Working Paper: Modeling the Infection Risk in Local Public Transport - Contact Number and Contact Time

Lea Fouckhardt[1], Sophie Laufer[2], Jochen Sauer[2], Johanna Koch[1], Natalie Schneider[1], Carsten Sommer[1]

1 University of Kassel

2 WVI Prof. Dr. Wermuth Verkehrsforschung und Infrastrukturplanung GmbH

DOI: 10.17170/kobra-202401259432

Abstract

To model the infection risk in local public transport, additional information on the personal contacts during a trip with local public transport as well as the contact duration with them is necessary. As morning student traffic in local public transport vehicles was assumed to hold a high infection risk during the COVID-19 pandemic, staggered school starting times were tested in the city of Herne in Germany. The data obtained during this time was used to create a calculation tool for the number of personal contacts during a ride in a bus, based on the maximum occupancy during the ride. The average contact time with a person was dependent on the total ride time. Both values were adjusted for different vehicle types in local public transport.

Keywords: infection risk, local public transport, LPT, contact number in LPT, contact time in LPT

Background

COVID-19 and Student Traffic

During the COVID-19 pandemic, people strived to avoid infections with the SARS-CoV-2 and the question where people were most likely to be infected emerged. Due to COVID-19 spreading via breath aerosols (Wang et al., 2021), distance was established as an important influence on infection risk in public transport (Schneider et al., 2022). Local public transport (LPT) was assumed to be a place of high infection rates due to being rather crowded and people were advised against using it. Especially the morning rush of students in vehicles of LPT was seen as a risk of infection (ABC News, 2020; BBC News, 2020; Süddeutsche Zeitung, 2020) and a case for measures reducing infection risk was made (#BesserWeiter, 2021).

A staggered timetable was introduced in the German city of Herne in November 2020 in order to ease morning school traffic. As there was no evidence at the time to what extent this measure could help reduce the risk of infection in public transport vehicles, an examination was conducted (Koch et al., 2024, in press). The concept provided to stagger the start of lessons at secondary schools in the city according to school type and year group. Previously, all schools had started lessons at 8:00 a.m. (with the exception of one school), whereas the start was now staggered between 7:30 and 8:45 a.m. As a result of the new teaching times, the timetables had been adjusted. The concept was developed through the coordination of the local transport company "Straßenbahn Herne – Castrop-Rauxel GmbH" (HCR), which provides bus services in the city.

To investigate the effects of the staggered times, an allocation model was developed, which could map the distribution of student demand using various input variables. For this purpose, a routing based on the timetable was carried out for all students, which resulted in the most likely journey from home to school with all partial routes used. (Koch et al., 2024, in press). With this data, the number of contacts for each student, as well as the contact time was calculated (for different staggered scenarios). Based on this, a calculation approach for other vehicles of LPT was aspired. The following text describes the process of this in greater detail. All calculations only apply to single trips without changeovers and trip chains. Therefore, the expression ***journey section*** will be used for trips without changeovers or trip parts without changeovers. To estimate the infection risk for a trip chain, the risks for each journey section can likely be summarized.

The Impact of Staggered School Starting Times on Infection Risk

The spread-out starting times of the school consequently meant that students used LPT over a longer time period in the morning. This again resulted in a lower average occupancy in the vehicles of LPT, which thus translates to less contacts when using the LPT vehicles. From this project, extensive data on student bus traffic in the morning in Herne was readily available.

Student Traffic in Herne

Results of the routing process – of the students considered in the five scenarios – yielded detailed information on a student's bus line, boarding and alighting. For other passengers (called base load in further explanations), only data from automated passenger counting systems was available and used to calculate the base load contacts for each student.

Determining Contact Number and Time

Student to Student Contacts

For each part of a student's journey section, other students on the same journey section were determined based on the overlap of bus line and bus stops used by the respective students (e.g. contact of student 123 with student 456 on drive 789 between stops 1 and 5). A scenario in which student 123 exited the bus before student 456 boards, thus didn't count as a contact. For simplification, past contacts between two students on earlier journey sections were not considered, instead this was assumed to be multiple contacts instead of one longer contact. A contact time was estimated for each student to student contact.

