarding and alighting process itself (Fernández et al., 2015; Thoreau et al., 2016).

The interaction between boarding and alighting passengers is mentioned by Harris (2005) as one of the possible reasons for the differing passenger flow rates. This interaction takes place in an area sometimes referred to as the platform train interface area, or PTI for short (Holloway et al., 2015; Rodríguez et al., 2015; Seriani, Fujiyama, & de Ana Rodríguez, 2016). One way to depict the platform train interface area is as a semi-circular area located in front of the train doors, as shown in Figure 7.

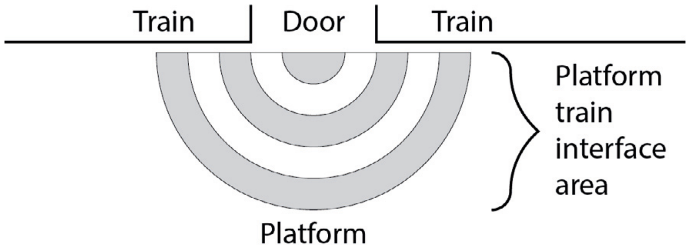

Figure 7: Schematic representation of the platform train interface area.

2.5.1 *What happens in the platform train interface?*

Two distinct processes that take place in the platform train interface area can be identified, this being the positioning of passengers in relation to the doors of a train and the formation of boarding and alighting lanes. Seriani and Fujiyama (2019) found that passengers waiting to board tend to stand more to the middle of the door when the number of boarding passengers is higher. This can, somewhat, be explained by findings suggesting that stress levels increase as a result of crowding when the number of boarders is large, resulting in passengers waiting to board to show less organized and civilized behaviour (Heinz, 2003; Hirsch & Thompson, 2014).

These situations where boarding etiquette is not followed can lead to passengers boarding a train before the alighting is finished (Seriani, Fujiyama, & Holloway, 2016) resulting in an overlap between boarding alighting passenger flows. This interaction between both flows of passengers slows down the boarding and alighting process, extending the time needed to complete the boarding and alighting process.

Another way in which boarding passengers obstruct the flow of alighting passengers is a result of passengers waiting to board leaning in. Passengers do this to try and spot a gap in the alighting flow of passengers (Harris et al., 2014; Heinz, 2003), this behaviour reduces the effective door width. This effect is illustrated in Figure 8, where the effective door width (shown in red) is much narrower compared to the designed door width (shown in green) as a result of how boarding passengers queue around the door. In practice, this means that even though the door is wide enough to allow for two lanes of alighting passengers, the effective door width only allows for a single lane of alighting passengers. Having more lanes means that more passengers can board or alight at the same time, speeding up the boarding and alighting process.

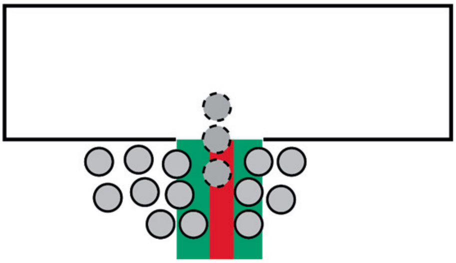

Figure 8: Designed (green) versus effective door width (red) when passengers crowd around the train door. Adapted from Harris et al. (2014)

2.5.2 The influence of the spread of boarding passengers

As stated by Heinz (2003), an easy mistake to make is to assume that passengers spread out evenly across all available doors. Although in theory this makes sense, in practice some doors are more used than others, a phenomenon known as concentrated boarding (Fox et al., 2017) and is illustrated in Figure 9. Concentrated boarding has an effect on the duration of dwell times since it leads to a minority of the available doors being utilized by a majority of the passengers, meaning that the boarding and alighting times at those doors will be extended. Studying the effect of concentrated boarding on dwell times Oliveira et al. (2019) found that dwell times can be extended by as much as 52 seconds when the degree of concentrated boarding is high. Looking back at Figure 6, it is possible to see that an uneven spread of boarding passengers will also effectively make the critical door more critical.

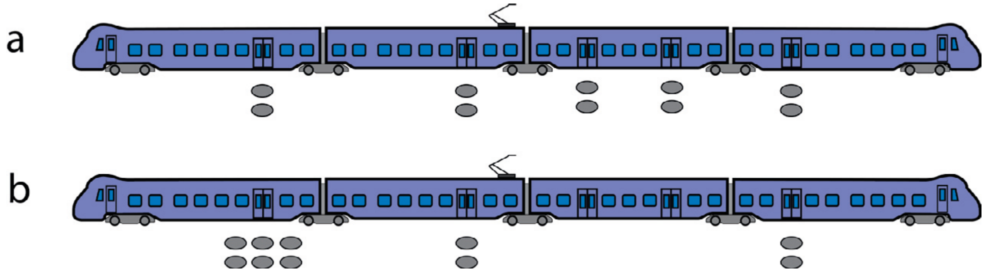

Figure 9: Example of an even spread of boarding passengers (a) and concentrated boarding (b).

2.5.3 Waiting positions of boarding passengers

The unevenness of the spread of boarding passengers is directly influenced by the waiting position of passengers on the platform. A study by Dell'Asin and Hool (2018) states that passengers are likely to use the door closest to them to board and Heinz (2003) states that passengers will only change their boarding door when the walking distance is less than 10 meters. This means that the waiting position of passengers on a platform is a good indicator of the spread of passengers between the available doors while boarding. The actual waiting position of passengers is the result of an interaction between passengers and the physical layout of both the platform and the station environment.

Observing the behaviour of passengers on station platforms, Nash et al. (2006) found that passengers tend to wait near objects on the platform and make use of the available roof coverage when waiting for a train to arrive. In addition, it has been found that passengers try to minimize their interaction with the flow of passengers on platforms, standing in such a way that they do not obstruct other passengers (Davidich et al., 2013). The physical layout of a platform in terms of the location of entrances and exits has also been shown to influence how passengers distribute themselves across platforms. Studies in several countries found that passengers are likely to wait near the entrance of a platform (Krstanoski, 2014; Lee et al., 2018; Peftitsi et al., 2020; van den Heuvel, 2016).

In contrast to this, some studies found that passengers are likely to base their waiting position on the location of the exit at their destination (Fang et al., 2019; Jusuf et al., 2017; Kim et al., 2014; Zheng, 2018). Rüger (2018) hypothesizes that this has to do with reducing the risk of being stuck in a queue while exiting a station.

Some studies have indicated that it is not just the exit location but also the surrounding area of a station that has an influence on where passengers position themselves on a platform. For example, when a certain area can only be reached through a specific exit (Bosina et al., 2017) or when a connecting mode of transport is more easily reached through a specific exit (Van Den Heuvel & Hoogenraad, 2014). A schematic example of this is shown in Figure 10, where the orange passenger wants to go to the city centre, and only one exit at the destination station leads to the city centre. In such a case, the orange passenger will stand in such a way that the door which is used for boarding lines up with this exit as much as possible.

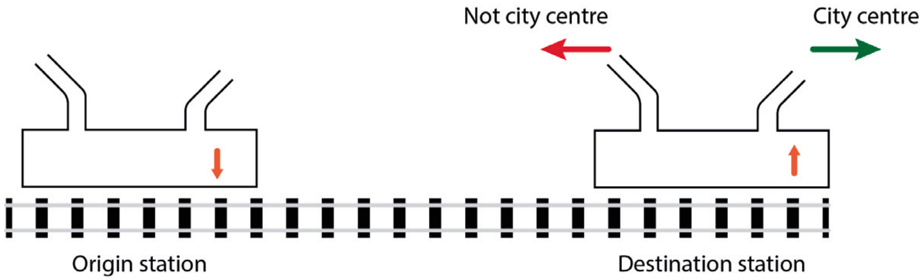

Figure 10: Example of waiting choices made by passengers using commuter trains.

This behaviour was found to be time-dependent, and passenger circulation slows down when the departure time of a train is near (Bosina et al., 2015; Fang et al., 2019; Fox et al., 2017; Wu et al., 2013). In such a case passengers opt to wait near the entrance location and board the nearest door. Rüger (2018) hypothesizes that this behaviour is due to passengers worrying about missing the train on the one hand, and passengers wanting to minimise their walking distance at their origin station on the other hand.

2.5.4 On-board congestion

The above-mentioned examples are related to what happens on the platform, but this is not the only way in which passengers can influence dwell times. As trains get more and more crowded the on-board congestion will start to play a role. As mentioned by Daamen et al. (2008) and Lee et al. (2018), increased vehicle occupancy or crowding will increase the difficulty for passengers to board and alight and extend the time needed for the boarding and alighting process. This has a strong link with on-board congestion. As described by Harris (2005), if a majority of passengers alight relatively large spaces are left inside the train which the boarding passengers can move into easily. If only a small number of passengers alight these spaces are not present making the boarding process more cumbersome.

Only a few studies have tried to uncover the impact of on-board congestion on boarding times. Studying the impact of passengers on dwell times in Tokyo, Palmqvist et al. (2020) found that increased on-board congestion resulted in longer boarding and alighting times. A recent study by Seriani et al. (2022) using a laboratory setting, also found that boarding and alighting times increase as the occupancy rate of a train carriage increases. Studying the impact of on-board density on the flow of passengers, Luangboriboon et al. (2020) found that there is a possible drop-off point for the passenger flow as the density in the vestibule increases, but do not find a definitive answer. Despite this, it is likely that the time needed for passengers to board and alight increases as on-board crowding increases.

2.5.5 Late arriving passengers

Dwell times can also be extended due to the behaviour of individuals near the end of the boarding procedure. An example of individual behaviour that has been identified in the literature is that of *door-holding*. Door-holding refers to the act of holding open a train door when the closing process has been initiated and is found to be most prevalent during rush hours (Hyun et al., 2016; Lindfeldt, 2017). Various reasons for door-holding have been identified in the literature.

Passengers trying to "squeeze" into a full train will extend the time a door is open, for example (Coxon & Bono, 2010). Other examples from the literature are situations where the train door is kept open to allow late-arriving passengers to board. The latter can also be due to staff members who hold the doors open to allow passengers arriving just before the departure time to still board (Coxon & Bono, 2010). The impact of door-holding is amplified when the closing sequence has to be reinitiated, which will extend dwell times even further (Harris et al., 2013).

2.6 Rolling stock design and passenger flows

The flow of passengers is not only affected by what happens in the platform train interface area but also by the design of the trains, also known as rolling stock, and platforms. The following section provides a brief overview of some of the findings made regarding the effect of the design of rolling stock and platform on dwell times.

2.6.1 Gaps and steps

When passengers alight or board a train they have to traverse a step between the platform and train, this can be either horizontally, vertically, or both horizontally and vertically depending on the design of both the platform and rolling stock. Several studies have indicated that decreasing the vertical and horizontal gap is beneficial for boarding and alighting speeds (Atkins, 2004; Daamen et al., 2008; Seriani & Fujiyama, 2019a) which is possibly due to people not having to slow down to accommodate any gaps (Heinz, 2003). Depending on the design of a train, a vertical hurdle in the form of steps can be present at the door as well which can slow down boarding and alighting times (Heinz, 2003). A possible explanation for this is that these steps increase the perceived difficulty of boarding a train making passengers slow down (Holloway et al., 2015). The effect of both the vertical and horizontal gaps is somewhat related to the direction of passengers, where the impact is larger for boarding passengers than for alighting passengers, (Holloway et al., 2015). This distinction is most clear when two steps are in place, in which case the boarding times are half a second longer per participant in comparison to the alighting times.

2.6.2 Door width

Although the step size has been indicated as an influential factor on the speed of boarding and alighting, Seriani and Fujiyama (2019a) found that the vertical gap is less relevant for the flow rate of passengers when the doors of the train are wide, allowing for a side-by-side flow. Similar remarks with regard to the benefit of wider doors are made by Wiggenraad (2001). It is unclear as to what the ideal train design should be with regard to the door width, however. Some results indicate a door width of 1.7 to 1.8 metres (Thoreau et al., 2016) to be the optimum, whereas other results indicate a smaller door width of 1.65 metres to be optimal (Fernández et al., 2015). This being said, Harris et al. (2014) found that the door width may not be as significant as commonly stated. This limited effect of the door width is potentially due to the positioning and behaviour of passengers, as previously described, which can be more influential on the boarding and alighting flows than the actual door width (Harris et al., 2014; Harris & Anderson, 2007).

2.6.3 On-board bottlenecks

It is not just the outside door which forms a potential bottleneck for boarding passengers, there are also on-board bottlenecks which slow down the flow of boarding passengers. A study in Switzerland showed that interior bottlenecks become normative for the boarding rates when the number of boarding passengers at a single door exceeds 10 people (Tuna, 2008). This is especially the case when the capacity of the interior doors is lower than that of the exterior doors (Heinz, 2003). Another on-board bottleneck is the width of the aisles between the seats. Somewhat counterintuitive, Sutton and Moncrieff (2015) found that narrower aisles resulted in lower boarding and alighting times due to fewer obstructions as people waited in the aisle and showed more cooperative behaviour. The authors nuance their findings by stating that only a limited number of observations were collected. Sutton and Moncrieff (2015) also studied the impact of perch seats in the vestibule and found that the presence of such seats negatively affected the boarding and alighting speeds. On-board crowding also forms a bottleneck, where the bunching of passengers in the train slows down the flow of passengers (Coxon et al., 2009). As with the gaps and steps, the effect of these bottlenecks depends on whether there the majority of passengers are boarding or alighting (Thoreau et al., 2016).

2.7 Collecting passenger flow data

In order to study the effect of passengers on dwell time it is important to have information about passenger flows. This is, however, not a straightforward task. Several methods exist to collect passenger flow information. In this context, passenger flows refer not only to the volume but also to other aspects such as waiting positions and the choice of carriages. The following overview is not an extensive list of all studies that made use of a specific method but rather serves as an overview to show different collection methods and some of their strengths and weaknesses.

2.7.1 Manual counts and video observations

A relatively straightforward way to collect passenger flow information is to make use of manual passenger counts where observers on the platform, or on board a train, count the number of passengers boarding and alighting. Although it is relatively simple to perform such manual counts it is labour-intensive and it is often not feasible to collect information on a door-by-door level or on a large selection of stops. In practice, this means that it is often only possible to collect a small sample of data.

An example of the use of manual counts is a study by Wiggenraad (2001) in which a total of 7 stations were observed, each for one day. The study made use of between 6 and 21 observers across these stations and a total of 130 trains dwelling were observed. This example shows how manually collecting data is indeed labour-intensive and the total sample size that can be collected is relatively small.

Some studies made use of camera observations to somewhat overcome the need for a large number of observers. A benefit of this over the use of manual counts is that it is possible to re-watch the videos. This will, in theory, help to improve the accuracy of the measurements. Camera observations were used by Oliveira et al. (2019) to study where passengers stand on the platform prior to boarding a train. In a similar type of study, Davidich et al. (2013) made use of video observations to track passengers on a platform in order to study their waiting behaviour. Video observations were also used by van den Heuvel (2016) to study the effect of changing the stopping location of trains at Schiphol Airport. In their study, the researchers made use of video observations in combination with visual recordings by the research team, this was done to ensure the accuracy of the measurements and identify events that might cause issues with the validity of their study. Although the process is less labour-intensive compared to manual counts, the data collection is still limited in its scope due to viewing angles, battery life, and the need to have observers present. Oliveira et al. (2019), for example, mentioned that cameras would turn off after an hour and that some doors were out of the field of vision of the cameras requiring manual counts. In addition to this, there can be privacy concerns related to recording individuals in a public setting. This can also mean that it is not always possible to make use of video observations as a way to collect data.

2.7.2 Automatic fare collection

The last few years have seen an increase in the implementation of new technologies in the public transport sector. Where in times of the past one had to buy an actual paper ticket, sometimes even from a person, most operators offer applications or smartcards instead these days. Smartcard data can be a rich source of information with regard to passenger volumes. Depending on the system, passengers have to either only tap-in upon entry or they are required to tap-in and also tap-out when leaving the system. The latter provides an incredibly rich data source since it allows for the construction of origin-and-destination matrixes in addition to providing a good estimate of the number of passengers on trains. Having access to both the origin and destination of a passenger allows for in-depth studies of travel behaviour, especially when automatic passenger count data is present as well.

An example of a study using smartcard data is the work done by Van Den Heuvel and Hoogenraad (2014) who made use of automatic fare collection at the exit of stations to study the flow of passengers through three stations in the Netherlands. In their study, the researchers assessed transfer times by comparing the arrival times of trains to tap-out times, allowing the researchers to study transfer times as a result of a new train schedule. The researchers also assessed desired exit locations based on where passengers tap-out. Although these examples are not directly related to passenger counts and dwell times, it does highlight how smartcard data is a rich source of data. An example closer related to dwell times is a study by Peftitsi et al. (2020) who used automatic fare collection from the metro system in Stockholm. In their study, they used this data, in combination with load-weight data to study the distribution of passengers on metro trains.

Although a rich source of data, automatic fare collection does have some practical drawbacks. A prerequisite for the use of automatic fare collection data, be it by means of smartcards or smartphones, is that passengers need to at least tap-in at their origin station. This is, however, not always the case. Where some stations in the Netherlands have been turned into a proverbial fortress, where one needs to tap-in and tap-out to even just cross the station, this is not the case in many other places in the world. For example, in Skåne, where the University of Lund is located, passengers make use of a smartphone app to buy tickets and there is no need to either tap-in or tap-out. In addition to this, if passengers are only required to tap-in at their origin station, as is the case in Stockholm for example, it is often not possible to make a distinction regarding the direction of travel. This means that if trains running in both directions halt at a station around the same time, it is not directly possible to identify which train is taken. Different data collection methods are needed in those cases, such as the use of systems that automatically count the number of passengers boarding and alighting a train.

2.7.3 Automatic passenger count data

Automatic passenger count data refers to passenger count data that is automatically collected on trains, hence the name. Several ways to count passengers exist. Examples of this are studies by Peftitsi et al. (2020) and Fang et al. (2019) who made use of load-weight data to study the preferred carriage of passengers in Stockholm and London respectively. Load-weight data provides an assumption about the number of passengers on board a train based on the weight of the train. To so do, the weight of the train upon departure minus the empty weight of the train can be divided by an average weight of passengers. When making use of load-weight data it is important to calibrate the collection method to the local context since passenger counts are not accurate if the average weight that is used is too high or too low. Peftitsi et al. (2020) state that the average weight is 78 kilograms per passenger including luggage, whereas Fang et al. (2019) use a value of 75 kilograms.

Although load-weight data has several important benefits over the use of manual counts or video observations, it still has some drawbacks. The first drawback is that the estimation of passengers based on weight is not always accurate since an assumption about the average weight has to be made[1]. Furthermore, due to the way in which load weight data is collected, often by making use of sensors in the suspension of a carriage, the retrieved information is limited to an aggregation on a carriage level. Only having access to passenger count data on a carriage level means that more detailed analyses such as the identification of the critical door or the spread of passengers between the doors are not possible. This is where another method to automatically count passenger volumes comes in, this being automatic passenger count systems which make use of sensors at each door.

Making use of sensors at each door to count passengers does not require an estimation of the average weight of passengers and makes it possible to collect information on a door-by-door level. That being said, this is not always the case that this kind of data is always available on a door-by-door level. A study by Palmqvist et al. (2020), using counts from commuter trains in Stockholm, was still limited to information on carriage level due to the way the data was collected and aggregated. A study by Buchmueller et al. (2008), using highly detailed automatic passenger count data from commuter trains in Switzerland, shows how having a highly detailed level of detail of automatic passenger count data on a door-by-door basis makes it possible to study dwell time at a greater level of detail. In their study, Buchmueller et al. (2008) not only had access to the number of passengers on a door-by-door level but also to the timestamps of both the first and last passenger through the door and timestamps for the opening and closing of each door. By having access to such detailed data it was possible to study the time it takes for each sub-process of dwell times, allowing the researchers to construct a figure as previously shown in Figure 6.

2.7.4 Using mobile phone data

Both automatic fare collection and automatic passenger count data are not always available, either the systems are not present or the data cannot be shared freely. This calls for different methods to collect information on the number of passengers on board a train. An example of a study that made use of a more exotic way to collect passenger counts is the study by Sørensen et al. (2018) who utilized cell phone data to measure the number of passengers on board a train. One reason to do so is that passenger count data is not easily made accessible by operators.

[1] I, for example, would count for 1.35 passengers in Stockholm at the time of writing.

To overcome this issue, the researchers collected information on the number of phones connected to selected base stations. The selected base stations were located in such a way that it was likely that the majority of the mobile phone traffic was generated by passengers. Their results show that using mobile phone data is a promising way to substitute other passenger counts. This being said, as with passenger count data, gaining access to mobile phone data can also be difficult which means that although the method proposed is promising, it might not always be feasible to collect passenger counts by making use of mobile phone data.

3 Scope of this thesis

3.1 Problem outline

Railways have the potential to play a big role in our shift towards a more sustainable way of travelling, both for medium and long distances. However, current levels of punctuality reduce the attractiveness of trains as a mode of transport and with this its competitiveness against other transport modes. The poor punctuality of railways in Sweden not only affects the experience of those who currently travel by rail but also hampers a modal shift away from private motorized transport, something which is necessary to help reduce greenhouse gas emissions from the transport sector.

As stated by Palmqvist et al. (2017) there is reason to believe punctuality issues are a result of errors in the timetable. Recent years have seen much effort being placed on introducing better timetable practices to ensure an on-time performance of trains, with advancements from operational research trickling down to practice (Caimi et al., 2017). Despite this, the process of dwell time scheduling has remained more or less stagnant and is mostly based on general assumptions in practice (Christoforou et al., 2020; Palmqvist, 2019), whilst an important part of the timetable planning process for commuter trains is to accurately estimate and schedule dwell times (Buchmueller et al., 2008).

An analysis of the scheduled dwell times in Skåne, southern Sweden, between 2012 and 2020 revealed that the scheduled dwell times have not changed for most stations during this period. At the same time, passenger volumes have increased which means that it is likely that more time is needed for passengers to board and alight a train, and the infrastructure is used more intensively which increases the risk of small disturbances to cascade into larger delays, making the system less robust and stable. This phenomenon of dwell time scheduling practices not being changed is not unique to Sweden as Nash et al. (2006) identified a similar problem in Switzerland where passenger volumes had increased but dwell times remained stagnant, resulting in a situation where the scheduled dwell times were impossible to maintain.

