

MIT Open Access Articles

*Identifying the Potential for Partial Integration of Private and Public
Transportation*

The MIT Faculty has made this article openly available. **Please share**
how this access benefits you. Your story matters.

Citation: Sustainability 13 (6): 3424 (2021)

Published Version: <http://dx.doi.org/10.3390/su13063424>

Publisher: Multidisciplinary Digital Publishing Institute

Permanent Link: <https://hdl.handle.net/1721.1/131354>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: <https://creativecommons.org/licenses/by/4.0/>



Article

Identifying the Potential for Partial Integration of Private and Public Transportation

Iva Bojic ^{1,*}, Dániel Kondor ^{1,†}, Wei Tu ², Ke Mai ², Paolo Santi ^{3,4} and Carlo Ratti ^{1,3}

¹ Singapore-MIT Alliance for Research and Technology, Singapore 138602, Singapore; dkondor@mit.edu (D.K.); ratti@mit.edu (C.R.)

² Guangdong Key Laboratory of Urban Informatics, Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518000, China; tuwei@szu.edu.cn (W.T.); mk@szu.edu.cn (K.M.)

³ Senseable City Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; psanti@mit.edu

⁴ Istituto di Informatica e Telematica del CNR, 56127 Pisa, Italy

* Correspondence: ivabojoic@mit.edu

† These authors contributed equally to this work.

Abstract: In this paper, we explore the potential for a partial integration of private and public transportation as provided by taxi and bus services on the level of individual trips. While these modes are typically separated for economic reasons, in the future with shared Autonomous Vehicles (AVs) providing cheap and efficient transportation services, such distinctions will blur. Consequently, optimization based on real-time data will allow exploiting parallels in demand in a dynamic way, such as the proposed approach in the current work. New operational and pricing strategies will then need to be considered, along with regulatory tools to limit congestion, encourage sharing and provide services in a more sustainable way by utilizing a dynamic landscape of urban transportation. In the current work, we evaluate existing parallels between individual bus and taxi trips in two Asian cities and show how exploiting these synergies could form the basis of achieving some of the above goals in the future.

Keywords: travel time savings; bus passengers matching; dynamic ride-sharing; Autonomous Vehicles



Citation: Bojic, I.; Kondor, D.; Tu, W.; Mai, K.; Santi, P.; Ratti, C. Identifying the Potential for Partial Integration of Private and Public Transportation. *Sustainability* **2021**, *13*, 3424. <https://doi.org/10.3390/su13063424>

Academic Editor: Kenji Doi

Received: 9 February 2021

Accepted: 16 March 2021

Published: 19 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In today's transportation industry, there is a large operational gap between private end-to-end services such as taxis (We note that the taxi market is often highly regulated and in some cases, taxi companies are government-owned. Due to these reasons, some authors consider taxis as part of public transportation. Nevertheless, for the purposes of the current work, we believe the important distinction is along the lines presented here.) and ride-sharing, and fixed-route public transportation such as buses and subways [1,2]. While the former provide more convenience for passengers and usually a significantly shorter travel time at a higher price, the latter operate on inflexible schedules and routes, with many intermediate stops that slow down the service, but at a significantly lower cost. In many countries, government policies in the past decades have favored private transportation by ensuring low fuel prices and investing in road infrastructure, despite the increasing evidence of the significant societal cost of private transportation, up to 28 times higher than public transportation [3]. Due to the large increase of vehicles on roads, cities around the world face significant problems due to congestion; in peak hours, public transportation options can provide a shorter travel time when separated from general traffic [4].