Base Load Contacts

Due to missing source-destination information for non-student passengers, these contacts were simplified. The contact number for each student was described by the occupancy at the point in time were the student makes an entrance to the bus. This consisted of both students and non-students. This number was elevated by the number of entrances of base load passengers until the student exited the vehicle. This was the number of base load contacts.

Calculation of the Average Contact Time

The students' data was aggregated by combining contacts with the same contact times. For each student and contact time the number of contacts was calculated (e.g. student 123 had two minutes of contact with two other students, and student 123 also had four minutes of contact with another student, …). For the base load, it was assumed that the distribution of contact times was identical to the distribution in the student to student contacts. The results of these assumptions were distributed over the time slices.

Based on these results, the average contact time for each student was calculated. This was done by averaging the contact number aggregated by time slices. It was weighted by the number of contacts.

As an example, for student 123 this meant: the student had two contacts of two minutes and one contact of four minutes. Their average contact time was therefore:

$$\frac{2 \cdot 2\,min + 1 \cdot 4 min}{(2+1)} = 2{,}67\ min$$

Generalization

A simple extrapolation of the complete calculation process from Herne to other data sources such as larger surveys was not possible in lieu of the EMILIA project. However, it is likely that demand situations vary drastically, dependent on the LPT's system design, development, station times and vehicle sizes used. The transfer on other cases and vehicles was based on several assumptions and therefore constitutes a model of calculation. Thus, important parameters influencing contact time and contact number were derived of the Herne analysis. These parameters were later adjusted for other vehicles.

Derivation of Contact Numbers

There was a linear relation between the individual maximum occupancy in a vehicle and the individual number of contacts during a journey section (Figure 1). This showed that the maximum occupancy could be used as a predictor for calculating the total number of contacts. To take the rise of total number of contacts with longer journey section durations under the same maximum occupancy into account, the Herne results were classified by journey section duration (class width 5 minutes). As expected, journey sections with longer durations showed higher increases in the number of contacts, due to passenger exchanges along the route and recorded maximum occupancy.

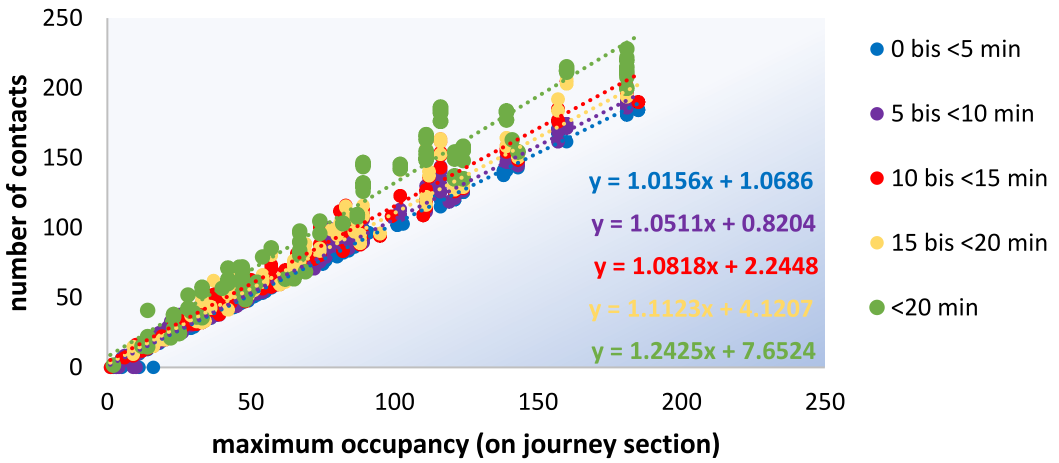

Figure 1: Relation between maximum occupancy of a journey section and total number of contacts during the journey section in a bus for students in Herne