Although it is easy to critique the static approach to dwell time scheduling, it must be said that deciding how much dwell time should be scheduled is a non-trivial task. The dwelling process involves several external factors which makes it subject to higher variability (Cornet et al., 2019) and makes it difficult to exactly know the necessary time upfront. The task of dwell time scheduling is made even harder due to a lack of understanding of the underlying causes of dwell time delays since dwell time delays are not well recorded (Harris et al., 2013; Pritchard et al., 2021; Volovski et al., 2021). Having a more in-depth understanding of how different factors influence dwell times is a necessary step towards more accurately scheduling dwell times, allowing for well-informed decision processes which will help to work towards reducing delays incurred by commuter trains along their journey.

3.2 Research gap

One reason why dwell times for commuter trains are not well understood is that most studies in the past were limited in either the size or the scope of their data. Studies on the behaviour of passengers during the boarding and alighting process often take place in laboratory or mock-up settings, see for example the work done by Daamen et al. (2008), Harris (2005), and Seriani and Fujiyama (2019a). Although valuable insights can be gained regarding the underlying processes and behaviour of passengers, results from laboratory studies do not always translate well into real-world scenarios (Dobbins et al., 1988). As Luangboriboon et al. (2020) mention, even though laboratory experiments allow for a more in-depth study into the influence of a single variable, they may not completely reflect a real situation. The authors state that people may behave differently due to the environment in which the study takes place or the instructions that are provided. To overcome these issues, studies can take place outside of laboratory settings. Although studies in real-world settings make it possible to study the dwell time process in a more natural environment, these studies often rely on a limited number of observations on a few stations, see for example the studies by Oliveira et al. (2019) and Wiggenraad (2001). The limited sample size and geographical spread means that some effects might not be clearly identified and the findings that are made can potentially be limited to only a small subset of all stations.

Using automatic data sources such as automatic passenger count data can help overcome the issue of having a limited sample since it enables the use of data from multiple stations, multiple railway services, and over a long period of time. However, as mentioned by Coulaud et al. (2023), only a few studies had access to both highly detailed passenger count and railway operational data, and most of these studies are limited to data on specific parts or lines of the network such as the studies by Christoforou et al. (2020), Cornet et al. (2019) and Coulaud et al. (2023), rather than the whole network.

The study performed by Palmqvist et al. (2020) is one of the few examples where dwell times were studied with a network-wide scope, using data from the railway network in Stockholm. Although the study also included data from Japan, this was again limited to only a small part of the network. As a result of this, there is a lack of studies into the relationship between passengers and dwell time on a network-wide level and over an extended period of time. Filling this gap can help to increase the understanding of the effect of passengers on dwell times, increase the understanding of dwell time delays, and help inform planners when scheduling dwell times in the future.

3.3 Research aim and questions

The overarching research aim of this thesis is to develop knowledge of how dwell time delays arise in order to identify and describe potential ways in which dwell time delays can be reduced. A specific focus is placed on the role of passengers in the dwelling process in relation to dwell time delay. In addition to this, a secondary aim is to identify how dwell times can be studied on a network-wide level. The following research questions form the basis of this thesis, and when answered, will help towards achieving the research aims:

Research question 1
What are the causes of dwell time delays for commuter trains?

Research question 2
How do boarding and alighting passengers influence the duration of dwell times for commuter trains?

Research question 3
How can dwell time delays be studied in a network with automatic passenger count data?

Research question 4
Which types of measures can be taken to reduce dwell time delays for commuter trains?

3.3.1 *Motivation and relationship of the research questions*

The four research questions that make up the backbone of the work presented in this thesis do not stand on their own and the way in which the research questions are connected is graphically shown in Figure 11. As the figure shows, research question 1 feeds through to both research questions 2 and 4, with the answer to research question 1 partially answering both questions. A similar relationship is present between research questions 2, 3, and 4 with research questions 2 and 3 partially answering research question 4.

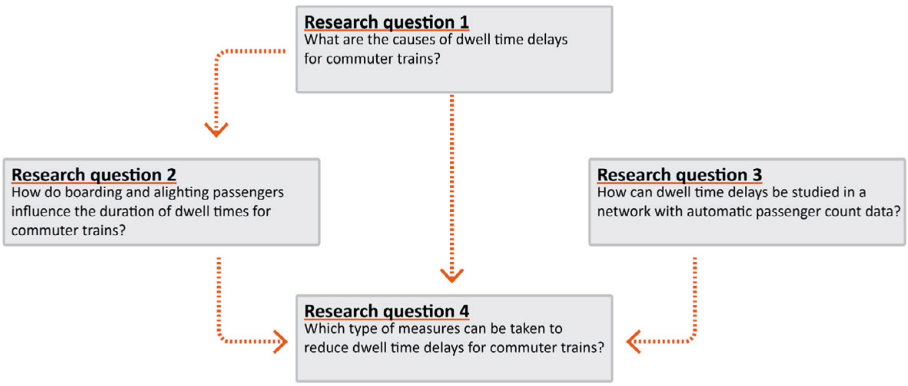

Figure 11: Relationship between the four research questions.

The development of research questions 1, 3, and 4 was guided by the overarching aim of this thesis. Research questions 1 and 4 arose from the need to define measures to reduce the risk of dwell time delays for which an understanding of the causes was necessary first. Research question 3 arose from the need to be able to better identify dwell time delays since current delay indexes obscure the presence of smaller delays. In contrast to this, research question 2 was defined based on the answer to research question 1. During the process of identifying the likely causes for dwell time delays it became clear that the influence of passengers plays a substantial role and more in-depth knowledge on this relationship is needed, thus forming the basis for research question 2.

4 Papers included in this thesis

4.1 List of included papers

Paper I

Kuipers, R. A., Palmqvist, C.-W., Olsson, N. O. E., & Winslott Hiselius, L. (2021). The passenger's influence on dwell times at station platforms: A literature review. *Transport Reviews*, 1–21. DOI: 10.1080/01441647.2021.1887960

Paper II

Kuipers, R. A., & Palmqvist, C.-W. (2022). Passenger Volumes and Dwell Times for Commuter Trains: A Case Study Using Automatic Passenger Count Data in Stockholm. *Applied Sciences*, 12(12), 5983. DOI: 10.3390/app12125983

Paper III

Kuipers, R. A., & Palmqvist, C.-W. (2022). The spread of passengers on platforms and dwell times for commuter trains: A case study using automatic passenger count data. *11th Triennial Symposium on Transportation Analysis conference* (TRISTAN XI), Mauritius.

Paper IV

Kuipers, R. A., & Palmqvist, C.-W. (2023). Impact of a lower demand during the COVID-19 pandemic on the frequency of dwell time delays. *Transportation Research Interdisciplinary Perspectives*, 21, 100911. DOI: 10.1016/j.trip.2023.100911

Paper V

Kuipers, R. A., Tortainchai, C., Tony, N.C., & Fujiyama T. Understanding dwell times using automatic passenger count data: A quantile regression approach (*article under review – Transportation Research Interdisciplinary Perspectives – first round of reviews)*

Paper VI

Kuipers, R. A. (2024). Understanding dwell times using automatic passenger count data: A quantile regression approach. *Journal of Rail Transport Planning & Management*, 29, 100431. DOI: 10.1016/j.jrtpm.2024.100431

4.2 Declaration of contributions

In *Paper I* Kuipers was the main author and responsible for the literature search and the synthesis of the literature along with the writing of the manuscript, this was done under the supervision of Palmqvist, Olsson, and Hiselius who assisted with the methodology and reviewing of the manuscript.

Kuipers was the main author for *Paper II* and the work was done in close collaboration with Palmqvist who assisted with the methodology, initial writing of the manuscript, and the later reviewing of the manuscript. The formal analysis was performed by Kuipers.

The idea of *Paper III* was initiated by and carried out by Kuipers under the supervision of Palmqvist. The same is the case for *Paper IV* for which Kuipers was the main author and initiator of the idea and research approach, and supervision was performed by Palmqvist.

The idea of *Paper V* was initiated by Fujiyama and the research approach was discussed in close collaboration with Tortainchai and Tony. Kuipers was the first author and along with Tortainchai responsible for the formal analysis, visualization and writing of the manuscript.

Kuipers was the sole author of *Paper VI*

4.3 Relationship between included papers

The relationship between the six included research papers is illustrated in Figure 12. Most papers, with the exception of Paper V, stem from the findings in Paper I which serves as an umbrella for the research that was conducted. The findings from Paper I revealed which passenger flow characteristics are relevant to focus on when studying the impact of passengers on the duration of dwell times. Based on this it was decided to focus on the volume and spread of passengers in Paper II and Paper III respectively. The findings from, and method used in, Paper II shaped the research conducted in Paper IV where the impact of the volume of passengers was further investigated. In contrast to the other papers where the focus was placed on the effect of passengers on dwell times, the study conducted in Paper V focussed more on understanding dwell times from an operational context. The knowledge gained along the way was combined into the study that makes up Paper VI, where variable choices were informed based on findings from the results of all papers, with the exception of Paper V.

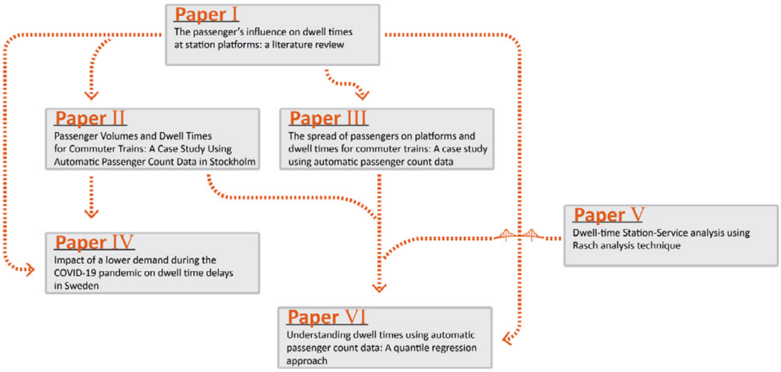

Figure 12: Relationship between the papers included in this thesis.

4.4 Relation between the research questions and papers

The research questions described in *Section 3.3* are answered by combining findings from the different papers included in this thesis. An overview of the relationship between the research questions and the included papers is shown in Table 1. All research questions are answered by combining the findings of at least two of the included papers. There is a heavy and deliberate emphasis on answering research question 2, given the important role that boarding and alighting passengers play in the dwelling process. A strong emphasis is also placed on research question 4, where the findings from several papers are combined to describe possible measures which can help to reduce dwell time delays

Table 1: Relation between the research questions and included papers.

Research question		Paper
1	What are the causes of dwell time delays for commuter trains?	I, VI
2	How do boarding and alighting passengers influence the duration of dwell times for commuter trains?	I, II, III, IV, VI
3	How can dwell time delays be studied in a network with automatic passenger count data?	II, IV, VI
4	Which type of measures can be taken to reduce dwell time delays for commuter trains?	I, II, III, IV, V

5 Methods

This chapter describes the methods that have been applied throughout the work done that makes up this thesis, and an overview of the different methods used in each paper is provided in Table 2. There is a heavy emphasis on the use of quantitative methods, with only the literature review being a qualitative method. The use of mostly quantitative methods was guided by the available data which lends itself well to the use of different quantitative approaches to study various relationships between passengers, operational conditions, and dwell times.

Table 2: Methods used for the papers included in this thesis.

Method	Paper
Literature review	I
Visual graphical analysis	II, IV
Statistical analysis	II, IV
Item response model	V
Regression analysis	III, VI

Although there exists a large body of literature that focuses on how passengers influence dwell times, much of these findings are dispersed across a multitude of different studies. A systematic literature review was, therefore, conducted in Paper I to collect and organize the state-of-the-art on how passengers influence dwell times. A visual graphical analysis was performed in Paper II and Paper IV to discover trends in the data. The graphical visual analysis was chosen here since it is a straightforward explorative data analysis technique where trends are observed by visually representing the data. The outcome of such an analysis is then used to guide further analyses. Following these visual graphical analyses, several different statistical analyses were performed in Paper II and Paper IV to test a series of hypotheses. The study in Paper V made use of an item response model, namely a Rasch analysis, to provide a framework to understand the combined effect of stations and railway services on dwell times. Regression analyses were chosen for Paper III and Paper VI to understand the impact of multiple variables on either the probability of a dwell time delay, presented in Paper III, or on the duration of dwell times as is the case for Paper VI.

5.1 Systematic literature review

Paper I

A systematic literature review was chosen as the method for Paper I since it allows to comprehensively record and assess the state-of-the-art on how passengers influence dwell times. The eight-step framework, schematically shown in Figure 13, as described by Denyer and Pilbeam (2013) was used to perform the systematic literature review. In addition to the eight steps proposed by Denyer and Pilbeam (2013), an additional step of backwards snowballing, as described by Jalali and Wohlin (2012), was included in the literature search. This additional step was necessary to overcome the issue of missing papers during the initial search as a result of an inconsistent use of terminology in the published literature. Terms such as *dwell time* and *passenger service time* are used by different authors, for example.

An important step in the literature review is the evaluation of the quality of the literature. The *critical appraisal tool for evidence-based librarianship* (Glynn, 2006) was used to perform this evaluation step in which a score is assigned to each article. These scores are based on predefined criteria that focus on the population being studied, the way in which data is collected, the design of the study, and how the results of the study are reported. Articles were then classified as having either a *good* or *questionable* quality in all four categories as well as an overall score. When a paper was determined to be of questionable quality it was re-examined and a decision was made to include or exclude the paper based on this critical re-examination.

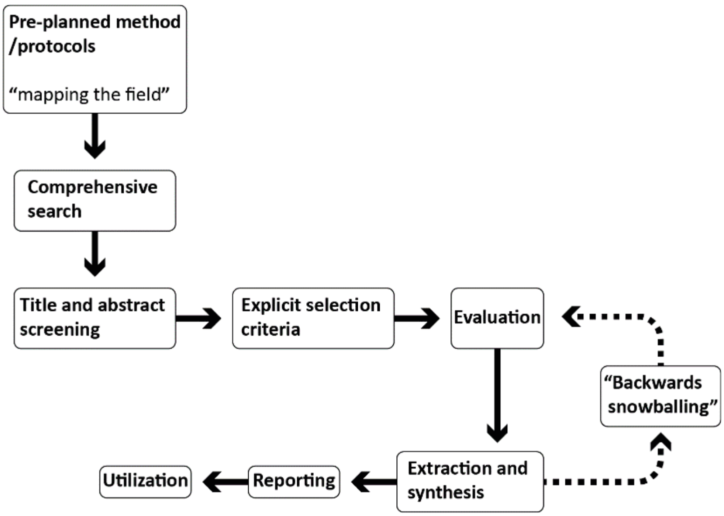

Figure 13: Schematic overview of the systematic literature search.

The aim of the literature search was to present the way in which passengers affect dwell times, both before and during the boarding and alighting process, and the underlying causes hereof. Given this aim, keywords related to *passenger distribution* and *boarding and alighting behaviour* were used to systematically search the databases. These keywords were combined with additional phrases such as *station* and *platform* to narrow down the results. This choice and phrasing of these keywords was based on the preliminary findings whilst mapping the field in step one of the framework that was used.

The literature search took place during March 2020, and the period of the publications was limited to January 2000 and March 2020. Papers were excluded when the results were based on a model without elaborating on passenger characteristics, used fixed boarding and alighting times, focused on busses, or focused on dwell times solely from a timetable perspective. In addition to these criteria, the location was added as an additional criterion for studies that focussed on passenger flow characteristics. This criterion was added since the literature review focussed on passenger behaviour within a European context and walking characteristics and the space maintained between pedestrians are not similar between all cultures. North American and European passenger flow characteristics are found to be relatively similar (Daamen, 2004), meaning that only studies originating from either of these two continents were included.

5.2 Visual graphical analysis

Paper II & Paper IV

The expected trends and distributions of the variables were not clear from the start for most papers, especially since the studies make use of a large amount of data. Paper II and Paper IV, therefore, included an exploratory data analysis by means of plotting the data and visually assessing trends, a technique known as *visual graphical analysis*. Visual graphical analysis is an explorative technique that can aid in the analysis and interpretation of the data by visually representing the information (Wainer & Thissen, 1981; Keim et al., 2006; Brown et al., 2007; Landmark et al., 2017). Visual graphical analyses have been applied within delay research in the past. For example, a common way to visualize the on-time performance of train lines is the use of delay profiles (van Oort & van Nes, 2009; van Oort et al., 2015) where each trip is plotted on a line graph to show the development of delays on a specific line. Another example of visual graphical analyses being used in railway studies is the use of heat maps to identify delay hot spots and to visualize characteristics of delays such as passenger loads and spatiotemporal factors (Christoforou et al., 2020; Huang et al., 2019).

The visual graphical analysis performed in Paper II consisted of plotting the frequency distribution of different sizes of dwell time delays for different groups of passenger volumes. A similar approach was used in Paper IV to compare the frequency of dwell time delays between the period before and during the COVID-19 pandemic. In addition to the frequency of delays, the median and interquartile range for the different passenger flow characteristics such as the volume of passengers were plotted as well in Paper IV. This was done to visually assess any potential differences in the passenger flow characteristics between the period before and during COVID-19.

5.3 Statistical analyses

Paper II & Paper IV

The following sections describe the different statistical analyses that have been performed, stating the paper in which the method was applied and the rationale for using the specified approach. There is an emphasis on the use of robust statistical analyses resulting in the use of mostly non-parametric statistical methods as opposed to the more generally applied parametric methods. The main reason for this is that both the data on dwell times as well as the data on passenger flows violate the assumption of normality present for parametric methods.

5.3.1 Chi-Square goodness of fit test

A Chi-Square goodness of fit test was applied in Paper II and Paper IV to study changes in the frequency of dwell time delays. The Chi-Square goodness of fit test is a non-parametric test that can be used to analyse the difference between groups when the dependent variable is measured at a nominal level (McHugh, 2013), such as counts or frequencies. In the case of Paper II, the interest was the difference in the distribution of delays for different groups of passenger volumes. The Chi-Square goodness of fit test was applied in Paper IV to assess the difference in the frequency distribution of delays between the period before and during the COVID-19 pandemic.

A transformation of the data was necessary in order to perform a Chi-Square goodness of fit test. This is necessary since the dependent variable needs to be measured on a nominal level, such as frequency counts, and dwell time data is collected on a continuous scale. The continuous variables of observed dwell time delays were transformed into frequency counts based on the size of the delay. It is important that the size of the buckets reflects a real-world situation, and that it allows to study changes in the frequency dwell time delays with sufficient detail. The choice was made to use group delays in buckets of *fifteen* seconds.

The scheduling regime in place makes use of steps of thirty seconds when planning process times. Having very small buckets, of one second, for example, would thus not reflect a real-world application. Having bucket sizes of thirty seconds was found to be too coarse to properly study dwell time delays, however, resulting in too few observations per group. One of the assumptions of the Chi-Square goodness of fit states that eighty per cent of the expected frequencies should at least have five occurrences (McHugh, 2013), which was not reached with bucket sizes of thirty seconds. Fifteen seconds was, therefore, chosen to be the happy middle.

5.3.2 Kruskall-Wallis and Wilcoxon signed-rank tests

Paper IV aimed to understand differences in both passenger flow characteristics and the length of dwell times between different groups, in this case, the period before and during the COVID-19 pandemic. The classical way to approach such a task would be to make use of a Student's t-test, ANOVA, or paired t-tests. Given the need for robustness in the statistical analyses, as described before, non-parametric equivalents of these tests were used. One of such tests is the Kruskall-Wallis test which is the non-parametric equivalent of the ANOVA-F test and can be used to compare independent samples (Ostertagová et al., 2014). Another test that was used is the Wilcoxon signed-rank test, which provides a non-parametric alternative to the paired t-test (King & Eckersley, 2019).

5.3.3 Effect sizes for statistically significant findings

The different tests above describe the approach taken for significance testing. It is, however, not enough to only report significant results and effect sizes should be reported as well (Brydges, 2019). Where significance testing is used to determine whether an effect is due to chance or not, the effect size is a measure of the practical relevance of a finding (Lantz, 2013). The type of effect size that is calculated is determined by the statistical test that is performed. The effect size associated with a Chi-Square goodness of fit test is *Cohen's W* for example, whereas effect sizes for a Kruskall-Wallis or Wilcoxon signed-rank test are based on *eta squared* and *Pearson's r* respectively.

To interpret the effect sizes the suggestions provided by Cohen (1988) for the labels of the effect sizes were used throughout the studies that make up this thesis. In general terms, these labels are small, medium, and large with the values associated with these labels being dependent on the effect size measure that is used (Cohen, 1988; Lakens, 2013). An overview of the effect sizes used and their respective labels is presented in Table 3.

Table 3: Effect size measured used.

Effect size measure	Small	Medium	Large
Cohen's *w*	0.01	0.3	0.5
Eta squared	0.01 < 0.06	0.06 < 0.14	>= 0.14
Pearson's *r*	0. 1 < 0.3	0.3 < 0.5	>= 0.5

5.4 Item response model

Paper V

The study presented in Paper V makes use of a so-called *item response model*, specifically the Rasch analysis technique. Initially developed by George Rasch (1980), the Rasch analysis technique is a psychometric technique which has been widely used in medical and health sciences and is designed to be used for questionnaire surveys in which participants are asked about their perceived difficulty to perform a specified task. The outcome of a Rasch analysis can be used to calculate intervals that represent the ability of a participant and the overall perceived difficulty to perform the specified task. Using both these indexes it is then possible to determine the likelihood of a given person successfully performing a given task. This is done by comparing the difference between the ability of a person and the difficulty of a task. This difference between ability and difficulty is central to the Rasch analysis technique. When the ability of a person is large compared to the difficulty of a given task, the probability of a successful response is larger and vice versa. A recent example of the use of the Rasch analysis technique within the domain of transportation is the study by Cheng and Chen (2015), who applied this technique to assess the accessibility of two cities in Taiwan.

Although the Rasch analysis technique is commonly used for questionnaire data it was used in Paper V to study the ability of a given train service (i.e. the participant) to have a punctual dwell time, which was compared to the difficulty for all train services on the line have a punctual dwell time at a specific station (i.e. the task). Doing so allowed to study the relative dwell time performance of a given train at a given station, in a single dimension. Taking both service and station performance into account in a single dimension makes it possible to not only identify problematic train services or stations separately but also jointly examine stations and services. This means that it is possible to identify which services perform poorly at a given station.