Recently, disruptive changes caused by the emergence of Transportation Network Companies (TNCs), such as Uber, DiDi or Grab contributed to an increase in congestion in cities across the world [5]. While originally perceived as competitors to taxis [6,7], it soon

became apparent that TNCs serve as competitors to public transportation as well [8–10]. With the expected availability of Autonomous Vehicles (AVs) in the near future [11–13], congestion trends are likely to continue at an increased pace. Eliminating drivers' salary can result in Autonomous Mobility On Demand (AMOD) services becoming dramatically cheaper than services with human drivers today [14,15], raising concerns if public transportation can remain competitive at all [16]. A shift to AMOD from public transportation can have drastic effects in terms of congestion and total energy use, even if we account for more efficient traffic flows [17] and greener drive-chain technologies [18,19]. At the same time, these cost savings will allow public transportation operators to explore innovative in-between solutions with new vehicle form factors to provide increased level of service for commuters without increasing costs [20,21]. Still, it is yet unclear what role will AVs play in the future of transportation. Previous research has considered cases of individually owned AVs [16,22–25], AMOD fleets, both with and without ridesharing [14,15,21], and also using AVs as first- and last-mile transportation options in a close integration with transit [20,26,27].

Given the above concerns, there will be a strong need for policies that limit demand and promote sustainable, shared forms of transportation. Multiple cities have recently started implementing congestion charges and regulation specifically aimed at TNC companies [5,28–30]. At the same time, there is significant research done on proposed new ways of collaboration between AMOD and public transportation [26,27,31]. In this regard, AVs present not only challenges, but opportunities as well: since the cost of operation will be much lower than of human-driven taxis of today [14,15], such services will have a lot of flexibility in terms of pricing, giving powerful tools for governments to influence demand and incentivize more sustainable solutions [32,33]. Thus, we envision a future where private (i.e., non-shared) on-demand transportation remains a “premium” option, with a large portion of the price being a congestion fee levied by governments, while a flexible array of shared options are presented with the aim to attract both private and public transportation users of today [1].

In this paper, we investigate the feasibility of an ad-hoc shared transportation option for private and public transportation users. With such a pairing, public transportation passengers can reduce their travel time, while AMOD passengers can be offered discounts; all of this will result in an increase of overall service quality without adding to the traffic volume. The idea of combining different modes of transportation for one trip is not a new one, with Mobility-as-a-Service (MaaS) term denoting exactly that—*bundling different transport means, public and private, into one easy-to-use package for customer* [34]. The potential for MaaS was explored in different countries such as Sweden [35], Finland [36] and Germany [37]. However, most of the existing research on MaaS is assuming that the modes of transportation are chosen before the trip starts and different modes are taken in sequence by the same passenger, while in this paper, we investigate the situation where the matching is done “on the go”, in an opportunistic manner, with faster and more convenient modes substituting for slower ones wherever possible. Moreover, unlike most of the existing work on AVs and public transportation, where AVs are used as a first/last mile solution [27,38,39], we allow the whole public transportation segment to be replaced with a shared trip in an AV.

We note that the sharing scheme proposed in our paper could be implemented even today with taxi and TNC operators instead of AVs. However, currently the cost and pricing structure of taxi and TNC operations are such that implementing the above-mentioned sharing solution makes little economic sense. This way, we only expect this concept to be relevant when AVs become a reality and government actors implement an integrated framework for regulation, pricing and demand management, as outlined above. With that being said, we aim to answer the following two research questions in this paper: (i) what is the percentage of public transportation passengers who could be matched with the private transportation passengers; and (ii) what is the benefit for matched passengers, measured as average travel time savings.

We base our analysis on the current demand of bus and taxi trips; while we acknowledge that introduction of AMOD will alter demand, using the current demand allows us to establish a baseline about the benefits that can be expected without increasing the road traffic. This will be important when considering the relationship between policy choices, future demand scenarios and outcomes. We expect that our methodology can be readily adapted to these scenarios in future works, and comparisons with the current results will help in understanding the benefits and trade-offs of various pricing and policy decisions. In the current work, we present results for two cities (Singapore and Shenzhen, China), showing the potential of generalizability of the concept of partial integration between private and public transportation and the evaluation methodology, at least in the context of Asian metropolises.

The rest of the paper is organized as follows. Section 2 gives an overview of the datasets used in our analysis. As we are comparing bus and taxi trips for two different cities, we have four different sources describing people's mobility patterns. Section 3 describes our methodology, formally defining the matching process between trips and explaining the steps of our analysis in detail. Section 4 shows our main findings answering the two research questions we pose in this paper, i.e., the percentage of bus trips that can be matched and the average travel time saved for those matched trips. Finally, Section 5 discusses our results and shows directions for future work.