$$A = \begin{cases} 1{,}0156 * b + 1{,}0686 \ for\ f \in [0; 5[\\ 1{,}0511 * b + 0{,}8204 \ for\ f \in [5; 10[\\ 1{,}0818 * b + 2{,}2448 \ for\ f \in [10; 15[\\ 1{,}1123 * b + 4{,}1207 \ for\ f \in [15; 20[\\ 1{,}2425 * b + 7{,}6524 \ for\ f \in [20; \infty[\end{cases}$$

with A = number of contacts

b = maximum occupancy during the journey section

f = own journey section duration

Derivation of Average Contact Time

Since universally accurate source-destination information of all passengers was not available, the average contact time was used instead of the distribution of all contact times of one passenger. The contact time depended on several factors. A person's journey section duration was a natural limit to their respective contact times, as long as shared changeovers were dismissed. Rather, contact times were shorter than one's own journey section duration. This was the case, if passengers entered later along the target person's journey section or exited sooner that the target person.

Analysis of the Herne results pointed to a positive relation between the target person's journey section duration and the average contact time (Figure 2). The contact time grew with rising journey section duration. For very long journey sections the average contact time stagnated as more passengers boarded and exited the vehicle within that time frame, therefore limiting the average contact time of the target person by the duration of their own journey sections. Due to this, a linear function wasn't sufficient as an explanation. To approximate the observed values, the function was divided into sections.

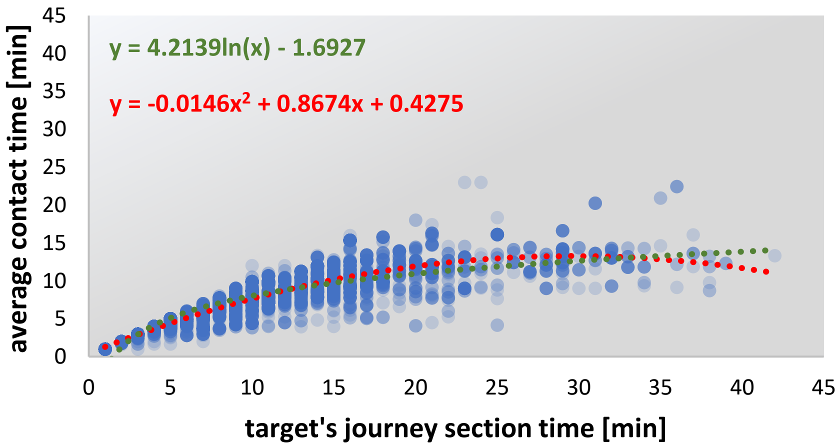

Figure 2: Relation between a target person's journey section time and average contact time for the students of Herne

The logarithmic function shown in Figure 2 describes the relationship between journey section duration and average contact time well for journey section durations longer that 13 minutes. However, for journey section durations between zero and five minutes this would yield contact times longer than the target's journey section duration. To avoid these errors, the assumption that changeover and passenger exchange play a negligible role in very short durations of up to five minutes was made, and therefore journey section duration and contact time can be equated in this interval. The quadratic function expresses the relation between five and 13 minutes of journey section duration. In reality, this interval played a significant role in journey section duration distributions (see also Figure 3 and Figure 4). The logarithmic function would overestimate the contact durations in this time interval. The limits were chosen as to avoid gaps in the function of the relation between journey section duration and contact time.

$$K_{city\ bus} = \begin{cases} f \ for \ f \in [0; 5[\\ -0{,}0146 * f2 + 0{,}8674 * f + 0{,}425 \ for \ f \in [5; 13[\\ 4{,}2139 * \ln(f) - 1{,}6927 \ for \ f \in [13; \infty] \end{cases}$$

with K = average contact time

f = target's journey section duration

Transfer to Other Vehicles of LPT

To transfer the results of the city bus system in Herne to other vehicles of LPT, two assumptions were made:

- The journey section time, the maximum occupancy and the passenger changeover are the decisive factors influencing the contact numbers and times.
- The contact times are very different between vehicles.