A challenge to use the Rasch analysis within the context of dwell times is that dwell times are measured on a continuous scale. This is problematic since the Rasch analysis technique requires polytomous data such as responses collected on a Likert-like scale. The continuous dwell time data does not fit this Likert-like scale directly, but it can be converted to fit. To make this conversion, the observed dwell time deviations were categorized according to predefined buckets. The size of these buckets is context-specific and it is important to keep operational constraints in mind as well as the observed data when defining the bucket size. The regime for labelling dwell time deviations used in Paper V is shown in Table 4. As can be seen, the continuous dwell time is converted into a five-point scale based on the size of the dwell time deviation, and the labels then be used in the Rasch analysis technique.

Table 4: Labelling regime to transform continuous dwell time data to polytomous data for use in the Rasch analysis technique

Dwell time deviation	*Label*
<= 0 seconds	5
> 1 and <= 30 seconds	4
>31 and <= 60 seconds	3
> 61 and <= 90 seconds	2
> 91 and <= 120 seconds	1
>120 seconds	0

5.5 Regression analyses

Paper III & Paper VI

In its simplest form, regression analyses allow to study the relationships between a dependent and one or several independent variables. Regression analyses are valuable since they allow the indication of whether there are statistically significant relationships between dependent and independent variables or not, indicate the strength of these relationships, and allow the user to make predictions (Sarstedt & Mooi, 2014). Two types of regression analyses were performed, this being logistic regression and quantile regression in Paper III and Paper VI respectively.

5.5.1 *Logistic regression*

A logistic regression was used in Paper III to study the influence of the spread of boarding passengers on the probability of a train incurring a dwell time delay. Logistic regression is similar to multiple linear regression with the difference being that the response variable is binomial, and the result relates to the conditional probability that an outcome occurs based on a set of explanatory variables (Sommet & Morselli, 2017; Sperandei, 2014). The basic model for the log odds takes the form of Equation 3 (Sperandei, 2014), and the probabilities of an outcome occurring based on these log-odds can be determined by making use of Equation 4 (Sommet & Morselli, 2017).

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0+\beta_1 X_1+\beta_2 X_2+.....\beta_m X_m \qquad (3)$$

$$\text{Probability} = \frac{\exp\,(\beta_0+\beta_1 X_1+\beta_2 X_2+.....\beta_m X_m}{1+\exp\,(\beta_0+\beta_1 X_1+\beta_2 X_2+.....\beta_m X_m} \qquad (4)$$

5.5.2 *Quantile regression*

The study performed in Paper VI made use of a quantile regression. Quantile regression is an extension of ordinary least squares regression with the major difference being that quantile regression allows the modelling of the entire conditional distribution rather than just the conditional mean (Rodriguez & Yonggang, 2017). This means that it is also possible to study the effects of the independent variables on the dependent variable in different parts of the conditional distribution (Karlsson, 2006) such as the tails of the distribution (Belaïd et al., 2020).

Quantile regression also allows to capture the variance in the relationship between the dependent and independent variables at different points in the conditional distribution as well as any asymmetry in the distribution (Waldmann, 2018). The basic model for the quantile regression is shown in Equation 5 (Rodriguez & Yonggang, 2017), and each quantile level, indicated by τ, yields a distinct set of regression coefficients.

$$Q_\tau(y_i) = \beta_0(\tau) + \beta_1(\tau)x_{i1} + \beta_2(\tau)x_{i2} + \cdots + \beta_p(\tau)x_{ip} \qquad (5)$$

Quantile regression was chosen as the method of choice for Paper VI for several reasons. First, dwell time data violates some of the assumptions of ordinary least square regression, most notably the assumption of normality. This is to be expected since dwell times are likely to be longer than shorter making the distribution inherently skewed. This non-normality is handled in quantile regression as it does not assume a parametric distribution (Hao & Naiman, 2007; Kourtit et al., 2022). Quantile regression is also better adapted to handle outliers (Hao & Naiman, 2007) as well as heteroscedasticity (Waldmann, 2018) compared to ordinary least square regression. The common approach when dealing with ordinary least squares regression is to remove outliers in the data due to their negative impact on the model fit. Although extreme values can be considered to be outliers, this does not mean that these values do not hold relevant information when studying dwell times. It is, for example, possible to have large dwell times with few passengers and vice versa. Simply removing such occasions would lead to the occlusion of important information.

In addition to quantile regression being better adapted to dealing with dwell time data compared to parametric approaches, the method was also chosen since it allows to study the relationship between dependent and independent variables outside of the mean (Kourtit et al., 2022). This is of interest when studying dwell times since the distribution holds more information than just the mean value. Quantile regression models provide the flexibility to identify how relationships between the dependent and independent variables change over different parts of the distribution (Cook & Manning, 2013; Staffa et al., 2019), the interpretation of the coefficients becomes nontrivial when using other non-parametric regression models. The interpretability of the influence of different variables on dwell time is important given objective of the Paper VI, and it was, therefore, chosen to make use of quantile regression.

6 Data used

Several different datasets were used, with a strong emphasis on operational data from the railway system. The data type and origin of the data are shown in Table 5. There is a mix of data from the national railway system, provided by the Swedish Transport Administrator, and the local operators of two regions in Sweden. This being Storstockholms Lokaltrafik who are responsible for the commuter trains in the Stockholm region, and Skånetrafiken who are the body responsible for public transport in the region of Skåne in southern Sweden. The following sections explain what the data consists of, how it has been used, and what some of the differences are between the data received from Storstockholms Lokaltrafik and Skånetrafiken.

Table 5: Overview of data used in this thesis.

Data type	*Origin of data*
Literature data	Scopus, Lund university library, experts
Train movement data (LUPP)	Swedish Transport Administrator
Timetable data (Trainplan)	Swedish Transport Administrator
Automatic passenger count data on a train level	Sorstockholms Lokaltrafik
Automatic passenger count data on a door-by-door level	Skånetrafiken

6.1 Literature data

The first data set used consists of literature data, collected by systematically searching the registries of Scopus and Science directly through the Lund University portal. The literature data was collected as part of the work in Paper I, which consisted of a systematic literature review. The body of literature was kept up to date during the years spent working on this thesis by making use of “Really Simple Syndication”, or RSS feeds for journals with relevant topics. This ensured that, to a large extent, the literature could stay up-to-date.

6.2 Train movement data

The train movement data that was used originates from track-side signal data, which can be accessed through a data layer called “LUPP” and is provided by the Swedish Transport Administrator. As explained in Palmqvist (2019) track-side signal data is coded by making use of track circuits to detect the presence of trains and timestamps are recorded when a train enters or exits the circuit. With regards to stations, these track-side detectors are often located at the edges of the station area, rather than in the middle of the platform, this principle is illustrated in Figure 14.

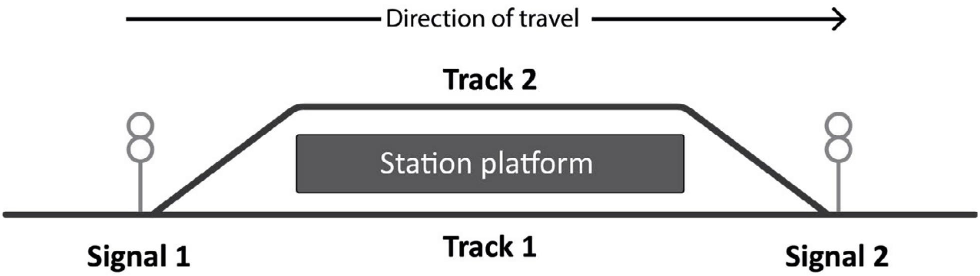

Figure 14: Schematic description of the location of data collection using track-side signals.

A train dwelling at a station is registered at *signal 1* upon entry. Once the dwell process is completed the train will depart and is registered at *signal 2*. The dwell time is then based on the difference in the time registered at signal 1 compared to the time at signal 2. The actual dwell time is thus not collected at the platform where the train halts but at the edges of the station. To somewhat overcome this issue, an automatic adjustment is made to account for the time it takes for the train to go from signal 1 to the platform, and from the platform to signal 2. These adjustments are usually set to be in the order of 10 to 20 seconds (Palmqvist, 2019).

Using these adjustments introduces a level of imprecision in the data. This is why the data which is stored in the LUPP database is truncated on a minute level to, somewhat, cover these imprecisions introduced in the data. This truncation does mean that LUPP data is not ideal when studying dwell times. The data does, however, still hold relevant information regarding train movements and the timetable. This was especially useful for the study making use of automatic passenger count data originating from Stockholm, Paper II, where information on the scheduled departure and arrival times of trains was not present in the automatic passenger count data.

6.3 Timetable data

Although the train movement data holds information on the scheduled times, these are not always the most accurate times since ad hoc changes can be made to the timetable. To make sure the correct timestamps for the scheduled arrival and departure times were used during the analyses, a data source called "Trainplan" containing up-to-date timetable data was utilized. Trainplan contains information on the actual scheduled arrival and departure times of trains, including any ad hoc changes. The inclusion of timetable data was especially important during the process of correcting for early arriving trains, which is described in *Section 7.2*. Correcting for early arriving trains is important since not correcting for longer dwell times due to an early arrival will lead to an overestimation of dwell time delays.

6.4 Automatic passenger count data

Automatic passenger count data was used in order to study the effect of passengers on dwell times. As the name suggests, this data holds information on passenger counts, more specifically the number of boarding and alighting passengers. In addition to the number of boarding and alighting passengers, the automatic passenger count data also provides the actual dwell time on a magnitude of seconds. This dwell time is recorded using an on-board system.

Two datasets holding automatic passenger count data were used. The first came from the commuter trains in Stockholm, operated by *Stockholms Lokaltrafik*, for the period between 2013 and 2016. The second dataset consists of automatic passenger count data collected on board the commuter trains operated in the region of Skåne, where the trains are operated by *Skånetrafiken*. This dataset spans between 2016 and 2020.

The commuter train network in Stockholm, shown in Figure 15a, serves 54 stations and consists of six lines with a combined length of 241 kilometres (Storstockholms Lokaltrafik, 2020). The commuter train network in Skåne, shown in Figure 15b, serves a much larger area with a total of 77 stations in the network. The network consists of nine lines with a combined length of 511 kilometres, with the shortest line being 15 kilometres and the longest being 107 kilometres.

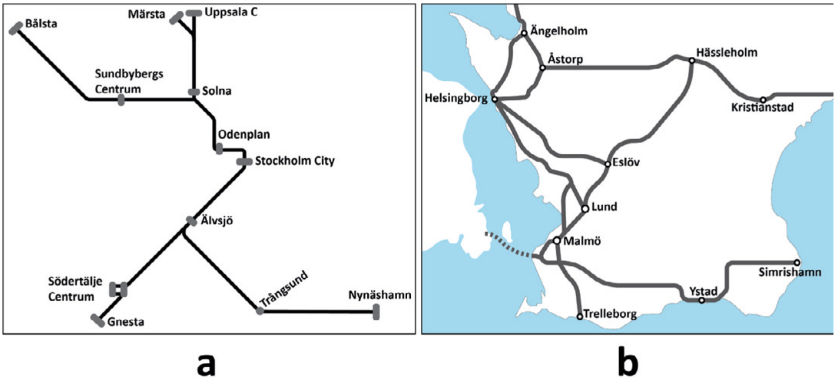

Figure 15: Schematic representation of the commuter train network in Stockholm (a) and Skåne (b) with some stations highlighted.

6.4.1 High-precision dwell time data

The automatic passenger count data provides the actual dwell time of a train in a magnitude of seconds, which is crucial when studying dwell times and dwell time delays. To illustrate why it is important to know the actual dwell time on a magnitude of seconds the percentage of stops with a given delay size is shown in Table 6, using both data from the track-side signals, i.e. the LUPP data, and dwell times from trains in Skåne.

Table 6: Percentage of stops with delays from both the signalling data (LUPP) and automatic passenger count data (APC) from commuter trains in the region of Skåne.

Year /Delay size	No delay		1 min.		2 min.		3 min.		4 min.		5 min.	
	LUPP	APC	LUPP	APC	LUPP	APC	LUPP	APC	LUPP	APC	LUPP	APC
2017	67%	24%	29%	65%	3%	9%	1%	1%	0%	1%	0%	0%
2018	68%	10%	27%	77%	3%	11%	1%	2%	0%	1%	0%	0%
2019	69%	15%	26%	71%	3%	12%	1%	2%	0%	1%	0%	0%

When making use of the LUPP data it seems that only a few dwell time delays occur, with about two-thirds of stops not having a delay and a third of the stops having a delay of up to one minute. However, when using data with dwell times measured on a scale of seconds it becomes clear that dwell time performance is not as good. The dwell time data collected on board the trains shows that less than a third of stops have no delay and the majority of stops have a delay of up to one minute.

6.4.2 The same but different

As mentioned, two different automatic passenger count datasets were used, one originating from the commuter trains in Stockholm and the other from the commuter trains in Skåne. Although both of these datasets consist of automatic passenger count data four important differences can be identified. First, is the way in which the data is collected. The trains that are operated in Skåne and Stockholm both use sensors to collect passenger counts, but their placement is different as schematically shown in Figure 16. The automatic passenger count system used in Stockholm, shown on the right, makes use of photocells which are located at floor level on the train doors. In contrast to this, the system used on the commuter trains in Skåne, shown on the left, makes use of ceiling-mounted infrared sensors. In the latter case, there is a minimum detection height of around one meter whereas the floor-level sensors do not have a minimum detection height.

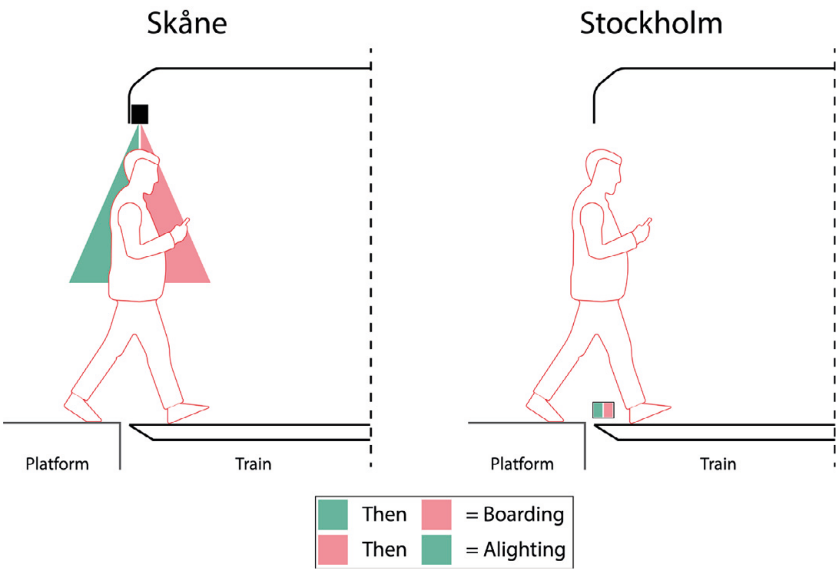

Figure 16: Difference between registration of boarding and alighting passengers for Skåne and Stockholm.

The different ways in which the data is collected come with their own drawbacks concerning the precision of data collection. In the case of the data originating from Stockholm, it is possible that the system reports more passengers than the actual number due to other objects, such as a suitcase or dogs, being recorded as well. In the case of the system used by Skånetrafiken, it is possible that the system underreports the number of passengers boarding and alighting since children could be too small to be recorded. If, for example, a class of four-year-olds takes the train for a school trip, it is possible that only the teacher will be counted.

The second difference between both datasets is the way in which the data is made available. The passenger count data collected in Stockholm is aggregated on a carriage level. This means that the passenger count data provides information on the total number of boarding and alighting passengers for a carriage at a given stop. The data collected on board the commuter trains in Skåne, on the other hand, is made available on a door-by-door level. This means that the data not only provides information on the total volume of passengers on a carriage level but also shows the volume of passengers per door. Having such precise data allows for more in-depth analyses, such as the spread of boarding passengers.

The third difference between both datasets which is worth mentioning is the number of trains equipped with an automatic passenger count system. For the trains in Skåne, all 99 trains in circulation are equipped with such a system. This is, however, not the case in Stockholm where approximately every seventh commuter train is equipped with an automatic passenger count system. This means that the data provides a sample of the population in the case of Stockholm, whereas the entire population is captured in Skåne.

The fourth difference between both datasets that is worth pointing out is the way in which the dwell time is measured by the on-board systems. In the case of the trains in Skåne, the dwell time is measured as the time between the doors unlocking and locking. This means that the measured dwell time effectively captures the start of the alighting process up to the time the train is ready to depart, omitting the time it takes to dispatch a train. In the case of the commuter trains in Stockholm, the dwell time is measured based on the wheel-stop and wheel-start time. Here the dwell time thus also includes the dispatching time, which can lead to somewhat overestimating the effect of passengers on dwell times as delays can be a result of slow dispatching or slow acceleration of trains.

7 Data processing

7.1 Combining the different data sets

Before the train movement, timetable, and automatic passenger count data could be used for the analyses they had to be combined. Combining the different datasets made it easier to handle them and extract the desired information. This process was necessary throughout all quantitative analyses performed as part of this thesis, since no single database holds all the relevant information, and the task of combining the datasets took up a large portion of the work that was done. Several scripts were written using SQL (structured query language) to combine the different datasets. SQL is a programming language designed to process information stored on a relational database, making it well-suited for data handling. In order to be able to merge the different datasets it was important to identify which variables are similar between the datasets. Table 7 shows the different variables which were used to merge the datasets, and whether or not they are present in a given dataset.

Table 7: Available variables to be used to merge different datasets.

	Train movement data	Timetable data	APC-data from Skåne	APC-data from Stockholm
Train ID	✓	✓	✓	X
Date	✓	✓	✓	✓
Arrival time	✓	✓	✓	✓
Departure time	✓	✓	✓	✓
Station	✓	✓	✓	✓
Origin	✓	✓	✓	✓
Destination	✓	✓	✓	✓

The rationale for the merging of datasets based on three variables is shown in Figure 17. For this example, the *target train* represents an observation from the signalling data that needs to be matched with an observation from the automatic passenger count data. If the match is made using only the Train ID this would return four observations as a match, the same is true when only using the date. Three matches would be returned if only the station is used as the matching criteria. A better matching criteria should thus be used, which can be done by triangulating the three criteria as this leads to a single match between both datasets.

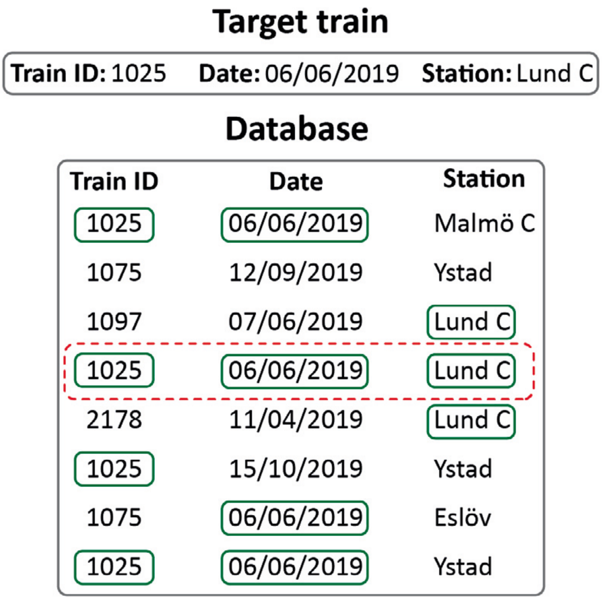

Figure 17: Rationale for the merging of datasets based on three variables using SQL.

The previous example is made possible since all datasets have a Train ID allowing for an easy match. The Train ID indicates the train number for that day and is used only once per day for a train running between its origin and destination. Combining the Train ID, station, and date thus leads to a unique identifier which can be used to make matches between different databases. The data originating from the automatic passenger count system in Stockholm lacks this Train ID. The merging of that dataset with other datasets is, therefore, more complicated. The rationale for this process is shown in Figure 18.

Target train

Date: 06/06/2019 **Station:** Stockholm C **Arrival time:** 09:06 ±60s
Origin: Märsta **Destination:** Södertälje centrum

Database

Date	Station	Arrival time	Origin	Destination
06/06/2019	Stockholm C	09:08:01	Märsta	Södertälje centrum
12/09/2019	Skogås	10:12:57	Västerhaninge	Bålsta
07/06/2019	Odenplan	09:06:35	Nynäshamn	Bålsta
06/06/2019	Stockholm C	10:12:12	Södertälje centrum	Märsta
11/04/2019	Älvsjö	16:08:47	Tumba	Upplands Väsby
15/10/2019	Sollentuna	12:32:45	Uppsala C	Älvsjö
06/06/2019	Stockholm C	09:05:55	Märsta	Södertälje centrum
06/06/2019	Stockholm C	09:06:12	Södertälje centrum	Märsta

Figure 18: Rationale for the merging of datasets without having the Train ID numbers.

The *target train* again represents an observation from the signalling data that needs to be matched with an observation from the automatic passenger count data. Since the Train ID is missing in the latter, a different approach was needed here. In this case, the arrival time, origin, and destination need to be used along with the date and station criteria. A margin of error of *sixty seconds* was allowed for the arrival time since the timestamps in the signalling data are aggregated on a minute basis, whereas the timestamps from the automatic passenger count data are recorded in seconds.

Again, if only using the date or station as a criterion, four matches would be returned. Using the arrival time or the combination of the origin and destination would return two matches in both cases. Using the date, station, and arrival time as criteria would still result in two matches. This is because it is possible for two trains running in the opposite direction of each other to halt at the same station around the same time. To overcome this all five criteria were used together to make sure only a single match between both datasets was made.

7.2 Correcting for early arriving trains

It is important to correct for early arrivals when dealing with dwell time deviations. Early arriving trains cannot depart before the scheduled departure time, leading to the dwell time being extended. These dwell time deviations are, however, not delays when the train still departs on time. The frequency of early, on-time, and late arrivals for each dataset is shown in Table 8, split out per year. Given the high frequency of early arriving trains, it is clear that not correcting dwell times for early arrivals can lead to severe overestimating the frequency and size of dwell time delays. It must be noted that the low portion of on-time arrivals in both datasets can be partially explained by the arrival punctuality being measured to the nearest second.

Table 8: Frequency for early, on-time, and late arriving trains based on the difference between the actual and scheduled arrival times, measured to the nearest second.