2. Data

In this paper, we compare the percentage of bus trips that can be shared with taxi trips in two cities, namely Singapore and Shenzhen, China. The taxi dataset for Singapore consists of more than 30 million of taxi trips recorded for 77 consecutive days, a total of 10 weeks. Distribution of the number of taxi trips per each day is given in Figure A1, where each day in a week is marked with a different color. What we can observe from the figure is that the number of taxi trips is always the largest on Friday and the smallest on Sunday, while between Sunday and Friday the number gradually grows. What is also evident is that there is a little of variance across the observed weeks, with an exception of the second to last week, where Thursday was a public holiday and consequently the number of taxi trips on the preceding Wednesday was higher and on Thursday itself it was significantly lower. From those 10 weeks, we chose the second week as a representative week for the further analysis. In Shenzhen, we used one week of taxi trips as well, that has 3 million trips in total, while in Singapore, we have about 2.8 million trips in the chosen week. Neither dataset includes any personal data of passengers.

Bus trips for Shenzhen were collected during the same week as taxi trips and include tap-in data for 15 million trips in total, made by 6.5 million anonymized passengers who are represented by randomly generated IDs used throughout the data collection period. Using the algorithm developed by Tu et al. [40], we were able to infer destinations for a total of 3.4 million trips that we use in the following analysis (for a more detailed description, see the Appendix A). Bus trips for Singapore were generated based on the aggregate counts of bus usage available in the DataMall public interface [41]. The data is based on smart card tap-in and tap-out events, and includes the hourly total number of bus trips made between any two bus stops in the city over the course of a month, separately counted for weekdays and weekends. In total, we have generated a bit more than 20 million bus trips for one week. We generated daily numbers based on this, assuming a Poisson distribution of individual daily counts and assigned trips in the hourly intervals based on a Poisson distribution of bus arrivals as described in more detail in the Appendix A.

In addition to the taxi and bus datasets, we also downloaded the road networks for the two cities from OpenStreetMap [42]. For Singapore, we used a bounding box that covers the whole island, and then manually excluded roads that provide connections to Malaysia, and finally, kept the largest connected component of the resulting road network, yielding the network shown in Figure 1. For Shenzhen, we used a bounding box that covers the official boundary of Shenzhen and then removed the connections to Hong Kong and kept

the largest connected component, similarly to the case of Singapore. We further processed the raw networks by performing a friend-of-friend clustering, grouping together nodes with a threshold radius of 20 m, reducing the network size to simplify processing and remove uncertainties from small errors in GPS data. After the clustering procedure, the road network for Singapore had more than 50,000 nodes and 120,000 edges, while the one for Shenzhen had almost 40,000 nodes and 100,000 edges.