Table 1: Average passenger numbers and conversion factors based on Herne's city bus system for urban and regional vehicles (source: WVI survey)

	vehicle	passengers	⌀ journey section duration [min]	conversion factor
urban	city bus	470.000	11.7	1
	tram	70.000	10.8	1
	subway	90.000	9.0	1
regional	regional bus	470.000	16.2	1.5
	suburban rail	510.000	16.9	1.5
	regional rail	330.000	26.8	2.5

Through comparisons of journey section durations between Herne and the traffic of city busses, regional busses, trams, subways, suburban and regional rail, conversion factors for the journey section duration classes were found. These factors enable the transfer of journey section duration classes to the other vehicles. Table 1 shows the average journey section durations for the different vehicles. The corresponding journey section duration distributions can be found in Figure 3 and Figure 4.

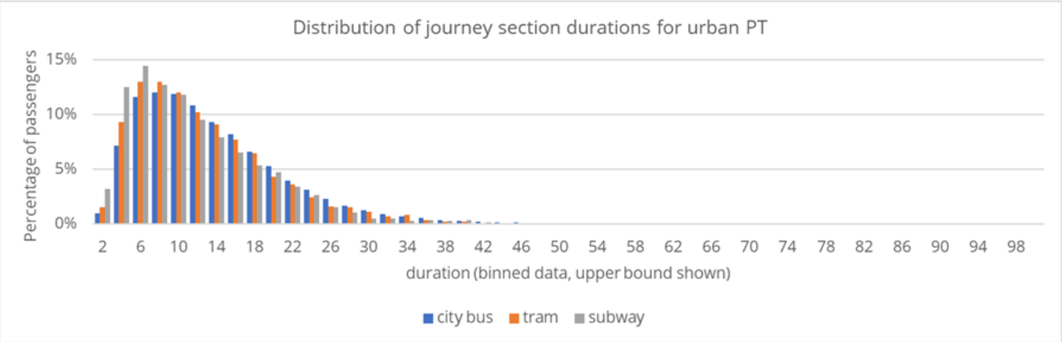

Figure 3: Trip duration distributions for urban PT (source: WVI survey)

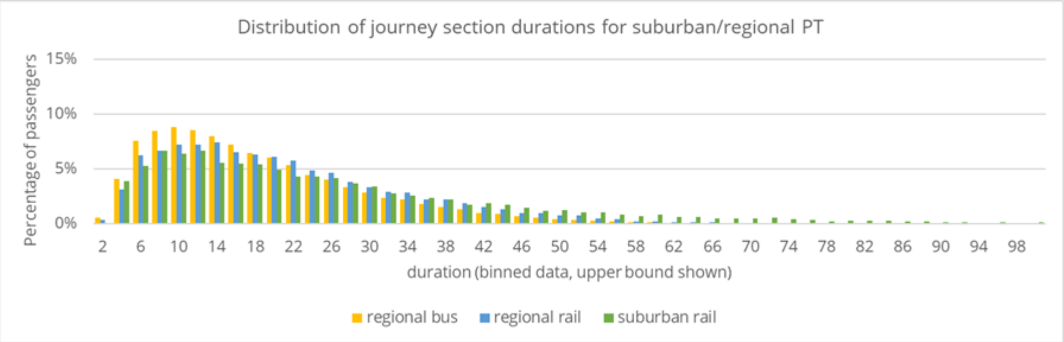

Figure 4: Trip duration distributions for regional PT (source: WVI survey)

The distributions of journey section duration for urban public transport systems were similar. The average journey section durations of 9 to 13 minutes coulf be approximated as being interchangeable. Similarly, regional public transport systems such as regional busses and suburban rail showed similar distributions for journey section times of 16 to 17 minutes, while journey section durations in regional rail were much longer on average at 27 minutes. The relations between the functions for the number of contacts in the city bus system could be transferred to other vehicles through the conversion factors.

Example: The conversion factor for regional rail is 2.5 in comparison to city busses (Herne). The function for calculating the number of contacts in journey section duration class 1 (zero to five minutes in Herne) is applied to 2.5-times the range (meaning: instead of zero to five minutes it is applied for zero to 12.5 minutes for regional rail). This in turn means that the function for journey section duration class 2 would only be applied for much higher values for regional rail. For the average contact time, the functions from the Herne data were also transferred via these conversion factors, for the target's own journey section duration and for the predicted average contact time.