APC data from Skåne				**APC data from Stockholm**			
Year	*Early*	*On-time*	*Late*	*Year*	*Early*	*On-time*	*Late*
2017	58%	8%	34%	2013	59%	1%	40%
2018	64%	1%	35%	2014	64%	1%	35%
2019	61%	1%	38%	2015	67%	1%	33%
2020	51%	1%	49%	2016	67%	1%	32%

The steps used to correct for an early arriving train are as follows:

Step 1: determine if a train arrived early or on time. If yes, continue to step 2a, if a train arrives late, the adjusted dwell time deviation should be calculated using step 2b instead.

Step 2a: determine whether the train suffered a departure delay. If yes, then the dwell time deviation is equal to the departure delay. If no, then the dwell time deviation is equal to zero.

Step 2b: the dwell time deviation is calculated by taking the difference between the scheduled dwell time and the actual dwell time.

A train arriving early and leaving on-time will thus be registered as having a dwell time deviation of zero, even though the actual dwell time is larger than the scheduled dwell time. If a train arrives early and leaves late, only the departure delay is measured as the dwell time delay and the additional time from leaving early is excluded from this dwell time delay.

8 Summary of the included papers

Paper I

The passenger's influence on dwell times at station platforms: a literature review
Kuipers, R. A., Palmqvist, C.-W., Olsson, N. O. E., & Winslott Hiselius, L. (2021); Transport Reviews

The study in Paper I aimed to present how passengers affect dwell times, both before and during the boarding and alighting process, along with the underlying causes. The chosen method was a systematic literature review which took place in March 2020 and included key terms related to passenger distribution as well as search terms related to boarding and alighting behaviour.

Reviewing 59 papers, published between January 2000 and March 2020, two distinct ways in which passengers influence dwell time during the boarding and alighting process were identified. The first aspect is the behaviour shown by passengers around the doors of a train. The reviewed studies showed that when boarding passengers crowd around the doors, the effective width of the door will be reduced which slows down the flow of alighting passengers. This behaviour was shown to be more prominent when a larger number of boarding passengers was present and is possibly related to increased stress levels. The second way in which passengers influence dwell times is the manner in which boarding passengers distribute themselves across the platform. When this spread is uneven, some doors will be over-used whilst other doors are under-used during the boarding process. This means that the boarding time will be extended at the busiest doors, whilst the boarding process is already finished at some other doors.

Based on these findings it was concluded that adaptations to the design of rolling stock will only be beneficial if the behaviour of passengers is adapted as well. Adding wider doors will only result in benefits if passengers do not crowd around the door, and adding more doors is only beneficial when passengers spread out more evenly. The results also indicate a possible connection between the spread of boarding passengers and the behaviour shown at the door itself, given that stress levels increase when more passengers are present. When boarding passengers are not evenly spread, some doors are more likely to experience crowding. This findings suggests that both the spread and behaviour of passengers at the door should be addressed when implementing measures to reduce dwell time delays.

Paper II

Passenger Volumes and Dwell Times for Commuter Trains: A Case Study Using Automatic Passenger Count Data in Stockholm
Kuipers, R. A., & Palmqvist, C.-W. (2022); Applied Sciences

The aim of Paper II was to understand the relationship between the volume of boarding and alighting passengers and the frequency of dwell time delays. The frequency of delays was chosen as the variable to study since this is more robust than using the actual size of dwell time delays. Furthermore, it can be argued that it is of more interest to understand how often a certain delay happens, and which delays are most common, compared to knowing the total or average delay size. The study made use of automatic passenger count data collected on board commuter trains in Stockholm between 2013 and 2016. The data provided both the volume of passengers and actual dwell times on a magnitude of seconds. Prior to the statistical analyses, the passenger volumes were grouped in bucket sizes of ten boarding and alighting passengers in order to define the sample space with respect to passenger volumes. The next step involved comparing the actual dwell times to the scheduled dwell time in order to count the number of dwell time delays. The dwell time delays were split then up into intervals of fifteen seconds. The frequency of dwell times given a specific volume of boarding and alighting passengers was then plotted and a visual graphical analysis was performed to highlight trends in the data.

A Chi-square goodness of fit test was used to test for the effect of the number of passengers on the frequency distribution of dwell time delays, comparing the frequency distribution of dwell time delays under different passenger volumes against an unconditional distribution of dwell time delays. The unconditional distribution was based on all observations without making a distinction in terms of passenger volumes. The hypothesis tested states that if passenger volumes do not change the frequency of dwell time delays, all frequency distributions for dwell time delays will follow the unconditional distribution regardless of the volume of boarding and alighting passengers. The initial Chi-square test was complimented with post-hoc pairwise testing when a statistically significant result was found and the effect sizes for significant results were calculated based on Cohen's W.

The results show that dwell time delays occur more often as passenger volumes increase, but the delay size itself does not necessarily increase as passenger volumes increase. This is in contrast with other studies which indicate a linear relationship between dwell times and passenger volumes, where the frequency of delays would thus increase as the volume of passengers increases. It is only after a certain number of boarding and alighting passengers is present that the size of delays increases as well, and based on the data used in Paper II, this threshold seems to be around 20 boarding and alighting passengers.

Paper III

The spread of passengers on platforms and dwell times for commuter trains: A case study using automatic passenger count data

Kuipers, R. A., & Palmqvist, C.-W. (2022); 11th Triennial Symposium on Transportation Analysis conference (TRISTAN XI), Mauritius

The findings in Paper I indicated that the spread of passengers across the train doors can have a negative effect on dwell times. The studies found during the literature review are, however, limited to small sample sizes which can mean that the effect is less prominent than previously reported. The automatic passenger count data collected on board commuter trains in Skåne allows to study the spread of boarding passengers on a network-wide scale, using a large number of observations.

Using this data, the work done in Paper III focussed on understanding the impact of the spread of boarding passengers on the probability of a dwell time delay on a network-wide level. This was done with the aim of gaining an understanding with respect to where it is relevant to make interventions aimed at spreading out passengers. A logistic regression model was used to determine the probability of a dwell time delay occurring given the spread of passengers, in combination with other variables such as the volume of passengers and the scheduled dwell time.

Comparing situations with an uneven and even spread of passengers under different combinations of passenger volumes, proportion between boarding and alighting passengers, and scheduled dwell times, it was found that an uneven spread of passengers is most influential at higher volumes of boarding passengers. The spread of boarding passengers was found to only have a limited effect when passenger volumes are low. Although the probability of delays still increases in such a case this increase was found to be only marginal. As such, it was concluded that the return on investment in terms of the costs of applying measures to spread out passengers versus the gain in dwell time punctuality will be greatest at stations where the expected volume of passengers is relatively large and such measures are less effective at stations with relatively low passenger volumes.

Paper IV

Impact of a lower demand during the COVID-19 pandemic on the frequency of dwell time delays

Kuipers, R. A., & Palmqvist, C.-W. (2023); Transportation Research Interdisciplinary Perspectives

Given the impact of passenger volumes shown in Paper II, it was of interest to understand what happens with dwell time delays if fewer passengers travel. Understanding the effect of a reduced volume of passengers can assist in understanding whether spreading the load of passengers across multiple trains can help to reduce dwell time delays. Previous studies have indicated positive benefits of flattening the curve of peak passenger demand on the need for rolling stock, but the impact on dwell time punctuality did not receive much attention.

The COVID-19 pandemic provided an unprecedented opportunity to study the effect of a sustained reduction in passenger volumes on the dwell time punctuality of commuter trains. This is especially the case since Sweden did not introduce hard lock-down measures, meaning there were still commuting passengers. To study the effect of a reduced passenger demand on dwell time delays, three years of automatic passenger count data from Skåne was used. For the purpose of this study, the period of 2018 and 2019 was used as the pre-COVID period and data from 2020 as the COVID period. The study was limited to observations during peak hours to capture passengers who had to travel during both periods and can be assumed to be familiar with the railway system.

The study made use of different statistical tests. A Chi-square goodness of fit test was used to understand whether the frequency of dwell time delays changed under COVID-19 conditions. A Kruskal-Wallis test was used to determine whether there were statistically significant changes in passenger flow characteristics. Inferences were then made based on the outcome of these tests. The initial findings indicated that there was a change in the frequency of dwell time delays, where a sharp decrease in the frequency of dwell time delays was observed during the beginning of the pandemic. The change in the frequency of dwell time delays coincided with a drop in the peak load of the number of boarding passengers, indicating a possible relationship. Further testing, by making use of a Wilcoxon signed-rank test, showed that the median length of dwell times reduced across a range of passenger volumes. This suggests that the increased number of non-delayed stops is not only caused by a reduction in passenger volumes. It was concluded that a reduced peak load of passenger volumes is likely to be a cause for the improvement found in dwell time delays, but this is not the sole reason. In order to reduce both the size and frequency of dwell time delays it can thus be beneficial to spread the load of passengers across multiple trains, but this will not be sufficient to ensure punctual dwell times and other measures are needed.

Paper V

Dwell-time Station-Service analysis using Rasch analysis technique

Kuipers, R. A., Tortainchai, C., Tony, N.C., & Fujiyama T; (article under review – Transportation Research Interdisciplinary Perspectives – first round of reviews)

The aim of Paper V was to find a way in which the dwell time performance of both stations and services on a specific railway line can be studied in a single dimension. The study made use of data from commuter trains in the UK and commuter trains in Skåne and the method proposed in this study is the Rasch analysis technique. The aim of the study was two-fold. The first part aimed to show the applicability of the Rasch analysis technique in dwell time evaluations, given that it is a tool normally used for questionnaire studies rather than within an operational context. In addition to this the study aimed to highlight how the output of the Rasch analysis can be used to study the dwell time performance of a railway line.

The Rasch analysis allows to combine both the station and service performance into a single dimension. This is done by comparing the relative ability of a service to have a punctual dwell time with the difficulty of all services to have a punctual dwell time at a given station. This indicator was introduced as the *dwell time performance score* in Paper V, where a higher score indicates a lower chance of dwell time delays and vice versa. Being able to calculate an indicator that combines both the relative station and service performance has a benefit over the more common approach of taking the average delay at a station since it allows one to identify which service, or group of services, is most likely to incur a delay.

The results of the study show that the Rasch model can be used within an operational context and can adequately reflect the high and low-performing stations and services. The dwell time performance scores also revealed that, within the railway lines studies in Paper V, there are no stations for which all services are likely to incur a dwell time delay. Such insights would not be available when using the average dwell time delay for stations or services. It was, therefore, concluded that the Rasch analysis technique can be applied within an operational context such as dwell time studies, and that the outcome of a Rasch analysis can help to identify problematic services. Having insights into which services are problematic in terms of dwell times will allow planners to more actively target efforts to reduce dwell time delays.

Paper VI

Understanding dwell times using automatic passenger count data: A quantile regression approach

Kuipers, R. A. (2024); Journal of Rail Transport Planning & Management

The objective of the study that formed Paper VI was to study the effect of different explanatory variables on the conditional distribution of dwell times. The study made use of a quantile regression model. Although a common way to study these relationships is to make use of ordinary least square regression models, such models are sensitive to outliers, can only model the conditional mean, and require data to be normally distributed. In contrast to this, quantile regression allows to study the conditional distribution across different percentiles. This means that in addition to modelling the conditional median, a quantile regression approach also allows to study the effects of the independent variables on the dependent variable in the tails of the distribution. The quantile regression is also more robust to outliers and heteroscedasticity, making it more suitable for dwell time research given the nature of dwell time data. The final model included variables relating to passenger flow characteristics, the operation of trains, and the historical dwell time at stations. These variables were previously identified to be relevant with respect to dwell times.

The model fit was found to be promising and summary plots showed that coefficients from the quantile regression model differ from those originating from an ordinary least squares regression model. Furthermore, the coefficients of the quantile regression model were found to change across the different percentiles. This indicates that the use of quantile regression over ordinary least squares regression is justifiable and beneficial when studying dwell times.

Numerical examples were used to further understand the impact of the different variables included in the model. The results of these numerical examples indicate that it is important to account for station-specific characteristics when scheduling dwell times, which is currently not common practice. This is shown by the impact of the passenger flow characteristics and the historical dwell times, which are all variables that are different from station to station. In addition to this, the results further highlight that the volume of boarding passengers is not the main determining factor for the duration of dwell times since the spread of passengers is found to play a larger role. The third observation is that arrival punctuality plays a major role in the duration of dwell time. Although trains arriving early do not necessarily lead to delays, the longer dwell times can be problematic in terms of capacity utilization.

9 Answers to the research questions

Research question 1

What are the causes of dwell time delays for commuter trains?
Paper I, VI

The first research question of this thesis is aimed at identifying possible causes for dwell time delays, specifically within the context of commuter trains. As shown in *Section 2.4*, several general causes for dwell time delays have been identified by Harris et al. (2013) and Pritchard et al. (2021). Their causes can be summarized in three distinct groups, these being the scheduling principles of dwell times, the operation of trains, and the presence of passengers. The results from Paper VI indicate that the operational variables such as the historical dwell time at a given station and arrival punctuality of a train can have a stronger influence on the length of dwell times compared to passenger flow related variables. Trains arriving early will lead to extended dwell times since such trains have to wait longer than scheduled before they can depart. Similar findings were reported by Kecman and Goverde (2015). It must be noted that a train can still depart on time in these instances, so although early arrivals can have a strong influence on dwell time deviations the impact on dwell time delays is likely less prominent.

The way in which dwell times are scheduled can be a cause for dwell time delays. Past studies have highlighted that when the scheduled dwell time is not realistic this increases the chance of a dwell time delay occurring. Scheduling realistic dwell times means that the scheduled dwell time should account for the dynamic nature of passenger flows and be aligned with the expected passenger demand. This latter aspect also suggests that the process of dwell time scheduling should be revised over time to reflect changes in passenger demand. In practice, this is often not the case, however. In Skåne, for example, the scheduled dwell time remained the same between 2012 and 2020. This whilst the passenger demand did change over the same period of time. Furthermore, current scheduling practices in Sweden dictate that dwell times are scheduled the same for peak and off-peak hours and all services in circulation. This is likely done in favour of cyclicity and symmetry in the timetable which has benefits when designing a timetable and makes the schedule easier to understand for passengers.

This rather static approach to scheduling dwell times does, however, not align with reality as passenger demand changes throughout the day and the necessary dwell time during peak hours can be different to that during off-peak hours at the same station. Current scheduling principles in Sweden also define the dwell time to be the same for most stations within a network, meaning that station-specific characteristics such as the passenger demand and layout of a station are not taken into account when scheduling dwell times.

With respect to the operation of trains, the literature reviewed in Paper I suggests that staff can influence the duration of dwell times by allowing late arriving passengers to board. Another aspect of the operation of trains which can be the cause of dwell time delays is the maintenance of the rolling stock. Broken doors mean that the effective number of doors available to boarding passengers is reduced, and this results in passengers having to reposition to the nearest door upon the arrival of the train.

Although dwell time scheduling principles and operation of trains can be considered to be case specific, meaning that the above-mentioned causes are not omnipresent, a consensus found in the literature reviewed in Paper I is that the time it takes for the boarding and alighting process to be completed effectively defines the duration of dwell times. More specifically, as the study by Buchmueller et al. (2008) showed, the boarding and alighting time at the most critical door defines the total dwell time of a train in a normal situation. In this case, the most critical door is the door where boarding and alighting take the longest. Since the boarding and alighting process defines the length of dwell times, this is also the most likely cause of dwell time delays.

To conclude, it can be stated that the dwelling process is complex and no single cause can be identified as to why dwell time delays occur. Nevertheless, the findings presented here suggest that the main cause for dwell time delays under normal operating conditions, i.e. without door failures, is the boarding and alighting process taking longer than what is accounted for in the scheduled dwell time. On the one hand, this is possibly a result of unrealistic dwell time scheduling, where the scheduled dwell time does not accurately reflect the actual time needed to complete the boarding and alighting process. On the other hand, it is possible that the boarding and alighting takes longer than scheduled as a result of the behaviour of passengers.

Research question 2

How do boarding and alighting passengers influence the duration of dwell times for commuter trains?
Paper I, II, III, IV, VI

Given the influence of passengers on the duration of dwell times, it is of interest to further understand how boarding and alighting passengers influence the duration of dwell times. The conceptual model of dwell times by Li et al. (2016) defines the flow of passengers as a function of the volume of passengers, passenger characteristics such as their mobility, and the capacity of the door. This definition does, however, not capture all factors that influence the flow of passengers. The behaviour of boarding and alighting passengers, such as how boarding passengers spread out and how queues are formed around the doors, are also important aspects concerning the time needed to complete the boarding and alighting process.

Building upon the conceptual model of dwell times presented by Li et al. (2016) and the hierarchy of passenger influences on dwell times presented in Paper I, a flow diagram of how passengers influence dwell times was defined. This flow diagram is presented in Figure 19. Two distinct situations in which passengers influence dwell times have been identified, the first being clustered boarding and alighting. Clustered boarding and alighting is characterized by boarding passengers being present on the platform before the train arrives. The second situation that has been identified is that of late arriving passengers, characterized by boarding passengers arriving near or at the end of the dwelling process close to the scheduled departure time of a train. The two situations can both be present during a dwelling process, with the majority of passengers being present ahead of the arrival of a train and a small number of boarding passengers arriving close to the departure time of a train.

The effect of late arriving passengers is rather straightforward. In such cases, delays arise when the door-closing sequence cannot be initiated since a train cannot depart before all doors are locked. One reason for the doors not closing is excessive service by a member of staff, holding a door open for passengers who arrive on the platform late to allow those passengers to still board the train. Another way in which late-arriving passengers can extend dwell times is through an act often referred to as *door holding* or *door forcing.* In such cases, a late arriving passenger holds or forces the door open which impedes the closing of the door. Door holding not only stops the doors from locking it can also necessitate the door locking procedure to be reinitiated, even further extending dwell times.

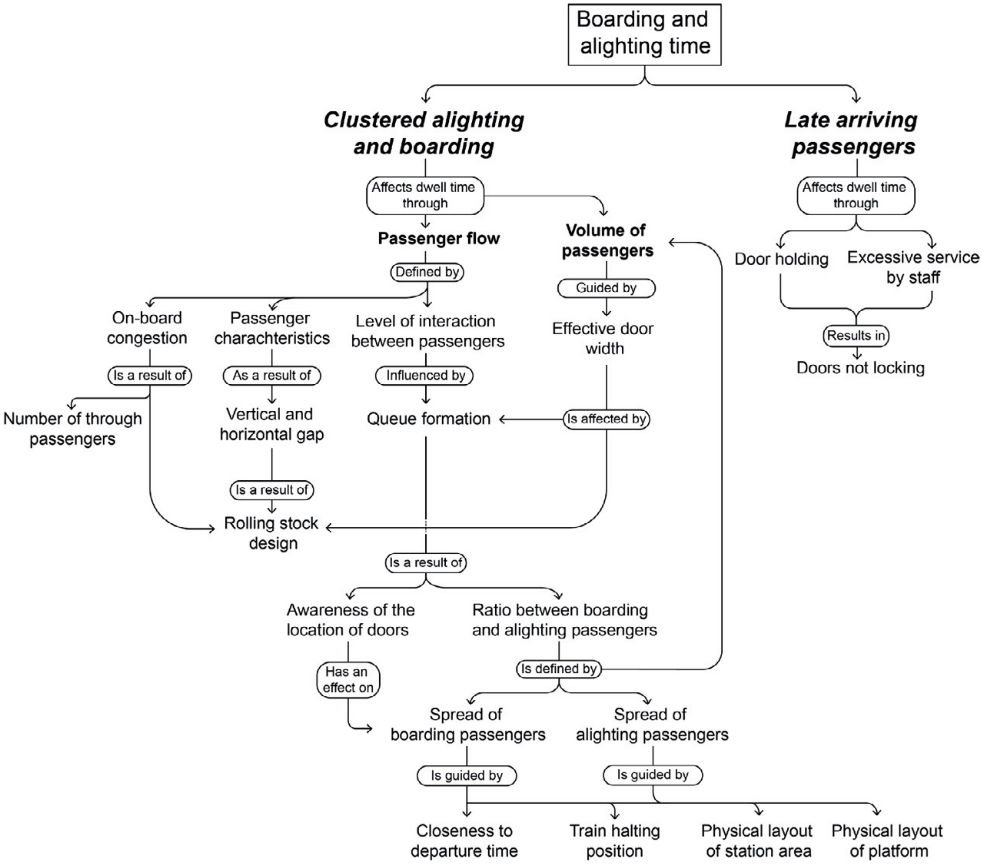

Figure 19: Flowchart for the influence of passengers on the boarding and alighting time.

The way in which passengers influence the length of dwell times is more complicated during the process of clustered boarding and alighting, where the flow of passengers and the volume of passengers define the time needed to complete the boarding and alighting process. The flow of passengers is defined by a combination of the level of on-board congestion, passenger characteristics such as the presence of luggage and their level of mobility, and the level of interaction between boarding and alighting passengers. The way in which these aspects define the flow of passengers and ultimately the boarding and alighting time is a result of both the design of rolling stock and passenger behaviour[2].

[2] For an in-depth explanation of the influence of both passengers and rolling stock design on dwell times, I refer back to *Section 2.5* and *Section 2.6* of this thesis.

The next paragraphs provide a brief description of the way in which the four aforementioned factors affect the flow of passengers and some of the underlying processes involved. These descriptions are based on the literature reviewed in Paper I.

The flow of passengers refers to the number of passengers that pass the door per second and is thus closely related to the speed at which passengers can either alight or board. The speed at which passengers can alight and board is a result of their characteristics and the horizontal and vertical gaps between a train and a platform. Passengers with reduced mobility or who carry luggage can have more difficulties traversing larger gaps between the train and the platform. As a result, the flow of passengers slows down in such cases, extending the time needed to complete the boarding and alighting process. These gaps are a result of the rolling stock in use, where design choices are made such as the presence of a level entry or not.

The flow of passengers is also defined by the interaction between boarding and alighting passengers, where increased levels of interaction will slow down the flow of passengers. These interactions take place in the platform train interface area, in front of the doors of a train. This level of interaction is influenced by the formation of queues by passengers waiting to board. When queues are formed in front of doors the level of interaction between the flow of alighting passengers and the queue of boarding passengers will increase as there is less space for alighting passengers to move through. The way in which boarding passengers position themselves around the doors of a train also affects the effective door width, which becomes narrower when passengers stand in front of the door as opposed to next to the door.