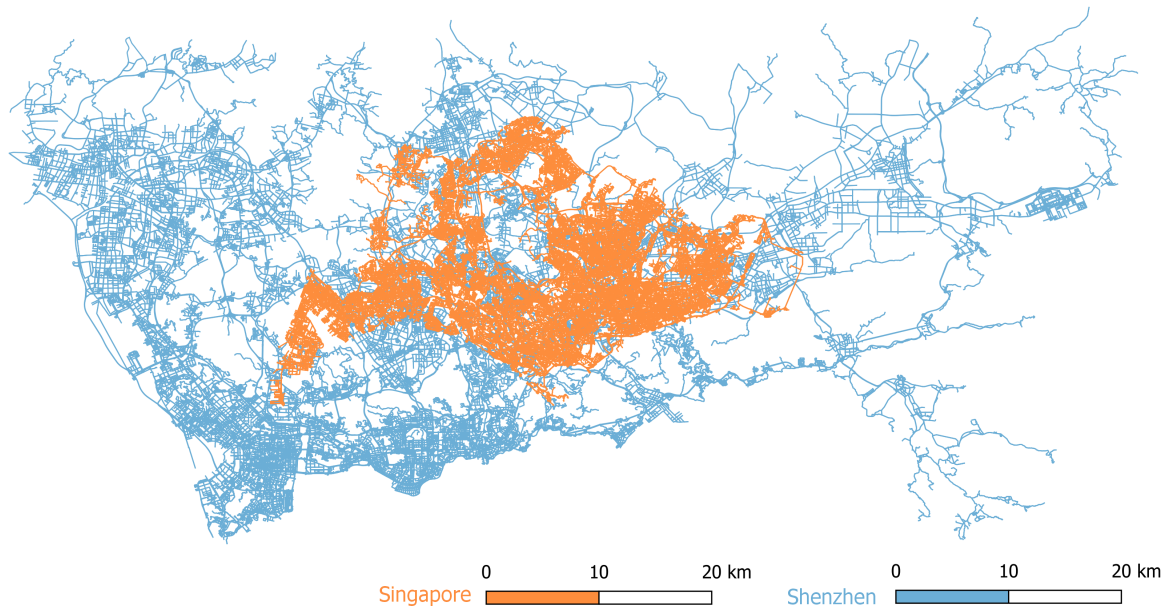


Figure 1. The road networks for Singapore and Shenzhen displayed on the same scale.

When comparing the numbers of records in Singapore and Shenzhen datasets, what we can conclude is that the number of taxi trips for both cities is comparable, i.e., around 3 million, with Singapore recording a bit less. However, due to limitation of Shenzhen bus dataset not including tap-out data, out of around 15 million records initially recorded in the dataset, we were able to use only a bit more than 25%. With that being said, what we can see is a large difference in the ratio of bus trips recorded in each city, with Singapore having around six times more bus trips than Shenzhen. However, on the other hand, in the Shenzhen dataset we have individual trips with the exact starting time, which is information that is missing from the Singapore dataset as there we only have aggregated numbers of people traveling between each pair of origin/destination bus stops within one hour. When comparing the numbers of nodes in each city's road network, we note that the area of Shenzhen is 2.8 larger than Singapore; this indicates that the road network of Shenzhen is more sparse, which is also evident in Figure 1.

3. Methodology

The taxi trip datasets for both Singapore and Shenzhen are in the same format of Global Positioning System (GPS) traces. As taxis move through the city, their geolocation (i.e., latitude and longitude) is recorded at irregular time intervals. In that sense, for each taxi trip in the dataset, there are multiple spatio-temporal points allowing us to reconstruct the route that it took. As the first step, we mapped the taxi trajectories to the road network, using the algorithm of Yang et al. [43]. The result of this procedure is an ordered list of network nodes that are present in the most likely trajectory corresponding to the trip. The advantage of this method is that we are not limited by the irregularity in recording GPS points, and we can thus identify all possible matching opportunities. A taxi trip T_i is then represented as an ordered set of tuples (n_{ij}, t_{ij}) , $j = 1, 2 \dots N_i$, where n_{ij} denotes the sequence of road network nodes identified as a part of the trajectory, t_{ij} are the estimated

timestamps for each node based on the GPS timestamps and finally N_i denotes the total number road network nodes in a trajectory T_i .

For each bus trip, we assign a set of road network nodes as *candidate sets* for the beginning and end of the trip, based on their proximity to the coordinates of the bus stop. This way, a bus trip B_i is represented as the following: $B_i = \{S_i, t_{i,s}, E_i, t_{i,e}\}$, where S_i and E_i are the candidate sets for the start and end of the trip respectively, and $t_{i,s}$ and $t_{i,e}$ are the estimated start and end times of the trip. We control the selection of the candidate sets with the parameter d that we refer to as the *spatial buffer*. For the special value of $d = 0$, the candidate sets only include the closest node to the bus stop. For $d > 0$, the candidate sets include all road network nodes within an Euclidean distance of d . In practice, we use $d = 0, 100$ m and 200 m. The significance of d is to allow a match where the pick-up does not exactly take place at the bus stop. Since the actual start of a passenger's trip is typically not the exact location of a bus stop, this buffer is interpreted in the sense that instead of going to the bus stop, a passenger would walk to a pick-up location that is within an acceptable distance of their original location.