Implementation of the Results

The approach for calculating both contact numbers and contact time for different vehicles in LPT is important groundwork for modelling the infection risk when using LPT, which uses the contact time and the number of passenger contacts as two variables (Fouckhardt et al., 2024, manuscript in preparation). It needs to be mentioned that the contact times and contact numbers were estimated with a student traffic sample. It is highly likely that non-student traffic differs from this.

However, data from automated passenger counting systems focuses on the number of passengers in relation to the vehicle, not in relation to other passengers. This means that while the occupancy of the vehicle at different times is reported, is it not known which section of the journey is undertaken by which passenger. Consequently, passenger to passenger contacts cannot be deduced from this kind of data. In fact, up until the COVID-19 pandemic, the contact time between passengers did not seem of much importance to the LPT systems. To our current knowledge, more adequate data that considers factors such as urban or rural environment, time of day and line type, is not available. Thus, the data on student traffic will be used as the best approximation to yet exist. Further studies on this exact topic are highly advised.

Additional Information

Both the approach to calculating contact number and contact time, as well as the examination of effects of staggered school starting times were part of the EMILIA project. EMILIA stands for „Entwicklung eines pandemieresistenten Öffentlichen Personennahverkehrs" (Development of a pandemic-resistent LPT) and is funded by the German Federal Ministry of Digitalization and Traffic.

This working paper is based on the internal text written by WVI for the EMILIA-project of the University of Kassel (Laufer & Sauer, 2023).

Literature

#BesserWeiter. (2021). *Verkehrsunternehmen fordern gestaffelten Unterrichtsbeginn: Kleine Maßnahmen mit großer Wirkung*. https://www.besserweiter.de/verkehrsunternehmen-fordern-gestaffelten-unterrichtsbeginn.html

ABC News (6th August 2020). School buses adapt to keep kids safe during COVID-19 crisis. *ABC News*. https://abcnews.go.com/GMA/News/school-buses-adapt-kids-safe-covid-19-crisis/story?id=72212386

BBC News (22nd September 2020). Covid: School bus drivers fearful of coronavirus infection risk. *BBC News*. https://www.bbc.com/news/uk-54203100

Fouckhardt, L., Schneider, N., Koch, J., Helfers, A., Sommer, C. & Vilfayeau, S. (2024). *How Safe is Public Transport in a Global Pandemic? Simulation-Based Insights on the Infection Probability with Sars-Cov-2*. Manuscript in Preparation.

Koch, J., Schneider, N., Henniger, J., Fouckhardt, L. & Sommer, C. (2024). Wirkungen einer pandemiebedingten Schulzeitstaffelung am Beispiel der Stadt Herne. *Heureka - Optimierung in Verkehr und Transport*.

Laufer, S. & Sauer, J. (2023). *Berechnungsansatz zur Kontaktwahrscheinlichkeit: Dokumentation Ermittlung Kontaktzahl und Kontaktdauer*. Forschungsprojekt EMILIA.

Schneider, N., Helfers, A., Reiserer, M. & Sommer, C. (2022). Subjektives Empfinden des Risikos im ÖPNV während der Pandemie: Sicherheit aus Fahrgastsicht. *Der Nahverkehr*(11/2022), 18–22.

Süddeutsche Zeitung (28th August 2020). Volle Busse im Schülerverkehr: Eltern wegen Corona in Sorge. *Süddeutsche Zeitung*. https://www.sueddeutsche.de/bildung/schulen-wiesbaden-volle-busse-im-schuelerverkehr-eltern-wegen-corona-in-sorge-dpa.urn-newsml-dpa-com-20090101-200828-99-338057

Wang, C. C., Prather, K. A., Sznitman, J., Jimenez, J. L., Lakdawala, S. S., Tufekci, Z. & Marr, L. C. (2021). Airborne transmission of respiratory viruses. *Science (New York, N.Y.)*, *373*(6558).