The effective door width is relevant with regard to the effect of the volume of passengers on the boarding and alighting time. When more passengers are present the time needed to complete the boarding and alighting process will likely be longer. The extent to which the volume of passengers defines the boarding and alighting time is guided by the effective door width. Wider doors can allow for multiple lanes of boarding and alighting passengers, effectively doubling the flow of passengers through a single door. The door width is, again, a result of the rolling stock in use where a choice between wider or narrower doors is made. This is, however, the designed width and is not always the same as the effective door width. As mentioned before, the effective door width is dependent on the formation of queues. When boarding passengers queue in front of the door it is not possible to have multiple lanes of alighting passengers, which slows down the flow of alighting passengers and will increase the time needed for all passengers to alight.

The formation of queues is a result of the ratio between boarding and alighting passengers as well as the awareness of passengers regarding the location of the train doors. When the location of doors is indicated through the use of platform screen doors, for example, queues are more likely to be formed next to instead of in front of the doors. When the ratio between boarding and alighting passengers is skewed in favour of boarding passengers, it is more likely that queues are formed in front of the door. This behaviour is hypothesized to be a result of increased stress levels and the fear of not having a seat. This is more prominent when the majority of passengers are boarding since there is less pressure from the volume of the opposite flow of alighting passengers. The ratio between boarding and alighting passengers is in part defined by the spread of passengers between the doors of a train, along with the volume of passengers at a given station. When there is an uneven spread, a majority of passengers either board or alight through a limited number of doors, and the ratio will be skewed at those specific doors.

The spread of boarding and alighting passengers is guided by the physical layout of the origin and destination stations. When passengers are aware of the layout of their destination station, specific doors are favoured to reduce the walking distance at their destination station. On the other hand, when passengers are not familiar with the layout at their destination they are likely to stand near already present obstacles on the platform such as pillars or ticket machines, make use of the available roof coverage, and wait near the entrance points of platforms. Since boarding passengers are likely to board through a door close to their waiting position and movement through a train is often limited, the waiting position of passengers defines which door is used to board a train as well as the door used to alight in most cases.

The spread of boarding passengers is also guided by the awareness of the train halting position and the closeness to the departure time when entering the platform. Circulation of passengers on platforms slows down as the departure time of a train is closer, meaning that more passengers will wait near entrance points and a concentration of boarding passengers will occur at those places. Furthermore, when passengers are unsure about where the doors of a train will be, a higher concentration of passengers can be found in central locations of the platform and platform entrances as passengers opt to wait in a location where they are most sure that the train will halt.

Once boarding passengers make it past the queues and through the door of a train it is the level of on-board congestion which defines the flow of passengers. On-board congestion is a result of the number of through passengers, i.e. the passengers that remain on the train, and the interior design of a train. Internal bottlenecks such as luggage racks and narrow doorways have been found to slow down the flow of passengers, extending the boarding time. It is not just the layout of a train but also the degree of on-board crowding which can cause congestion, where the flow of passengers is slowed down due to a lack of space for the boarding passengers to occupy.

Quantifying the effect of passengers on dwell times

The explanations above provide an overview of previous research into the effect of passengers on dwell times. The next step is to quantify the effect of passengers on dwell times. Based on the available data from the automatic passenger counting systems it was possible to quantify the effect of the volume of passengers, the ratio between boarding and alighting passengers, and the spread of boarding passengers on the occurrence of dwell time delays. The following sections describe these findings.

The impact of the volume of passengers

A common assumption is that the volume of passengers is the main factor that defines the length of dwell times, and several studies have shown that the volume of passengers has a strong influence on the duration of dwell times. In most cases, these studies assume a linear relationship between the volume of passengers and the duration of dwell times (see for example: Antognoli et al., 2018; Lee et al., 2018; Palmqvist et al., 2020). This linear relationship means that a given amount of additional dwell time is to be expected for every additional passenger. However, the findings presented in Paper I show that such a linear relationship is likely, not present since multiple studies have found that the flow of passengers can change throughout the boarding and alighting process due to various reasons.

The findings from Paper II also suggest that such a linear relationship between the volume of passengers and dwell time delays is indeed not present. Studying the frequency of dwell time delays under different volumes of passengers it was found that the frequency of delays does increase when the volume of passengers increases, but the size of these delays does not necessarily increase along with this. Taking the dwell time process as a whole, one additional passenger does thus not result in a fixed amount of additional dwell time. This being said, there is an effect of the volume of passengers on the frequency of a delay occurring, indicating that there is some effect of the volume of passengers on the likelihood that a dwell time delay will occur.

Based on the effect found in Paper II, it is possible to assume that if fewer passengers travel, the frequency of dwell time delays will be reduced. Testing this hypothesis is not as simple as comparing the peak to the off-peak hours as the type of passenger would be different, commuters versus non-commuters, but requires a period in which there is a reduction in the volume of the same type of passengers. The COVID-19 pandemic and the Swedish government not introducing a hard lockdown provided exactly this opportunity. During the pandemic fewer people commuted, making it possible to study the effect of a sustained lower volume of passengers on dwell times. This was done in Paper IV by comparing dwell times and passenger flows of the two years prior to the COVID-19 pandemic to the first year of the pandemic.

The results in Paper IV show that an improvement in dwell time punctuality occurred in the same period as passenger volumes decreased. This initial result pointed towards the drop in the volume of passengers to be the main way in which passenger volumes have an effect on the frequency of dwell time delays. However, a pairwise comparison of the median observed dwell times revealed that an improvement was present across all passenger volumes, both lower and higher volumes, meaning that the reduction in passenger volumes was not the sole cause for the reduction in dwell time delays.

To conclude, an increased volume of passengers leads to an increased frequency of dwell time delays. The size of these delays does, however, not change directly with an increase in the volume of passengers. This indicates that an increase in the number of boarding and alighting passengers does not always result in longer dwell times. Furthermore, a reduction in passengers was not enough to explain the improvement in terms of dwell time punctuality during the COVID-19 pandemic. This indicates that although the volume of passengers plays a role in the length of dwell times, other factors play an important role as well.

The impact of concentrated boarding

The literature reviewed in Paper I identified the spread of boarding passengers as an important factor with regard to dwell times. Studies by Fox et al. (2017) and Oliveira et al. (2019) introduce the term *concentrated boarding* for this phenomenon and show that it leads to substantially longer dwell times. Both of these studies made use of a limited number of video observations to study the effect of concentrated boarding on dwell times, however. To better understand the impact of concentrated boarding on dwell times it is of interest to study this relationship on a larger sample size, which was done in Paper III in which the probability of a delay under different combinations of the spread of boarding passengers, the volume of passengers, and the ratio between boarding and alighting passengers was studied. The findings reinstate the effect found by Fox et al. (2017) and Oliveira et al. (2019), be it less prominent.

The results presented in Paper III show that the probability of a dwell time delay increases as a result of an uneven spread when passenger volumes are high, but less so in conditions with lower passenger volumes. Although concentrated boarding thus leads to longer dwell times and vice versa, this effect seems to only be present when the volume of boarding passengers is relatively high. These findings are quite intuitive since having a very small number of passengers waiting to board a train through a single door should indeed not require much more time compared to the same volume of passengers boarding through multiple doors.

The ratio between boarding and alighting passengers

The ratio between boarding and alighting passengers was used as a proxy for the interaction between passengers in the platform train interface area. A ratio skewed towards boarding passengers is likely to result in an increased level of interactions between boarding and alighting passengers and thus longer dwell times. An effect that was previously also stated by Seriani et al. (2019) who found that the boarding and alighting time increases as the ratio is shifted towards a majority of boarding passengers.

The ratio of passengers was, therefore, also included in the studies performed in Paper III and Paper VI. In both cases, the findings show that the impact of the ratio of passengers is not as prominent as the spread and volume of passengers. Furthermore, where the studies reviewed in Paper I indicate that a ratio in favour of boarding passengers will likely lead to longer dwell times, the findings in Paper VI contrast this. Here a negative relationship between dwell times and the proportion of boarding passengers was found. This negative relationship indicates that dwell times are likely to be shorter when the ratio between boarding and alighting passengers is shifted towards a majority of boarding passengers. No clear explanation for this reversed effect of the ratio of boarding passengers can be given, however. To conclude, the findings from both Paper III and Paper VI indicate that the ratio between boarding and alighting passengers on its own does not have a strong influence on the duration of dwell times.

The volume of passengers as an accelerator

Where previous studies have indicated that the volume of passengers is the main determinant for the duration of dwell times, the findings presented in this thesis suggest that the volume of passengers is not the main cause of dwell time delays. Instead, it acts as an accelerator for other effects such as the degree of concentrated boarding. To illustrate this, an excerpt from Paper III is presented in Table 9. The table shows the probability of a dwell time delay under different combinations between the volume and spread of passengers for trains arriving on time, during peak hours, and with a median ratio between boarding and alighting passengers.

Looking at the probabilities in the table the following can be noted. When the spread of boarding passengers is even, an increase in the volume of boarding passengers does not have a strong effect on the probability of a dwell time delay. In such cases, the probability increases by one percentage point between the 25th percentile and 75th percentile for the observed volume of passengers. Furthermore, comparing the probability of a dwell time delay between an even and uneven spread at the lower end of the passenger volumes shows that there is only a small effect of the spread of passengers on the probability of a dwell time delay. Such cases see an increase in the probability of a dwell time delay of five percentage points.

Table 9: Probability of a dwell time delay when a train arrives on time, during peak hours with a median ratio of boarding versus alighting passengers.

	Even spread of passengers	**Uneven spread of passengers**
Volume of boarding passengers= 25th percentile	53%	58%
Volume of boarding passengers = median	53%	63%
Volume of boarding passengers = 75th percentile	54%	76%

In contrast to this, the probability of a delay increases rapidly as passenger volumes increase when the spread is uneven, showing an increase of eighteen percentage points between the lower and upper end of the volume of boarding passengers. Furthermore, where the difference in probabilities between an even and uneven spread is only five percentage points for the 25th percentile for the observed volume of passengers, this effect is more prominent when passenger volumes are high. In such a case, having an uneven spread of boarding passengers increases the probability of a dwell time delay by twenty-two percentage points compared to a situation with an even spread of boarding passengers. This shows how the volume of passengers can act as an accelerator for the negative impact of an uneven spread of passengers on the probability of dwell time delays.

Swiss cheese model for passenger induced dwell time delays

In addition to showing how the volume of passengers can act as an accelerator, the example above also highlights how a large volume of passengers on its own does not necessarily lead to an increased probability of dwell time delays when the spread of passengers is favourable. This suggests that it is a combination of unfavourable factors, rather than a single variable, which leads to an increased risk of dwell time delays. Such a relationship can be visualized using a Swiss cheese model.

The Swiss cheese model, proposed by Reason (2000), is a way to visualize how the interdependencies of different elements within a complex system relate to the risk of errors. Since its publication, the Swiss cheese model has become the dominant paradigm for analysing errors in systems (Perneger, 2005). According to the metaphor, several layers of a system need to line up for an error to occur and each layer of the system comes with its own weaknesses, or holes as in a Swiss cheese. It is only when all these weaknesses, or holes, align that an actual error will occur (Perneger, 2005; Reason, 2000). Extending this metaphor to dwell time delays, a "Swiss cheese model for passengers induced dwell time delays" is presented in Figure 20. The layers are represented by the volume of passengers, the spread of boarding passengers, the formation of queues in the platform train interface area, and the level of on-board congestion.

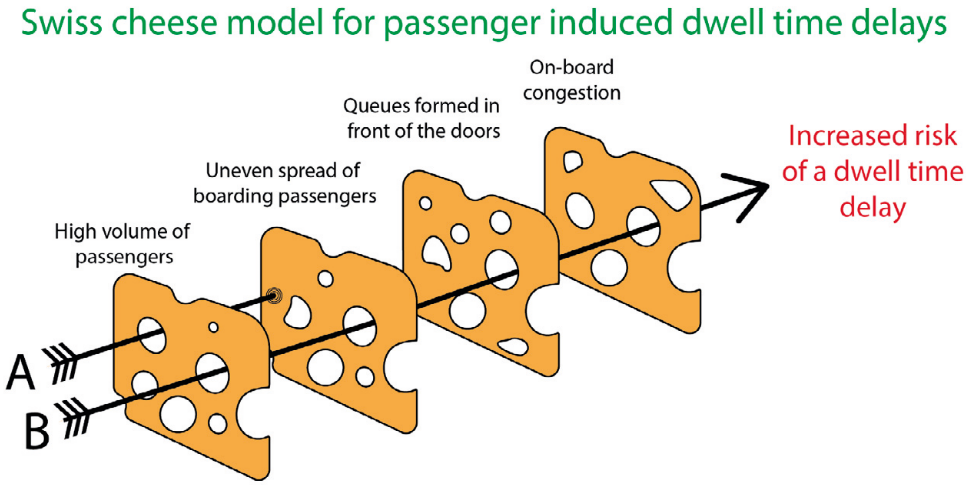

Figure 20: Swiss cheese model for passenger induced dwell time delays.

The "Swiss cheese model for passengers induced dwell time delays" works as follows. A high volume of passengers increases the risk of dwell time delays, however, with an even spread of boarding this additional risk will be somewhat mitigated as the findings in Paper III show, leading to a situation as indicated with arrow A. Here there is no increased risk of dwell time delays since the spread of boarding passengers is favourable.

When there is both a high volume of passengers and a strong uneven spread, both weaknesses line up. In this case, the next layer is the formation of queues around the doors and this will be the deciding factor for the risk of a dwell time delay. Based on the literature reviewed in Paper I, it is possible to assume that if boarding queues are formed in such a way that alighting passengers can form two rows of passengers, and the same can occur again with respect to the flow of boarding passengers, the risk of a dwell time delay will be lowered. On the other hand, if the queues are formed in such a way that the flow of alighting passengers is obstructed, dwell times are more likely to be extended. The final layer in the Swiss cheese model is on-board congestion. Even if boarding passengers can enter the train in a swift manner, the flow of boarding passengers will slow down if there is a high degree of on-board congestion.

It is worth pointing out that although late-arriving passengers have been identified to have an influence on dwell times as well, they are excluded from this model. The reason for this is that a passenger arriving late can increase the risk of a dwell time delay regardless of whether or not the clustered boarding and alighting was completed on time.

Research question 3

How can dwell time delays be studied in a network with automatic passenger count data?
Paper II, IV, VI

To understand how to reduce dwell time delays it is important to understand the magnitude of the problem and measure to what extent the actual dwell times exceed the scheduled times. The methods that are commonly used to measure punctuality do, however, not provide an adequate way to measure dwell time delays. One reason for this is that punctuality is often measured at the final stations thus hiding fluctuations along the journey of a train (Olsson & Haugland, 2004). Furthermore, dwell time delays are often shorter than the threshold used for punctuality statistics which are commonly around five minutes. Previous research by Palmqvist (2019) showed that 80% of delays at stations in Sweden are smaller than three minutes, thus falling well below the delay threshold used to measure punctuality. A similar observation is made in this thesis, with a large portion of delays being less than thirty seconds. Since current techniques to measure punctuality obscure the presence of dwell time delays, there is a need for a different approach to measuring the punctuality of dwell times.

A term often used alongside punctuality is that of reliability. A railway system can be considered reliable when the trains run on time most of the time (Vromans, 2005) and a similar definition can be applied to the reliability of dwell times. The dwell time is reliable when the actual dwell time corresponds with the scheduled dwell time, most of the time. Several different methods to measure reliability have been proposed in the literature. Rietveld et al. (2001) list the following indexes to measure reliability:

- The probability of a train arriving with a delay
- The probability of an early departure
- The mean difference between the expected arrival and the scheduled arrival time
- The mean arrival delay
- the mean arrival delay given that a train arrives more than a predefined amount of minutes late
- The standard deviation of arrival times

Although the list by Rietveld et al. (2001) is useful when the aim is to measure the reliability of a train going between its origin and destination stations, the proposed measurements cannot all be directly applied to measure dwell time punctuality or reliability. This being said, some of the proposed measurements are still applicable. It is possible to make use of the mean and standard deviations of dwell times, for example, as was done by Gysin (2018) and Pedersen et al. (2018). The mean dwell time can then be compared to the scheduled dwell time as a measure of punctuality, and the standard deviation provides an indicator of the stability of dwell times.

Using the mean and standard deviation is, however, not a robust way to measure dwell time delays. Robustness in the measurements of punctuality is important since the distribution of delays is often non-normal with a substantial amount of smaller delays and very few larger dwell time delays. In such a case using the mean as a measure of central tendency is not fitting and a more robust measure such as the median should be favoured (Rousseeuw, 1991). It must be noted that using a measure of central tendency, be it the mean or median, does not provide a complete picture of dwell time punctuality since it will not show how often a delay occurs. To gain more information it is important to go beyond just using the average delay.

Beyond the average delay

Several different approaches to measure dwell time punctuality that do not make use of a measure of central tendency exist. One of these is to make use of the probability of dwell time delays, see for example the study by Palmqvist (2022). The probability is deemed to be more robust compared to using the average dwell time delay since large delays are less influential. A downside of using probabilities to measure dwell time delays is that they only provide a binary approach. This means that a train is either delayed or not, and a train that is 10 minutes delayed is counted the same as one that is 1 minute delayed (van Loon et al., 2011).

A different way to measure dwell time punctuality is to use the probability, or frequency, of different delay sizes. This can be done by counting the occurrences of a delay within a given interval. An example of this is the study presented in Paper II where delays were split up into intervals of fifteen seconds. As with using the probability of a delay, counting how often a delay occurs is more robust to outliers since extreme values are only counted once and thus have less of an effect on the distribution. Different to using a binary probability, counting the frequency of delays allows to show how often a delay within a predefined interval happens. This has benefits over the use of a binary approach as it provides a better understanding of the size of the problem.

Consider the following dwell time observations, measured at a station with a scheduled dwell time of 60 seconds.

60 – 65 – 59 – 64 – 60 – 150 – 75 – 65 - 63

The median for these dwell times is 64 seconds, with the first and third quartiles being 60 and 65 seconds respectively. The conclusion, in this case, would be that small dwell times occur at this station. Calculating the probability of a delay occurring at this station by dividing the instances where the dwell time is larger than 60 seconds by the total number of observations, a the probability of a dwell time delay of 78% is found, indicating that delays are likely to occur at this station. Although both of these conclusions are correct, both cases obscure two important observations.

First, the median dwell time provides no indication of how often dwell time delays occur, and second, both cases obscure the occurrence of some rather large dwell time delays. This is why using the frequency of different delay sizes becomes more informative. When looking at the frequency counts of the observed dwell time, shown in Table 10, it can be observed that extreme delays do occur but not often, and that small deviations are quite common on the other hand. Having this insight can help inform the decision-making process since the focus could be placed on these small delays, for example.

Table 10: Example of frequency counts for dwell time delays.

Delay size	Shorter than scheduled	No delay	Max +5 seconds	Max + 10 seconds	Max + 15 seconds	> 15 seconds
Frequency	1	2	3	0	1	1

Measuring station and railway service performance in a single dimension

As stated by Pritchard et al. (2021) an important step in developing effective strategies to reduce dwell time delays is to identify hotspots where stations and/or railway services are likely to incur dwell time delays. Several different ways in which to identify such hotspots have been identified in Paper V, such as the use of clustering algorithms (Stoilova & Nikolova, 2017; Zemp et al., 2011; Zhou et al., 2022). Such clustering techniques are, however, not able to account for different types of data such as both passenger demand and service performance data at the same time. Data Envelopment Analyses and other data frontier approaches have been proposed to, somewhat, overcome this issue (Khadem Sameni et al., 2016; Tortainchai et al., 2022). Although these approaches can evaluate the efficiency of each factor involved, these methods are not usable to represent interactions between different factors such as the interaction between stations and railway services.

The described methods are applicable to study dwell times but are limited to either the performance of a station or that of a service, thus ignoring the possible interaction between both. This interaction is important since it is possible that only a few services cause a station to perform poorly in terms of dwell time, and vice versa. To understand this, there is a need for a tool which can take both station and service performance into account simultaneously. The method proposed in Paper V is the use of a Rasch analysis technique, a tool more commonly used in questionnaire studies. The results from Paper V suggest that the Rasch analysis technique can also be used within an operational context to study dwell times on a line level. The model output also shows that the Rasch analysis technique can adequately reflect the expected variability in dwell times for both railway services and stations.

Comparing the output from the Rasch analysis technique to the median delay, the results in Paper V show that the Rasch analysis allows to better identify which services are likely to incur a dwell time delay at a given station. By having access to such information, planners can make more informed decisions when making adjustments to reduce the likelihood of dwell time delays. It is, for example, possible that the output from a Rasch analysis reveals that adjusting the scheduled dwell time for a few services halting at a specific station can go a long way, rather than scheduling longer dwell times for all services halting at that specific station. Although methods such as the Rasch analysis technique are more difficult to implement compared to measuring the average dwell time delay or recording the frequency of delays, the additional insights allow for a more accurate identification of problematic situations in terms of dwell time delays.

Robust regression analyses on dwell times

When studying dwell time delays it is not just interesting to know the size of a delay, but also which factors have an effect on the duration of dwell times. This is commonly done by making use of regression models, see for example the work by (Coulaud et al., 2023; Palmqvist et al., 2020; Pritchard et al., 2021) which provide insights into the conditional mean of dwell times given a set of explanatory variables. Although the conditional mean is of interest, it does not provide information on the distribution of dwell times and suffers from the same limitations as the measures of central tendency described before. Furthermore, and more importantly, dwell time delays are likely to violate the normality assumptions associated with these regression approaches. The distribution of dwell times can be characterised by having a substantial amount of smaller delays and very few larger dwell time delays. The same is often true regarding passenger volumes where smaller volumes are much more common than large volumes of passengers.

Several different approaches exist to overcome this non-normality issue such as the use of non-parametric regression models. However, the outcome of non-parametric regression models is difficult to interpret. This interpretability is important when studying the influence of different factors on dwell times, rather than when the aim is to predict dwell times. A quantile regression model was proposed in Paper VI, to overcome the non-normality issue whilst still retaining the interpretability of the outcome. Quantile regression is also a useful approach when the conditional distribution of the dependent variable is of interest as it allows to study the relationship between dependent and independent variables outside of the mean (Kourtit et al., 2022). Furthermore, quantile regression is also better adept at handling outliers in the data (Hao & Naiman, 2007) and at handling heteroscedasticity (Waldmann, 2018).

The results in Paper VI indicate that the use of quantile regression for dwell time research is justifiable, as the results show different effects of the explanatory variables across the different percentiles, indicating that the assumptions of ordinary least square regression are not met (Staffa et al., 2019). Furthermore, the coefficients for the variables included in the model were found to change throughout the distribution, differing significantly from the ordinary least square coefficients in most cases. This indicates that using an ordinary least square regression model only provides limited information regarding the effect that the explanatory variables chosen in Paper VI have on dwell times. These findings suggest that more robust regression methods are more fitting when studying dwell times.

The importance of effect sizes

The final point worth making with regard to the use of large automatic passenger count data sets is related to the reporting of statistical inferences. As mentioned by Palmqvist (2019), researchers can study delays and delay causes with increasing levels of detail as a result of the growth of the volume of data that is being collected within the railway system. Automatic passenger count systems are an example of such improvements in data accessibility. Although having access to a large amount of data has several benefits, it is also possible to fall prey to the *large sample size fallacy*. In such cases, statistical significance is a result of having a large sample size but the findings have limited practical relevance (Lantz, 2013). This is where effect sizes come in. Effect sizes provide a measure of practical significance alongside the measure of statistical significance, and more importantly, effect sizes are not affected by large sample sizes (Lantz, 2013).

The practical significance of findings is especially relevant for the studies conducted in the research presented here since there is a focus on the real-world impact of the findings. As the discussion on the findings from the studies will show, the large sample size fallacy would have been present in the work presented here as well if effect sizes were not calculated. It is in general recommended to report effect sizes as they provide additional information such as the magnitude of a difference or association (Brydges, 2019; Wilkinson & Task Force on Statistical Inference, 1999). Furthermore, knowledge about effect sizes is an important aspect when planning a study (Lakens, 2013). When reporting effect sizes this not only guards researchers against making claims that hold little to no practical relevance, it also provides important information for the relevant field of research.

Research question 4

Which type of measures can be taken to improve dwell time punctuality for commuter trains?
Paper I, II, III, IV, V

The fourth research question aims to define different types of measures which can be taken to improve dwell time punctuality for commuter trains. Based on the literature and the studies conducted as part of this thesis three distinct avenues to improve dwell time punctuality have been identified. The first is a different approach to scheduling dwell times, where the current static approach to dwell time scheduling should be abandoned in favour of a more dynamic scheduling approach. The second avenue is to apply platform management measures to influence the behaviour of passengers in such a way that it has less of an impact on the duration and variability of dwell times. The third avenue that can be identified is that dwell time scheduling should take a more network-wide approach rather than treating stations as a single entity.

A different approach to scheduling dwell times

One way to improve dwell time punctuality according to Nash et al. (2006) is to conduct a complete and systematic revision of dwell times. A similar point can be made based on the findings presented in this thesis, where the current approach to dwell time scheduling has been found to lead to actual dwell times exceeding the scheduled times on a regular basis. Given the large amount of dwell time delays it is clear that the dwell times that are currently scheduled are too short. A straightforward approach to solve this issue is to schedule longer dwell times.

Although this comes at the cost of longer travel times for passengers, past studies have indicated that punctuality has been found to be more important than total travel time (Parbo et al., 2016; Yabuki et al., 2017). Furthermore, it is worth pointing out that these changes to dwell times do not need to be large. The findings in both Paper II and Paper IV indicate that most dwell time delays are relatively small, often being less than 30 seconds, and small adjustments to the scheduled dwell time can thus go a long way towards reducing dwell time delays.

This being said, extending dwell times is likely not the best way forward. Firstly, it can actually result in the dwelling process taking more time than what it does currently. This is due to a behavioural response to having additional time where more time is taken for a task if more time is given to said task (Carey, 1998). Furthermore, scheduling more dwell time could lead to unwanted behaviour such as an increase in door holding (Pritchard et al., 2021; Volovski et al., 2021). Extending dwell times can also lead to issues regarding the capacity of a railway system.

Longer dwell times reduce the available capacity since trains occupy stations longer and one should thus be thoughtful when scheduling longer dwell times. It is also worth pointing out that extending dwell times is likely not beneficial at all stations or across an entire day, leading to trains dwelling for an unnecessarily long time in some cases.

A different, and arguably more fitting approach is, therefore, to adopt a dynamic approach towards scheduling dwell times. The results presented in Paper V, using a Rasch analysis, show that there is a case to be made that extending dwell times at all stations for all trains is indeed not necessary. Instead, specific problematic stations or services can be targeted when scheduling dwell times rather than the one-size-fits-all approach currently used when scheduling dwell times. Doing so will extend dwell time there where it is needed, without increasing dwell times for cases where the current dwell time is actually sufficient. This approach does require continuous monitoring to follow up on changes made and see if they were effective or not. Whilst this increases the workload regarding the scheduling of dwell times, it will likely help to reduce dwell time delays in the long run.

The case for platform management

The findings presented in this thesis suggest that the dwelling process should not only be seen as an engineering problem but the behaviour of passengers should also be taken into account. On the one hand, this could be done by reducing the passenger volumes during peak hours, flattening the curve, to reduce the risk of a dwell time delay. However, the results from Paper IV suggest that this alone will not be sufficient. Based on the findings presented here, a strong case can be made in favour of platform management instead, actively changing the behaviour of passengers during the boarding and alighting process. As the logic behind the Swiss cheese model for passengers induced dwell time delays suggests, having passengers spread out evenly and queue next to the doors will likely reduce the risk of dwell time delays, even at higher passenger volumes.

Platform management consists of measures taken on the platform to change the behaviour of passengers. Examples of this are measures to spread out boarding passengers more evenly, reducing the level of concentrated boarding and the impact of the critical door, as well as measures aimed at instructing passengers so that queues should not be formed in front of the doors of a train. These kinds of measures can be implemented in different ways, be it the use of markings, instructions by staff or physical measures such as the use of platform gates or markings.

Although platform management can have the potential to reduce dwell time delays, not every station will benefit from such kind of measures. One way to identify stations where platform management measures can be beneficial is by making use of the expected volume of passengers. As mentioned in the answer to research question 2, the volume of passengers acts as the accelerant for the negative impact of the spread of boarding passengers. Platform management measures should, therefore, be introduced at stations with a high expected number of passengers, as it is those stations where real benefits can be gained.

A network-wide approach to dwell times

Both of the previous approaches are aimed at solving the problem of dwell time delays within the context of a single station. Such an approach does, however, not account for interdependencies between different stations and between stations and services. As the findings from Paper I indicate, for example, there is an interdependency between stations which influences the spread of boarding passengers. When scheduling dwell times or introducing measures to reduce the impact of passengers on dwell times it is important to take these interdependencies into account. This is in line with statements made by Nash et al. (2006) who argue that it is important to coordinate the implementations of measures to reduce dwell time delays since uncoordinated measures will likely be less effective in the long term.

In cases where a certain train service is often delayed at specific stations this can be a result of something occurring at a station upstream, and scheduling more dwell time at the problematic station might not address the actual cause for dwell time delays. Instead, it is important to understand the way stations are connected to uncover the underlying causes of dwell time delays. Making a change to the halting position at one station can, for example, help to reduce the chance of a dwell time delay at a downstream station.

10 Reflection on the data and methods used

10.1 The use of automatic passenger count data

The thesis presented here relied heavily on the use of automatic passenger count data. The choice to make use of automatic passenger count data was guided by its availability, as well as by the features present in the data. First, in addition to information on the volume of passengers, the on-board systems used to collect the passenger count data provided the actual dwell time on a magnitude of seconds. As previously shown in *Section 6.4.1* having data on a level of seconds rather than minutes is important when studying dwell times. In addition to this, the automatic passenger count data provides a large amount of data to work with. In the case of the data from Stockholm, there is information for an average of around 200,000 station stops per year, and in the case of data from Skåne, this was an average of around 1.4 million station stops per year. The passenger count data used here also covers a rather large time span, this being six and four years in Stockholm and Skåne respectively, and a large geographical spread since the entire network is included in both datasets. Having access to such a large amount of data meant that different phenomena could be studied across a wide range of conditions. Some further reflections on the automatic passenger count data are presented below.

10.1.1 Dwell times from the on-board systems

Given that the focus in the work presented here is placed on the effect of passengers on dwell times it is important to reflect on what the measured dwell time does and does not include. In the case of the data from trains in Skåne, the dwell time is measured based on the time between the doors unlocking and locking. The dwell time provided by the on-board system thus reflects the boarding and alighting time. In contrast to this, the data collected on board trains in Stockholm is measured based on the wheel-stop and wheel-start time. In this case, the end of the dwell time is registered upon the departure of a train and thus includes both the boarding and alighting time as well as the dispatching time, measuring the total dwell time.

Since the data collected in Stockholm also contains the dispatching time, it is possible that the effect of passengers is somewhat overestimated in Paper II. This is because a dwell time delay can also be caused by a slow dispatch rather than by passengers in these cases. It was, however, not possible to examine the dispatching time, and this effect could thus not be corrected in the data. The dispatching time is not included in the data collected in Skåne since the end of the dwell time is registered when the doors lock, thus before the train departs. This makes the data originating from trains in Skåne more accurate in light of the studies conducted in this thesis. It is worth noting that a similar trend in terms of dwell time delay sizes was found between both datasets, suggesting that the inclusion of the dispatching time did not lead to any major problems.

10.1.2 Passenger counts from the on-board systems

Some reflections can be made regarding the passenger count part of the data. The first is how accurately the systems count passengers. With regards to the data collected on board trains in Skåne, a detection accuracy of 98% is indicated by the provider of the automatic passenger count data. The exact accuracy of the system used in Stockholm was not disclosed, but the provider of the data mentioned that the counts from the on-board system are corrected based on historical passenger counts when necessary. Using this historical data ensures the accuracy of the data according to the provider of the automatic passenger count data. In order to exclude situations where an unrealistically large number of passengers was counted, an upper limit for the volume of both boarding and alighting passengers was used during the data handling procedures. The upper limit was based on suggestions from the data providers.

The next aspect worth reflecting on is data aggregation. The system used in Stockholm provides data on a carriage level, whereas the data collected in Skåne provides passenger counts on a door-by-door level. Where the former meant that it was only possible to study the effect of passenger volumes on dwell time, the door level data also allowed for studies into the spread of boarding passengers and to identify a critical door. As a result of this difference, no comparison was made between Stockholm and Skåne since the findings presented here suggest that information on the spread of passengers is required when studying dwell times, and only comparing the effect of the volume of passengers between both regions was deemed to be less valuable.

It is also worth reflecting on the number of carriages equipped with an automatic passenger count system, as this has an influence on what is measured in terms of the sample space. All carriages used in Skåne are equipped with an automatic passenger count system, providing data on the entire population. This is not the case in Stockholm, where every seventh carriage is equipped with an automatic passenger count system and the system thus provides a sample of the population.

The limited number of carriages equipped with a passenger count system also causes a challenge in terms of interpreting the data. Trains in Stockholm can be composed of multiple carriages and the number of carriages equipped with an automatic passenger count system can thus vary per train. In practice, this means that a single train can provide either none, one, or multiple passenger counts depending on its composition. When taking the sum of passengers this can lead to situations where some stops involve an extreme volume of passengers compared to other stops, due to there being multiple carriages present that are equipped with an automatic passenger count system. To avoid such errors, the average of the volume of passengers was used instead. This does come with a risk of effectively excluding stops with a high volume of passengers. If a train consists of two carriages, one with 50 and one with 100 passengers, the new total would be 75 passengers rather than 150 passengers. At the same time, this approach does still take the dwell time associated with this observation into account thus associating it with a lower passenger volume. This being said, the trend described in Paper II is still valid, and would likely be even more prominent if the sum of passengers was used instead.

10.1.3 Counts are not observations

Whilst the automatic passenger count data is a rich source of data, it only provides passenger counts and does not provide observations on the boarding and alighting process itself. This means that the inferences that can be made are based on passenger counts and not observations, and effects such as passenger characteristics and late arriving passengers have not been identified in, or excluded from, the data. Although having such information is valuable, a trade-off was made between having a large data set on many services and stations as is the case with passenger count data, and having in-depth information on the dwelling process at a limited number of stations as would have been the case when making use of observations. The choice was made in favour of the former, partly due to the availability of the data and partly since observational studies have been conducted in the past and were limited in their sample sizes. The work presented here, on the other hand, provides insights into the dwelling process over a long period of time for a large number of stations and services. This meant that it was possible to study the effect of previously highlighted phenomena, such as the impact of concentrated boarding on dwell times, across a wider range of conditions than what was previously done. This would have not been feasible when making use of observational data.

10.2 The choice for robust analyses

The various analyses performed as part of the thesis make use of so-called robust statistics, a choice that was primarily guided by the dwell time data not having a normal distribution. Most common parametric tests, such as a *student's t-test* for example, rely on a normal distribution of the data, and test for statistically significant differences between the mean of two groups. However, as stated by Rousseeuw (1991), very few distributions that occur in practice have a perfectly Gaussian distribution and real-world data thus frequently departs from the assumptions. The same is true for the data on passenger volumes and dwell times used here.

To illustrate why it was important to make use of robust statistical approaches, it is possible to take a look at the use of the median in favour of the mean. A common measure of central tendency is that of the mean, or average, value (Khorana et al., 2022), although easy to interpret the mean can lead to misinterpretations when outliers are present in the data. Inspired by an example provided by Rousseeuw (1991), the problem can be explained as follows.

Take the examples below with five trains dwelling at two stations with the following volume of boarding passengers:

30 – 35 – 28 – 31 – 80

30 – 35 – 28 – 31 – 5

The first four values are the same in both examples, with only the *fifth value* being changed. In the first case, the mean number of boarding passengers is 41, which is substantially higher than most observations made. In the second case, the mean volume of boarding passengers is 26. The median volume of boarding passengers, on the other hand, is 31 and 30 in the first and second cases respectively and much closer to the actual volume of passengers that is observed.

Furthermore, even though the only difference between both stations presented here is the fifth value, the mean volume of boarding passengers shifts from 41 to 26, a difference of 15 passengers. The median value, on the other hand, is much more stable, although it is still affected by the extreme fifth value, the change between both examples is only 1 passenger.

The impact of these kinds of extreme values and the presence of skewness in the data can pose problems when performing statistical analyses. Acknowledging the potential issues regarding the non-normality of the data, the quantitative analyses made use of non-parametric tests. Another name for such tests is *distribution-free statistical procedures*, indicating that there are no assumptions made about the distribution of the data (Scheff, 2016). Non-parametric tests are less powerful compared to their parametric counterparts when the assumptions for the latter are met and require larger sample sizes to achieve statistical significance (King & Eckersley, 2019; Ostertagová et al., 2014). However, when these assumptions are not met, as is the case with the data used here, non-parametric tests should be used (Ostertagová et al., 2014) especially since the use of parametric tests could cause misleading results (King & Eckersley, 2019).

10.3 The effect sizes used

When using large sample sizes, one can fall prey to the *large sample size fallacy*. In such cases, statistically significant findings are a result of having a large sample size but those findings have limited practical relevance (Lantz, 2013). An important decision when making use of effect sizes is specifying what small, medium, and large effects. The studies presented in this thesis make use of the suggestions provided by Cohen (1988) for the labels of the effect sizes. Using these suggested effect sizes is not the best approach, however, since these are guidelines that should only be used when domain-specific effect sizes are not available (Brydges, 2019; Cohen, 1988; Correll et al., 2020). As argued by Correll et al. (2020) the definitions of small, medium, and large effect sizes provided by Cohen are arbitrary and inconsistent and these definitions were proposed so that a researcher could make an informed judgment in the absence of other data.

Ideally, the effect sizes used would thus have been based on previously reported effect sizes within the domain of railways or dwell times. However, effect sizes are often not reported within the domain of dwell time research. This means that it was necessary to make use of the suggestions made by Cohen (1988), as these were considered to be the best option given the lack of alternative ways to construct effect sizes. In practice, this means that what is considered to be a small effect might actually be large or vice versa.

11 Contributions of the thesis

11.1 Contributions to research

The contributions to research from this thesis are related to how dwell time delays can be studied using large automatic passenger count data sets. The key takeaways are the identification of the need for robust statistical approaches, the need to report effect sizes, and the need to describe dwell time delays in terms of frequencies rather than a measure of central tendency. Although the datasets used in this thesis are somewhat unique in both their size and level of detail, similar large datasets regarding passenger volumes and operational aspects might become more commonplace within the domain of railway research. When this is the case, it is important to understand the limitations of the data and potential pitfalls that arise as a result of the volume of data. Non-parametric testing might become the norm as real-world data tends to take on a non-normal distribution, for example. Furthermore, statistical significance becomes more likely when sample sizes increase, and only reporting the result of statistical tests can lead to an overestimation or misrepresentation of the importance and practical relevance of findings. It is, therefore, important to report effect sizes along with the results of statistical tests. In fact, this notion can be extended beyond just dwell time research into the field of railway research and transportation research in general.

In addition to this, the work presented here shows the value of having data on more than just the volume of passengers when studying dwell times. The volume of passengers was found to not be the main cause for dwell time delays but act as an accelerator. When studying dwell times it is thus not sufficient to just focus on the volumes of passengers, but it is important to also include other aspects such as the spread of boarding passengers, and the ratio between boarding and alighting passengers when possible. When such data is not available, this limitation should be acknowledged during the analysis process.

11.2 Contributions to identifying dwell time delays

As the Dutch saying goes "meten is weten" which translates into measuring is knowing, the same is true for dwell time delays. However, the current way to measure punctuality obscures the presence of dwell time delays. With dwell time delays not being accurately measured, the magnitude of the problem of dwell time delays is likely to be unknown. Three aspects are important to improve upon in order to better measure dwell time punctuality and with it identify where dwell time delays occur. These three aspects are the location, measurement scale, and the way in which the data is aggregated.

The findings presented in this thesis point towards differences in terms of dwell time performance between different trains halting at the same station, and between the same trains halting at different stations. In order to better identify dwell time delays it is, therefore, important to measure dwell time punctuality at each station rather than just at a selection of stations or the final station. In addition to this, the dwell time performance should ideally be measured for each individual train service to better understand when dwell time delays occur.

When measuring dwell times it is essential to use a sufficient level of detail in the measurement scale. Train movement data is often used to determine the punctuality of trains. This data is aggregated on a minute-by-minute level. Measuring punctuality on a scale of minutes means that a large portion of dwell time delays are actually not measured, as the overview in Table 6 (*Section 6.4*) shows. Having data on a level of seconds, on the other hand, allows for a more granular analysis and makes it possible to identify even small dwell time delays and with it measure dwell time punctuality more accurately.

Accurately measuring dwell times is only half of the solution though, the next challenge is how to accurately reflect dwell time delays. To do so, it is important to have the right level of data aggregation. The current binary approach to punctuality will show whether a train is delayed or not, without providing insight into the size of the delay. Dwell time delays should ideally be represented in terms of their size and frequency instead. Doing so will not only show how large the dwell time delays are but also how often these delays occur.

To summarize, in order to better identify dwell time delays there is a need to make use of data on a scale of seconds to allow for a sufficient level of detail in the observations. Dwell time delays should then be measured based on both their size and frequency. These measurements should ideally include all stations within a network as well as the different services that are operated in order to identify where dwell time delays pose a problematic situation. These analyses should ideally take place periodically and be formally incorporated into the process of timetabling. Doing so allows for a more case-specific approach to dwell time scheduling which will be beneficial to help reduce dwell time delays.

11.3 Policy implications

Although there is no silver bullet for the question of how much dwell time should be scheduled, several points of attention can be raised to guide the process towards scheduling dwell times that are more aligned with the real-world situation. Working on these points can benefit the decision-making process for scheduling dwell times and will ultimately help improve dwell time punctuality for commuter trains. It is worth pointing out that even though the research presented in this thesis was conducted within a Swedish context, the points raised here can be applicable in other countries as well.

11.3.1 A more dynamic approach to dwell time scheduling

One aspect to include in the process of scheduling dwell times is to allow for a more dynamic approach. Currently, dwell times in Sweden are scheduled using a static approach where the same dwell time is applied to multiple stations and no difference is made between peak and off-peak hours. This approach assumes that the necessary dwell time is stable. The findings in this thesis suggest that this is likely not the case since there are differences in passenger flows between peak and off-peak hours as well as between stations. These differences are currently not captured. To better align the scheduled dwell time with the actual dwell time, a more dynamic approach is necessary where different dwell times are scheduled between peak and off-peak hours and more importantly between different stations and between different services halting at the same station. It should be acknowledged that this change increases the workload when designing a timetable, but it is likely to have large benefits in terms of the on-time performance of railways.

An argument in favour of the current way of scheduling dwell times is that it allows for a cyclical timetable, something that is preferred as it makes the timetable easier to understand for passengers (Robenek et al., 2016). A more dynamic approach to dwell time scheduling would mean that it is not always possible to have a cyclical timetable. This should, however, not be as much of a problem in practice since a majority of dwell time delays are rather small, being thirty seconds or less. Such a small change can be included in the operational timetable without presenting this to passengers. The train scheduled to depart at 9:01 can still be shown to depart at 09:01:00 but can be scheduled to actually depart at 09:01:15, for example.

11.3.2 Making use of the available data

In order to implement a more dynamic approach to dwell time scheduling it is important to make use of the available data, such as the highly detailed passenger count data used throughout this thesis. Despite this data being available it is currently not sufficiently used, and dwell time delays can remain hidden in punctuality statistics. A rather straightforward solution for this is to actually measure dwell time punctuality, using the available data. As mentioned in the answer to Research question 3, this would ideally be done by making use of the frequency of different delay sizes rather than a single threshold as is the case for the punctuality statistics currently in use. Adopting this approach can be the first step towards reducing dwell time delays since the magnitude of the problem will become visible.

11.3.3 Understanding and adapting (to) passenger behaviour

The findings presented in this thesis point towards the importance of the behaviour of passengers for dwell times. This means that the dwelling process should not be viewed as a technical problem that can be fixed by timetable or rolling stock based interventions. Instead, the behaviour of passengers during the boarding and alighting process, such as the formation of queues, should also be taken into account when introducing measures to reduce dwell time delays.

In addition to this, more effort should be placed on understanding how passengers move through a railway network, and what the implications for dwell times are of this behaviour. Having such an understanding will allow railway operators to adapt the behaviour of passengers to fit better with the system, or vice versa. A concrete example of this would be changes to the stopping position of trains in such a way that the doors close to the entrance are not always lined up with carriages that are already quite full given the boarding behaviour at previous stations.

11.3.4 An ideal way to collect passenger flow data

Although this thesis does not necessarily aim to show the ideal way in which passenger flow data should be collected it is still worth elaborating on this subject. The studies presented here relied heavily on the availability of automatic passenger count data. Based on this experience, it is fair to say that an automatic way to collect data is of great importance to ensure sufficiently large sample sizes to conduct in-depth studies on dwell times in a real-world setting. Although the passenger count data used here has a high level of detail, improvements can be made.

In addition to the number of boarding and alighting passengers per door, automatic passenger count data should ideally also include timestamps on when these passengers board or alight the train. Having access to information on when passengers board makes it possible to identify the effect of late-arriving passengers on dwell times.

Having information on the exact boarding and alighting times of passengers also makes it possible to study the concept of *minimum dwell time* and dwell time margins, as described in *Section 1.3*. It is currently necessary to make some concessions to measure minimum dwell times, such as only making use of delayed trains as done by Pedersen et al. (2018). This not only limits the available data, it also requires the train to depart as soon as the boarding is completed for this assumption to hold. Studying the minimum dwell time is relevant as it sheds light on dwell time margins and capacity assessments of networks, something which will become more important as networks become more heavily utilized.

It is also worthwhile to collect detailed origin-destination data alongside passenger count data, as this would make it possible to study carriage choice in greater detail. Although some studies delved into this, see for example the work done by Fang et al. (2019) and Peftitsi et al. (2020), these studies were limited to using data aggregated on a carriage level. Given this limitation, these studies could not study the door choice. On the other hand, the study presented in Paper III had no access to detailed origin-destination data, meaning that it was not possible to study the spread of boarding passengers in relation to travel behaviour. Having access to both origin-destination data and door-by-door passenger count data would bring together two valuable data sources.

To conclude, operators have the opportunity to collect an incredible amount of highly detailed data when the necessary systems are put in place. These systems should be considered when acquiring a new system or trainset. Retrofitting the necessary technology can become a rather costly exercise, whereas the costs are likely to be relatively low when part of a larger investment into new rolling stock. Data collection methods should, therefore, be taken into account in addition to other criteria when buying new, or updating existing, rolling stock. Having access to highly detailed information will allow for a more in-depth study of passenger flows within a railway network. This is not only relevant for dwell time research but can be beneficial for a wider area of transportation research as well, for both academia and practitioners.

12 Limitations

As with many things in life, this thesis comes with some limitations. The main limitations are related to the data that was used and the data that was not available or collected during the completion of this thesis. The main limitation of the passenger count data, it being counts and not observations, has been shortly addressed in *Section 10.1.3* but this deserves further elaboration. As mentioned before, the count data only provides the sum of passengers and does not provide information on when passengers board or alight. This can be problematic since late arriving passengers can have a noticeable effect on the length of dwell times. Take for example a situation where the scheduled dwell time is 60 but the actual dwell time is 75 seconds, and a total of 50 boarding passengers are registered. Two possible ways in which this number of passengers for this dwell time delay can be observed are shown in Figure 21.

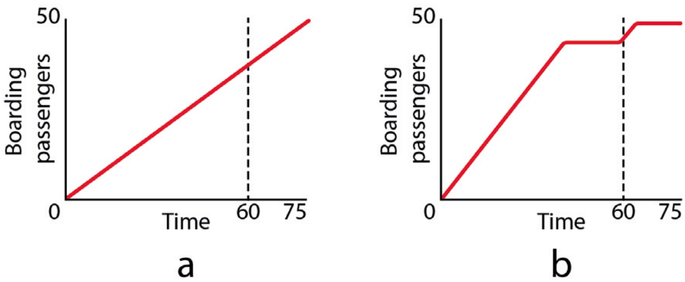

Figure 21: Examples of different ways in which passenger volumes can be collected. The dashed line shows the scheduled dwell time.

Figure 21a shows a continuous flow of passengers that results in exceeding the scheduled dwell time. Here it is possible to state that the boarding process took too long and this caused a dwell time delay. Another possible situation is shown in Figure 21b where the initial cluster of boarding passengers is handled within the scheduled time, but a late arriving passenger causes a dwell time delay to happen. Both situations show up the same in the datasets used in this study, whilst the actual cause for the dwell time delay is not the same. This is a limitation which is a result of the system used by the operator of the trains to collect the data and could thus not be addressed within the scope of this thesis.

Another limitation concerning data access is the lack of data on dwell time margins. Planners can include margins in the scheduled dwell time, scheduling more time than needed to allow trains to make up time and reduce their delay. There is, however, no formal way to schedule such dwell time margins in Sweden and as a result of this the presence of these margins is somewhat hidden in the timetable. Not having information on the presence and size of dwell time margins can result in an underestimation of dwell time delays, where some of the effects found in the work presented here might be stronger than reported.

There are also some limitations concerning the data that has not been included or collected. Although the layout of a station, both in terms of the platform and the station area was highlighted as being influential, such information was not collected during the work completed as part of this thesis. The absence of this information means that some assumptions had to be made regarding station-specific characteristics to capture this effect, which resulted in including the historical dwell time of stations in Paper V, for example. If information on the layout of platforms and information on the station area was present such assumptions would not have been necessary and the effect of station-specific characteristics could be studied with more detail. The absence of this information is the result of a prioritization of other avenues of research which result in the work presented in both Paper IV and Paper V.

Along a similar line of thought, it is worth discussing the absence of qualitative data such as questionnaires or interviews in the studies presented in this thesis. It was deemed to be unnecessary to perform an interview study into planning principles given the work presented in Palmqvist et al. (2017) in which comments regarding dwell time scheduling were made. These scheduling principles have likely not undergone drastic changes between the writing of this thesis and the publication of the paper by Palmqvist et al. (2017). Evidence of this is the rather stagnant scheduled dwell time where dwell times are similar over a period between 2012 and 2020.

Although questionnaire studies and observational studies on the behaviour of passengers on the platform were planned early on in this PhD journey and could have provided valuable additional information, such studies were not conducted in the end. The main reason for this was the COVID-19 pandemic, which not only meant that a significant portion of the time spent completing this PhD was spent working from home, but it also meant that the way in which passengers behave on platforms was likely to be different from what is considered to be the normal situation. Collecting qualitative data during this period would have been of interest within the context of COVID-19, but not necessarily reflect normal conditions.

13 Future research

Given the limitations previously described, there is naturally also room for future research and improvements. Three main themes can be identified within which future research could take place, these being a network and planning theme, a behavioural theme, and an unplanned event theme.

Regarding the first, the network and planning theme, two concrete ideas can be formulated. The first proposal falls under the planning side, where the proposed dynamic scheduling of dwell times can be studied. This can be done either within a simulated environment, be it a micro/macro simulation or mathematical simulation model, or in a real-world scenario. The second line of research is related to understanding the interdependencies of stations in relation to the critical door of a train. This line of work is inspired by the work presented in Fiebag (2019) where hotspots of boarding and alighting passengers were identified within a subway network in Germany. Such a line of research should include the effect of the area in which a station is located, such as the direction to a city centre or connecting transport modes. Knowing where hotspots of boarding and alighting passengers overlap can guide interventions such as changes to the halting position of a train.

As mentioned in the limitations, the studies presented in this thesis did not make use of qualitative data such as observations of passenger behaviour. This leaves room for future studies, such as studies into the behaviour of passengers during the boarding and alighting process. Furthermore, it is of interest to study the effect of the proposed platform management measures. This behavioural theme can also include the behaviour of train staff, both drivers and train attendants and their role in dwell time delays. A potential avenue here is to study the driving behaviour of train drivers to understand the impact of the train slowing down and speeding up again on dwell times.

The third theme of possible future research is that of studying unplanned events such as door failures and platform changes. These unplanned events can have potentially large effects on dwell times. For example, a broken door means that the boarding passengers are spread between fewer doors and can cause the boarding and alighting process to take much longer at the doors nearest to the one that is broken. Furthermore, passengers potentially require additional time between realizing a door is broken and choosing a new door to board through. The impact of this on dwell time is, however, not well studied. Knowing the impact of broken doors on dwell times can help to inform rolling stock maintenance schedules, for example, by identifying the importance of making sure that all the doors of a train operate as intended. Another unplanned event is that of a last-minute platform change. This requires passengers to move to a different platform and can result in passengers making use of the doors closest to the platform entrance to board the train as a result of stress due to the fear of missing the train, a phenomenon which has been previously identified. In such a case the platform change induces a degree of concentrated boarding, which will likely extend dwell times. No research on this phenomenon has been identified during the work done in this thesis, however. This whilst understanding the impact of track changes on dwell times can help inform dispatchers during real-time rescheduling tasks by raising awareness of what an unplanned platform change does to the dwell time.

14 Conclusion

The overarching aim of the thesis presented here is to develop knowledge of how dwell time delays arise in order to identify and describe potential ways in which dwell time delays can be reduced. In addition to this, a secondary aim is to identify how dwell times can be studied on a network-wide level. The following research questions have been answered to achieve these goals: 1) What are the causes of dwell time delays for commuter trains? 2) How do boarding and alighting passengers influence the duration of dwell times for commuter trains? 3) How can dwell time delays be studied in a network with automatic passenger count data? 4) Which type of measures can be taken to reduce dwell time delays for commuter trains?

The main findings are as follows. The dwell time process is complex and different causes for dwell time delays were identified. A consensus can be found in the literature that the presence of passengers is one of the major causes of dwell time delays. Although the volume of passengers is often regarded as the main way in which passengers have an effect on dwell times, the findings presented in this thesis indicate that the volume of passengers instead acts as an accelerator for the negative impact of other aspects, such as the behaviour of passengers. The risk of dwell time delays is thus not driven by the volume of passengers alone but by a combination of unfavourable factors such as a high volume of passengers in combination with an uneven spread of boarding passengers.

Different ways in which dwell time delays can be studied in a network with automatic passenger count data have been presented. In order to better highlight the problem of dwell time delays it is important to measure the actual dwell times across all stations and services. Dwell time delays should be represented in terms of their size and frequency to gain a better understanding of the magnitude of the problem, not only revealing how large delays are but also how often they occur. Furthermore, methods such as a Rasch analysis can be introduced to study dwell time performance for stations and services in a single dimension. Having such insights can help identify when and where dwell time delays arise, allowing for better informed measures to be taken to reduce dwell time delays.

In general, it can be stated that there is no silver bullet to avoid dwell time delays, and it is important to better understand the nature of dwell times and dwell time delays. This can be achieved by gaining a systematic understanding of the dwelling process. One aspect of this is the need for better delay indexes for dwell times and a systematic and formal monitoring of dwell time performances across stations and services. Doing so can help to point towards hot spots where delays are more likely to happen and efforts to reduce dwell time delays can make a real impact. This should ideally lead to a more dynamic approach to dwell time scheduling where dwell times are tailored to specific stations and specific services. Although in some cases this means that dwell times need to be extended, and thus extend travel times, the findings in this thesis suggest that small changes in the dwell time can go a long way so the actual impact will be limited.

The solutions towards reducing dwell time delays should, however, not be limited to a timetable or engineering scope. Adding more dwell time will likely not be the answer to solving dwell time delays. Furthermore, engineering solutions such as widening doors will only be beneficial when passengers queue next to these doors. Instead, the behaviour of passengers both on the platform and the way they travel through the network should be considered and studied as well. Having a better understanding of how often and where dwell time delays occur and the behaviour of passengers can help make informed decisions that have the potential to make a real difference and result in fewer dwell time delays. It is this combination of scheduling dwell times more dynamically and accounting for the users in the system that will have the biggest potential to make a real difference in terms of dwell time punctuality.

15 References

Ahn, S., Kim, J., Bekti, A., & Cheng, L. (2016). Real-time Information System for Spreading Rail Passengers across Train Carriages: Agent-based Simulation Study. *Australasian Transport Research Forum 2016*.

Annema, J. anne. (2009). Weerstanden van verplaatsingen: Tijd, kosten en moeite. In *Verkeer en Vervoer in hoofdlijnen* (pp. 107–126). Coutinho.

Antognoli, M. A., Girolami, F., Ricci, S., & Rizzetto, L. (2018). Effect of passengers' flows on regularity of metro services: Case studies of Rome lines A and B. *International Journal of Transport Development and Integration*, *2*(1), 1–10. https://doi.org/10.2495/TDI-V2-N1-1-10

Atkins. (2004). *Significant steps—Technical analysis*. Department for Transport.

Baee, S., Eshghi, F., Hashemi, S. M., & Moienfar, R. (2012). Passenger boarding/alighting management in urban rail transportation. *2012 Joint Rail Conference, JRC 2012*, *September 2014*, 823–829. https://doi.org/10.1115/JRC2012-74102

Belaïd, F., Youssef, A. B., & Lazaric, N. (2020). Scrutinizing the direct rebound effect for French households using quantile regression and data from an original survey. *Ecological Economics*, *176*, 106755. https://doi.org/10.1016/j.ecolecon.2020.106755

Black, D., Kolesnikova, N., & Taylor, L. J. (2012). Why Do So Few Women Work in New York (And So Many in Minneapolis)? Labor Supply of Married Women across U.S. Cities. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1129982

Bosina, E., Meeder, M., & Weidmann, U. (2017). Pedestrian flows on railway platforms. *17th Swiss Transport Research Conference*, 24.

Bosina, E., Meeder, M., Weidmann, U., & Britschgi, S. (2015). *Distribution of passengers on railway platforms*. 22. http://www.strc.ch/conferences/2015/Bosina_EtAl.pdf

Brons, M. R. E., & Rietveld, P. (2008). *Rail mode, access mode and station choice: The impact of travel time unreliability* [Research Report for the project TRANSUMO BTK]. VU University.

Brown, C. G., McGuire, D. B., Beck, S. L., Peterson, D. E., & Mooney, K. H. (2007). Visual Graphical Analysis: A Technique to Investigate Symptom Trajectories Over Time. *Nursing Research*, *56*(3), 195–201. https://doi.org/10.1097/01.NNR.0000270029.82736.5a

Brydges, C. R. (2019). Effect Size Guidelines, Sample Size Calculations, and Statistical Power in Gerontology. *Innovation in Aging*, *3*(4), igz036. https://doi.org/10.1093/geroni/igz036

Buchmueller, S., Weidmann, U., & Nash, A. (2008). Development of a dwell time calculation model for timetable planning. *WIT Transactions on the Built Environment*, *103*, 525–534. https://doi.org/10.2495/CR080511

Caimi, G., Kroon, L., & Liebchen, C. (2017). Models for railway timetable optimization: Applicability and applications in practice. *Journal of Rail Transport Planning & Management*, *6*(4), 285–312. https://doi.org/10.1016/j.jrtpm.2016.11.002

Cambridge Dictionary. (n.d.). *Punctuality*. Retrieved July 11, 2023, from https://dictionary.cambridge.org/dictionary/english/punctuality

Carey, M. (1998). Optimizing scheduled times, allowing for behavioural response. *Transportation Research Part B: Methodological*, *32*(5), 329–342. https://doi.org/10.1016/S0191-2615(97)00039-8

Carey, M., & Carville, S. (2000). Testing schedule performance and reliability for train stations. *Journal of the Operational Research Society*, *51*(6), 666–682. https://doi.org/10.1057/palgrave.jors.2600939

Christoforou, Z., Chandakas, E., & Kaparias, I. (2020). Investigating the Impact of Dwell Time on the Reliability of Urban Light Rail Operations. *Urban Rail Transit*. https://doi.org/10.1007/s40864-020-00128-1

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Cook, B. L., & Manning, W. G. (2013). *Thinking beyond the mean: A practical guide for using quantile regression methods for health services research*.

Cornet, S., Buisson, C., Ramond, F., Bouvarel, P., & Rodriguez, J. (2019). Methods for quantitative assessment of passenger flow influence on train dwell time in dense traffic areas. *Transportation Research Part C: Emerging Technologies*, *106*, 345–359. https://doi.org/10.1016/j.trc.2019.05.008

Correll, J., Mellinger, C., McClelland, G. H., & Judd, C. M. (2020). Avoid Cohen's 'Small', 'Medium', and 'Large' for Power Analysis. *Trends in Cognitive Sciences*, *24*(3), 200–207. https://doi.org/10.1016/j.tics.2019.12.009

Coulaud, R., Keribin, C., & Stoltz, G. (2023). Modeling dwell time in a data-rich railway environment: With operations and passenger flows data. *Transportation Research Part C: Emerging Technologies*, *146*, 103980. https://doi.org/10.1016/j.trc.2022.103980

Coxon, S., & Bono, A. D. (2010). Design strategies for mitigating passenger door holding behavior on suburban trains in Paris. *33rd Australasian Transport Research Forum Conference*.

Coxon, S., Bono, A., & Napper, R. (2009). The effect of suburban train carriage design upon punctuality, ingress and egress occlusion and passenger comfort. *32nd Australasian Transport Research Forum*, 1–11.

Cruijf, J. (n.d.). https://www.ad.nl/sport/de-beste-cruijffiaanse-uitspraken-da-s-logisch~a4182c68/?referrer=https%3A%2F%2Fwww.google.com%2F

Daamen, W. (2004). *Modelling Passenger Flows in Public Transport Facilities*. Delft University Press.

Daamen, W., Lee, Y. C., & Wiggenraad, P. (2008). Boarding and alighting experiments: Overview of setup and performance and some preliminary results. *Transportation Research Record*, *2042*, 71–81. https://doi.org/10.3141/2042-08

Davidich, M., Geiss, F., Mayer, H. G., Pfaffinger, A., & Royer, C. (2013). Waiting zones for realistic modelling of pedestrian dynamics: A case study using two major German railway stations as examples. *Transportation Research Part C: Emerging Technologies*, *37*, 210–222. https://doi.org/10.1016/j.trc.2013.02.016

Dell'Asin, G., & Hool, J. (2018). Pedestrian Patterns at Railway Platforms during Boarding: Evidence from a Case Study in Switzerland. *Journal of Advanced Transportation*, *2018*, 11. https://doi.org/10.1155/2018/4079230

Denti, E., & Burroni, L. (2023). Delay Indices for Train Punctuality. *Information*, *14*(5), 269. https://doi.org/10.3390/info14050269

Denyer, D., & Pilbeam, C. (2013). *Doing a literature review in business and management* [Presentation to the British Academy of Management Doctoral Symposium]. https://www.ifm.eng.cam.ac.uk/uploads/Research/RCDP/Resources/Working_with_literature_for_Cambridge.pdf

Dinmohammadi, F., Alkali, B., Shafiee, M., Bérenguer, C., & Labib, A. (2016). Risk Evaluation of Railway Rolling Stock Failures Using FMECA Technique: A Case Study of Passenger Door System. *Urban Rail Transit*, *2*(3–4), 128–145. https://doi.org/10.1007/s40864-016-0043-z

Dobbins, G. H., Lane, I. M., & Steiner, D. D. (1988). A note on the role of laboratory methodologies in applied behavioural research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior*, *9*(3), 281–286. https://doi.org/10.1002/job.4030090308

Douglas, N. (2012). *Modelling train and passenger capacity* [Technical report]. DOUGLAS Economics, Transport for NSW.

European Commission. (2021). *Seventh monitoring report on the development of the rail market.pdf*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021DC0005

European Environment Agency. (2022). *Transport and environment report 2022: Digitalisation in the mobility system : challenges and opportunities.* Publications Office. https://data.europa.eu/doi/10.2800/47438

Fang, J., Fujiyama, T., & Wong, H. (2019). Modelling passenger distribution on metro platforms based on passengers ' choices for boarding cars. *Transportation Planning and Technology*, *42:5*, 442–458. https://doi.org/10.1080/03081060.2019.1609218

Fernández, R., Swett, C., & Campo, D. (2008). *Data collection and calibration of passenger service time models for the transantiago system*. European Transport Conference 2009.

Fernández, R., Valencia, A., & Seriani, S. (2015). On passenger saturation flow in public transport doors. *Transportation Research Part A: Policy and Practice*, *78*, 102–112. https://doi.org/10.1016/j.tra.2015.05.001

Fiebag, T. (2019). *Innenraumgestaltung für S-Bahn Fahrzeuge am Beispiel der S-Bahn München* [Master thesis]. Fachhochschule St. Pölten.

Fox, C., Oliveira, L., Kirkwood, L., & Cain, R. (2017). Understanding users' behaviours in relation to concentrated boarding: Implications for rail infrastructure and technology. *Advances in Transdisciplinary Engineering*, *6*, 120–125. https://doi.org/10.3233/978-1-61499-792-4-120

Givoni, M., Brand, C., & Watkiss, P. (2009). Are Railways Climate Friendly? *Built Environment*, *35*(1), 70–86. https://doi.org/10.2148/benv.35.1.70

Glynn, L. (2006). A critical appraisal tool for library and information research. *Library Hi Tech*, *24*(3), 387–399. https://doi.org/10.1108/07378830610692154

Goverde, R. M. P. (2005). *Punctuality of Railway Operations and Timetable Stability Analysis* [PhD thesis, TU Delft]. https://repository.tudelft.nl/islandora/object/uuid%3Aa40ae4f1-1732-4bf3-bbf5-fdb8dfd635e7

Goverde, R. M. P., Hansen, I. a, Hooghiemstra, G., & Lopuhaa, H. P. (2001). Delay Distributions in Railway Stations. *The 9th World Conference on Transport Research*.

Grechi, D., & Maggi, E. (2018). The importance of punctuality in rail transport service: An empirical investigation on the delay determinants. *European Transport/Trasporti Europei*, *70*.

Gysin, K. (2018). *An Investigation of the Influences on Train Dwell Time* [Master thesis]. Swiss Federal Institute of Technology, ETH.

Hansen, I. A. (2010). Railway Network Timetabling and Dynamic Traffic Management. *International Journal of Civil Engineering*, *8*(1), 14.

Hao, L., & Naiman, D. (2007). *Quantile Regression*. SAGE Publications, Inc. https://doi.org/10.4135/9781412985550

Harris, N. G. (2005). Train boarding and alighting rates at high passenger loads. *Journal of Advanced Transportation*, *40*(3), 249–263. https://doi.org/10.1002/atr.5670400302

Harris, N. G. (2015). A European Comparison of Station Stop Delays. *International Congress on Advanced Railway Engineering*, 1–6.

Harris, N. G., & Anderson, R. J. (2007). An international comparison of urban rail boarding and alighting rates. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *221*(4), 521–526. https://doi.org/10.1243/09544097JRRT115

Harris, N. G., de Simone, F., & Condry, B. (2022). A Comprehensive Analysis of Passenger Alighting and Boarding Rates. *Urban Rail Transit*. https://doi.org/10.1007/s40864-021-00161-8

Harris, N. G., Mjøsund, C. S., & Haugland, H. (2013). Improving railway performance in Norway. *Journal of Rail Transport Planning and Management*, *3*(4), 172–180. https://doi.org/10.1016/j.jrtpm.2014.02.002

Harris, N. G., Risan, Ø., & Schrader, S.-J. (2014). *The impact of differing door widths on passenger movement rates*. 53–63. https://doi.org/10.2495/CRS140051

Harris, N., Graham, D. J., Anderson, R. J., & Li, H. (2014). The Impact of Urban Rail Boarding and Alighting Factors. *TRB 2014 Annual Meeting*, 13.

Heinz, W. (2003). *Passenger service times on trains, theory measurements, and models* [Licenciate]. KTH Royal Institute of Technology, Sweden.

Hirsch, L., & Thompson, K. (2014). I can sit but I'd rather stand: Commuter's experience of crowdedness and fellow passenger behaviour in carriages on Australian metropolitan trains. *ATRF 2011 - 34th Australasian Transport Research Forum*, *January*.

Holloway, C., Thoreau, R., Roan, T.-R., Boampong, D., Clarke, T., Watts, D., & Tyler, N. (2015). Effect of vertical step height on boarding and alighting time of train passengers. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *230*, 1234–1241. https://doi.org/10.1177/0954409715590480

Holmgren, J. (2020). The effect of public transport quality on car ownership – A source of wider benefits? *Research in Transportation Economics*, *83*, 100957. https://doi.org/10.1016/j.retrec.2020.100957

Hyun, D., Ko, D., & Robinson, S. E. (2016). *Reducing Dwell Time: London Underground Central Line* [Bachelor thesis]. Worcester Polytechnic Institute.

Jalali, S., & Wohlin, C. (2012). Systematic literature studies: Database searches vs. Backward snowballing. *International Symposium on Empirical Software Engineering and Measurement*, 29–38. https://doi.org/10.1145/2372251.2372257

Järnvägsbranschens samverkansforum. (2021). *TTT – Tillsammans för Tåg i Tid Årssammanfattning 2021*. https://bransch.trafikverket.se/contentassets/de2780dd12d847a6a5bae5c5f74907db/ttt_arssammanfattning_2021.pdf

Joborn, M., & Ranjbar, Z. (2022). Understanding causes of unpunctual trains: Delay contribution and critical disturbances. *Journal of Rail Transport Planning & Management*, *23*, 100339. https://doi.org/10.1016/j.jrtpm.2022.100339

Jones, H., Moura, F., & Domingos, T. (2017). Life cycle assessment of high-speed rail: A case study in Portugal. *The International Journal of Life Cycle Assessment*, *22*(3), 410–422. https://doi.org/10.1007/s11367-016-1177-7

Jusuf, Ç., Nemec, A., & Zahradnik, C. (2017). Optimization of Passenger Distribution at Metro Stations Through a Guidance System. *International Conference on Computer Aided Systems Theory*, 397–404. https://doi.org/10.1007/978-3-319-74727-9

Karlsson, A. (2006). *Estimation and inference for quantile regression of longitudinal data: With applications in biostatistics* [PhD thesis, Acta Universitatis Upsaliensis]. https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-7186

Kawabata, M. (2003). Jobs accessibility by travel mode in U.S. metropolitan areas. *Theory and Applications of GIS*, *11*(2), 165–172. https://doi.org/10.5638/thagis.11.165

Kecman, P., & Goverde, R. M. P. (2015). Predictive modelling of running and dwell times in railway traffic. *Public Transport*, *7*(3), 295–319. https://doi.org/10.1007/s12469-015-0106-7

Keim, D. A., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006). Challenges in Visual Data Analysis. *Tenth International Conference on Information Visualisation (IV'06)*, 9–16. https://doi.org/10.1109/IV.2006.31

Khadem Sameni, M., Preston, J., & Khadem Sameni, M. (2016). Evaluating efficiency of passenger railway stations: A DEA approach. *Research in Transportation Business & Management*, *20*, 33–38. https://doi.org/10.1016/j.rtbm.2016.06.001

Khorana, A., Pareek, A., Ollivier, M., Madjarova, S. J., Kunze, K. N., Nwachukwu, B. U., Karlsson, J., Marigi, E. M., & Williams, R. J. (2022). Choosing the appropriate measure of central tendency: Mean, median, or mode? *Knee Surgery, Sports Traumatology, Arthroscopy*, *31*(1), 12–15. https://doi.org/10.1007/s00167-022-07204-y

Kim, H., Kwon, S., Wu, S. K., & Sohn, K. (2014). Why do passengers choose a specific car of a metro train during the morning peak hours? *Transportation Research Part A: Policy and Practice*, *61*, 249–258. https://doi.org/10.1016/j.tra.2014.02.015

King, A. P., & Eckersley, R. J. (2019). Inferential Statistics III: Nonparametric Hypothesis Testing. In *Statistics for Biomedical Engineers and Scientists* (pp. 119–145). Elsevier. https://doi.org/10.1016/B978-0-08-102939-8.00015-3

Kourtit, K., Nijkamp, P., Türk, U., & Wahlstrom, M. (2022). City love and place quality assessment of liveable and loveable neighbourhoods in Rotterdam. *Land Use Policy*, *119*, 106109. https://doi.org/10.1016/j.landusepol.2022.106109

Krstanoski, N. (2014). Modelling passenger distribution on metro station platform. *International Journal for Traffic and Transport Engineering*, *4*(4), 456–465. https://doi.org/10.7708/ijtte.2014.4(4).08

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00863

Landmark, A. D., Seim, A. A., & Olsson, N. (2017). Visualisation of Train Punctuality – Illustrations and Cases. *Transportation Research Procedia*, *27*, 1227–1234. https://doi.org/10.1016/j.trpro.2017.12.092

Lantz, B. (2013). The large sample size fallacy. *Scandinavian Journal of Caring Sciences*, *27*(2), 487–492. https://doi.org/10.1111/j.1471-6712.2012.01052.x

Lee, J. (Brian), Yoo, S., Kim, H., & Chung, Y. (2018). The spatial and temporal variation in passenger service rate and its impact on train dwell time: A time-series clustering approach using dynamic time warping. *International Journal of Sustainable Transportation*, *12*(10), 725–736. https://doi.org/10.1080/15568318.2018.1432731

Li, D., Daamen, W., & Goverde, R. M. P. (2016). Estimation of train dwell time at short stops based on track occupation event data: A study at a Dutch railway station. *Journal of Advanced Transportation*, *50*(5), 877–896. https://doi.org/10.1002/atr.1380

Li, X., & Preston, J. (2015). Reassessing the financial and social costs of public transport. *Proceedings of the Institution of Civil Engineers - Transport*, *168*(4), 356–369. https://doi.org/10.1680/tran.12.00096

Lindfeldt, O. (2017). The impact of platform screen doors on rail capacity. *International Journal of Transport Development and Integration*, *1*(3), 601–610. https://doi.org/10.2495/TDI-V1-N3-601-610

Luangboriboon, N., Seriani, S., & Fujiyama, T. (2020). The influence of the density inside a train carriage on passenger boarding rate. *International Journal of Rail Transportation*, 1–16. https://doi.org/10.1080/23248378.2020.1846633

Luethi, M., Hürlimann, D., & Nash, A. (2005). *Understanding the timetable planning process as a closed control loop* [Application/pdf]. 13 p. https://doi.org/10.3929/ETHZ-A-005704049

Mackett, R. L., & Thoreau, R. (2015). Transport, social exclusion and health. *Journal of Transport & Health*, *2*(4), 610–617. https://doi.org/10.1016/j.jth.2015.07.006

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, 143–149. https://doi.org/10.11613/BM.2013.018

Monsuur, F., Enoch, M., Quddus, M., & Meek, S. (2021). Modelling the impact of rail delays on passenger satisfaction. *Transportation Research Part A: Policy and Practice*, *152*, 19–35. https://doi.org/10.1016/j.tra.2021.08.002

Nash, A., Weidmann, U., Bollinger, S., Luethi, M., & Buchmueller, S. (2006). Increasing Schedule Reliability on the S-Bahn in Zurich, Switzerland. *Transportation Research Record: Journal of the Transportation Research Board*, *1955*(1), 17–25. https://doi.org/10.1177/0361198106195500103

Ng, W.-S., & Acker, A. (2018). *Understanding Urban Travel Behaviour by Gender for Efficient and Equitable Transport Policies* (International Transport Forum Discussion Papers 2018/01; International Transport Forum Discussion Papers, Vol. 2018/01). https://doi.org/10.1787/eaf64f94-en

Noland, R. B., & Polak, J. W. (2002). Travel time variability: A review of theoretical and empirical issues. *Transport Reviews*, *22*(1), 39–54. https://doi.org/10.1080/01441640010022456

Økland, A., & Olsson, N. O. E. (2020). Punctuality development and delay explanation factors on Norwegian railways in the period 2005–2014. *Public Transport*. https://doi.org/10.1007/s12469-020-00236-y

Oliveira, L. C., Fox, C., Birrell, S., & Cain, R. (2019). Analysing passengers' behaviours when boarding trains to improve rail infrastructure and technology. *Robotics and Computer-Integrated Manufacturing*, *57*, 282–291. https://doi.org/10.1016/j.rcim.2018.12.008

Olsson, N. O. E., & Haugland, H. (2004). Influencing factors on train punctuality—Results from some Norwegian studies. *Transport Policy*, *11*(4), 387–397. https://doi.org/10.1016/j.tranpol.2004.07.001

Ostertagová, E., Ostertag, O., & Kováč, J. (2014). Methodology and Application of the Kruskal-Wallis Test. *Applied Mechanics and Materials*, *611*, 115–120. https://doi.org/10.4028/www.scientific.net/AMM.611.115

Palmqvist, C. W., & Kristoffersson, I. (2022). A Methodology for Monitoring Rail Punctuality Improvements. *IEEE Open Journal of Intelligent Transportation Systems*, *3*, 388–396. https://doi.org/10.1109/OJITS.2022.3172509

Palmqvist, C. W., Tomii, N., & Ochiai, Y. (2020). Explaining dwell time delays with passenger counts for some commuter trains in Stockholm and Tokyo. *Journal of Rail Transport Planning and Management*, *14*, 100189. https://doi.org/10.1016/j.jrtpm.2020.100189

Palmqvist, C.-W. (2019). *Delays and Timetabling for Passenger Trains* [Doctoral thesis, Lund University Faculty of Engineering, Technology and Society, Transport and Roads]. http://portal.research.lu.se/ws/files/70626078/Carl_William_Palmqvist_web.pdf

Palmqvist, C.-W. (2022). Excess probability of dwell time delays from train meets and passes. *Journal of Rail Transport Planning & Management*, *21*, 100298. https://doi.org/10.1016/j.jrtpm.2022.100298

Palmqvist, C.-W., Olsson, N. O. E., & Hiselius, L. W. (2017). Punctuality problems from the perspective of timetable planners in Sweden. *Transportation Research Procedia*, 9.

Palmqvist, C.-W., Olsson, N. O. E., & Winslott Hiselius, L. (2018). The Planners' Perspective on Train Timetable Errors in Sweden. *Journal of Advanced Transportation*, *2018*, 1–17. https://doi.org/10.1155/2018/8502819

Parbo, J., Nielsen, O. A., & Prato, C. G. (2016). Passenger Perspectives in Railway Timetabling: A Literature Review. *Transport Reviews*, *36*(4), 500–526. https://doi.org/10.1080/01441647.2015.1113574

Pedersen, T., Nygreen, T., & Lindfeldt, A. (2018). *Analysis of temporal factors influencing minimum dwell time distributions*. 447–458. https://doi.org/10.2495/CR180401

Peftitsi, S., Jenelius, E., & Cats, O. (2020). Determinants of passengers' metro car choice revealed through automated data sources: A Stockholm case study. *Transportmetrica A: Transport Science*, *16*(3), 529–549. https://doi.org/10.1080/23249935.2020.1720040

Perneger, T. V. (2005). The Swiss cheese model of safety incidents: Are there holes in the metaphor? *BMC Health Services Research*, *5*(1), 71. https://doi.org/10.1186/1472-6963-5-71

Pritchard, J., Sadler, J., Blainey, S., Waldock, I., & Austin, J. (2021). Predicting and mitigating small fluctuations in station dwell times. *Journal of Rail Transport Planning & Management*, *18*, 100249. https://doi.org/10.1016/j.jrtpm.2021.100249

Puong, A. (2000). *Dwell Time Model and Analysis for the MBTA Red Line*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.583.7667&rep=rep1&type=pdf

Reason, J. (2000). *Human error: Models and management*. *320*.

Rietveld, P., Bruinsma, F. R., & van Vuuren, D. J. (2001). Coping with unreliability in public transport chains: A case study for Netherlands. *Transportation Research Part A: Policy and Practice*, *35*(6), 539–559. https://doi.org/10.1016/S0965-8564(00)00006-9

Robenek, T., Maknoon, Y., Azadeh, S. S., Chen, J., & Bierlaire, M. (2016). Passenger centric train timetabling problem. *Transportation Research Part B: Methodological*, *89*, 107–126. https://doi.org/10.1016/j.trb.2016.04.003

Rodríguez, G. de A., Seriani, S., & Holloway, C. (2015). The impact of platform edge doors on passengers boarding and alighting time and platform behaviour. *Transportation Research Board 95th Annual Meeting*, 17.

Rodriguez, R. N., & Yonggang, Y. (2017). *Five Things You Should Know About Quantile Regression* (Paper SAS525). SAS Institute Inc.

Rousseeuw, P. J. (1991). Tutorial to robust statistics. *Journal of Chemometrics*, *5*(1), 1–20. https://doi.org/10.1002/cem.1180050103

Rüger, B. (2018). How platform infrastructure influences passenger behaviour. *International Journal for Traffic and Transport Engineering*, 9. http://dx.doi.org/10.7708/ijtte.2018.8(4).04

Sarstedt, M., & Mooi, E. (2014). Regression Analysis. In M. Sarstedt & E. Mooi, *A Concise Guide to Market Research* (pp. 193–233). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-53965-7_7

SBB CFF FFS. (n.d.). *Punctuality*. Retrieved August 8, 2023, from https://reporting.sbb.ch/punctuality

Scheff, S. W. (2016). Nonparametric Statistics. In *Fundamental Statistical Principles for the Neurobiologist* (pp. 157–182). Elsevier. https://doi.org/10.1016/B978-0-12-804753-8.00008-7

Seriani, S., Barriga, J. M., Peña, A., Valencia, A., Aprigliano, V., Jorquera, L., Pinto, H., Valenzuela, M., & Fujiyama, T. (2022). Analyzing the Effect of Crowds on Passenger Behavior Inside Urban Trains through Laboratory Experiments—A Pilot Study. *Sustainability*, *14*(22), 14882. https://doi.org/10.3390/su142214882

Seriani, S., Fernandez, R., Luangboriboon, N., & Fujiyama, T. (2019). Exploring the Effect of Boarding and Alighting Ratio on Passengers' Behaviour at Metro Stations by Laboratory Experiments. *Journal of Advanced Transportation*, *2019*. https://doi.org/10.1155/2019/6530897

Seriani, S., & Fujiyama, T. (2019a). Exploring the Effect of Train Design Features on the Boarding and Alighting Time by Laboratory Experiments. *Collective Dynamics*, *4*, A22. https://doi.org/10.17815/CD.2019.22

Seriani, S., & Fujiyama, T. (2019b). Modelling the distribution of passengers waiting to board the train at metro stations. *Journal of Rail Transport Planning & Management*, *11*(August), 100141. https://doi.org/10.1016/j.jrtpm.2019.100141

Seriani, S., Fujiyama, T., & de Ana Rodríguez, G. (2016). Boarding and Alighting Matrix on Behaviour and Interaction at the Platform Train Interface. *Rail Research UK Association (RRUKA) 2016 Annual Conference*, *November*.

Seriani, S., Fujiyama, T., & Holloway, C. (2016). Exploring the pedestrian level of interaction on platform conflict areas at metro stations by real-scale laboratory experiments. *Transportation Planning and Technology*, *40*(1), 100–118. https://doi.org/10.1080/03081060.2016.1238574

Sommet, N., & Morselli, D. (2017). Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS. *International Review of Social Psychology*, *30*(1), 203–218. https://doi.org/10.5334/irsp.90

Sørensen, A. Ø., Bjelland, J., Bull-Berg, H., Landmark, A. D., Akhtar, M. M., & Olsson, N. O. E. (2018). Use of mobile phone data for analysis of number of train travellers. *Journal of Rail Transport Planning and Management*, *8*(2), 123–144. https://doi.org/10.1016/j.jrtpm.2018.06.002

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia Medica*, 12–18. https://doi.org/10.11613/BM.2014.003

Staffa, S. J., Kohane, D. S., & Zurakowski, D. (2019). Quantile Regression and Its Applications: A Primer for Anesthesiologists. *Anesthesia & Analgesia*, *128*(4), 820–830. https://doi.org/10.1213/ANE.0000000000004017

Stoilova, S., & Nikolova, R. (2017). Classifying railway passenger stations for use transport planning – application to Bulgarian railway network. *Transport Problems*, *11*(2), 143–155. https://doi.org/10.20858/tp.2016.11.2.14

Storstockholms Lokaltrafik. (2020). *Fakta om: SL och länet 2020*. https://www.regionstockholm.se/globalassets/2.-kollektivtrafik/fakta-om-sl-och-lanet/fakta-om-sl-lanet-2020.-pdf.pdf

Thoreau, R., Holloway, C., Bansal, G., Gharatya, K., Roan, T. R., & Tyler, N. (2016). Train design features affecting boarding and alighting of passengers. *Journal of Advanced Transportation*, *50*(8), 2077–2088. https://doi.org/10.1002/atr.1446

Tortainchai, N., Wong, H., Winslett, D., & Fujiyama, T. (2022). Train Dwell Time Efficiency Evaluation with Data Envelopment Analysis: Case Study of London Underground Victoria Line. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(3), 728–739. https://doi.org/10.1177/03611981211056640

van den Heuvel, J. (2016). Field Experiments with Train Stopping Positions at Schiphol Airport Train Station in Amsterdam, Netherlands. *Transportation Research Record: Journal of the Transportation Research Board*, *2546*(1), 24–32. https://doi.org/10.3141/2546-04

Van Den Heuvel, J. P. A., & Hoogenraad, J. H. (2014). Monitoring the performance of the pedestrian transfer function of train stations using automatic fare collection data. *Transportation Research Procedia*, *2*, 642–650. https://doi.org/10.1016/j.trpro.2014.09.107

van Lierop, D., Badami, M. G., & El-Geneidy, A. M. (2018). What influences satisfaction and loyalty in public transport? A review of the literature. *Transport Reviews*, *38*(1), 52–72. https://doi.org/10.1080/01441647.2017.1298683

van Loon, R., Rietveld, P., & Brons, M. (2011). Travel-time reliability impacts on railway passenger demand: A revealed preference analysis. *Journal of Transport Geography*, *19*(4), 917–925. https://doi.org/10.1016/j.jtrangeo.2010.11.009

Volovski, M., Ieronymaki, E. S., Cao, C., & O'Loughlin, J. P. (2021). Subway station dwell time prediction and user-induced delay. *Transportmetrica A: Transport Science*, *17*(4), 521–539. https://doi.org/10.1080/23249935.2020.1798555

Vromans, M. J. C. M. (2005). *Reliability of railway systems*. Erasmus Research Inst. of Management (ERIM).

Wainer, H., & Thissen, D. (1981). Graphical Data Analysis. *Annual Review of Psychology*, *32*(1), 191–241. https://doi.org/10.1146/annurev.ps.32.020181.001203

Waldmann, E. (2018). Quantile regression: A short story on how and why. *Statistical Modelling*, *18*(3–4), 203–218. https://doi.org/10.1177/1471082X18759142

Weber, K., Zangl, M., Zahid, M. U., & Holzner, M. (2022). *Environmental Impact Evaluation of a European High Speed Railway Network along the 'European Silk Road'* (Research Report 459). The Vienna Institute for

International Economic Studies. https://wiiw.ac.at/environmental-impact-evaluation-of-a-european-high-speed-railway-network-along-the-european-silk-road-dlp-5837.pdf

Weston, J. G. (1989). *Train service model-technical guide* (Note 89/18; London Underground Operational Research). London Underground.

Wiggenraad, P. B. L. (2001). Alighting and boarding times of passengers at Dutch railway stations. *TRAIL Research School, Delft*, *December*, 1–21.

Wilkinson, L. & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*(8), 594–604. https://doi.org/10.1037/0003-066X.54.8.594

Wu, J., Ph, D., & Ma, S. (2013). Division Method for Waiting Areas on Island Platforms at Metro Stations. *Journal of Transportation Engineering*, *139*(4), 339–349. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000484.

Yabuki, H., Takatori, Y., Koresawa, M., & Tomii, N. (2017). Increasing robustness of timetables by deliberate operation of trains to shorten headways. *WIT Transactions on State-of-the-Art in Science and Engineering*, *94*, 126–135. https://doi.org/DOI: 10.2495/TDI-V1-N3-432-441

Yamamura, A., Koresawa, M., Adachi, S., & Tomii, N. (2012). *Identification of causes of delays in urban railways*. 403–414. https://doi.org/10.2495/CR120341

Yang, J., Shiwakoti, N., & Tay, R. (2019). Train dwell time models – development in the past forty years. *Australasian Transport Research Forum 2019 Proceedings*, 12.

Zemp, S., Stauffacher, M., Lang, D. J., & Scholz, R. W. (2011). Classifying railway stations for strategic transport and land use planning: Context matters! *Journal of Transport Geography*, *19*(4), 670–679. https://doi.org/10.1016/j.jtrangeo.2010.08.008

Zheng, M. C. (2018). Empirical study on congested subway transfer traffic patterns. *International Journal of Transport Development and Integration*, *2*(3), 258–270. https://doi.org/10.2495/TDI-V2-N3-258-270

Zhou, Y., Zheng, S., Hu, Z., & Chen, Y. (2022). Metro station risk classification based on smart card data: A case study in Beijing. *Physica A: Statistical Mechanics and Its Applications*, *594*, 127019. https://doi.org/10.1016/j.physa.2022.127019