We then compile a set of *potential matches* between bus and taxi trips, \tilde{M} , as pairs of trips where a taxi trip includes road network nodes from the start and the end candidate set of a bus trip in the correct order. We also require that the node in the bus start set is visited by the taxi within a short time interval, t_B , defined as a *time buffer* within the start of the bus trip and that the end node is visited earlier than the end of the bus trip, allowing time savings for the bus passenger. Formally, we define:

$$\tilde{M} = \{(T_i, B_j, \tau_{ij})\} \quad \forall (i, j), \quad \text{where } \exists k, l$$

$$\text{such that } \begin{cases} (1) & 1 \leq k < l \leq N_i \\ (2) & n_{ik} \in S_j \text{ and } n_{il} \in E_j \\ (3) & |t_{ik} - t_{j,s}| < t_B \\ (4) & t_{j,e} - t_{il} > 0 \\ (5) & \tau_{ij} \equiv (t_{j,e} - t_{j,s}) - (t_{il} - t_{ik}) > 0 \end{cases} \quad (1)$$

The conditions listed here guarantee that (1) road network nodes are visited in the correct order; (2) both the start and the end of the bus trip are visited by the taxi trip; (3) the bus passenger can take the taxi within a t_B temporal buffer of the start of their original trip; (4) they arrive earlier than with their original bus trip; and (5) that the actual travel time is shorter, where we define τ_{ij} as the travel time saving achieved. Note that this matching procedure does not require a bus trip origin/destination (O/D) pair to be exactly the same as a taxi trip O/D pair in order for two trips to be matched, but simply that a subset of the taxi trip matches the bus trip O/D pair.

The selection of the time buffer t_B assumes that a passenger would go to the bus stop t_B time before boarding the bus and is willing to wait up to t_B after the original departure time if matched with a taxi trip that still arrives earlier at their destination. Of course, in a more detailed model, t_B could be selected on a per-trip basis, if an estimate of the actual waiting time for the bus and individual tolerance for the extra waiting time could be established. The travel time saving τ_{ij} is defined as the *actual* travel time saving, i.e., how much faster the trip is realized with the taxi than the original bus trip. Notably, we do not include any time savings due to the taxi trip starting earlier. This likely underestimates the *total* time savings achievable for passengers, but since we do not have a reliable estimate of the waiting times for bus passengers, we chose to only focus on the part of the trip spent traveling. We acknowledge that minimizing waiting time could be an important additional goal of any combined on-demand mobility service.

Pairs in \tilde{M} represent all possible sharing opportunities and form a bipartite graph, where the τ_{ij} time savings are interpreted as edge weights. Potentially, any trip can have multiple match candidates (i.e., will be present with a degree > 1). Using this graph, we then calculate a maximal weighted matching [44,45] to arrive at an ideal assignment of trips that maximize time saved for the bus passengers while respecting the condition that

each trip can be matched only at most once. To be able to provide a tractable solution and also to limit inconvenience to the taxi passengers, we do not consider the possibility of a taxi trip being matched to multiple bus trips consecutively; if there are multiple such candidates, we choose the one that contributes to maximizing time savings globally. In future work this could be extended in such a way that instead of maximizing the total travel time savings on a system level, we could prioritize to match trips considering the purpose of the travel.

4. Results

The results of our analysis are presented with two main figures, each one answering one main research question. Figure 2 shows the percentage of bus passengers who were able to be matched with the taxi passengers, while Figure 3 shows the average travel time saved per a matched bus trip expressed in minutes. Results for the percentage of matched trips and average time savings are calculated for one hour windows during the time period of a significant bus service, between i.e., 6 a.m. and 11 p.m. in both cities; correspondingly, x -axes are limited between 6 a.m. and 10 p.m. y -axes show the percentage of matched trips in Figure 2 and the time savings in minutes in Figure 3. Each main figure is divided into eight sub-figures denoting results for Singapore and Shenzhen separately, as well as for different time buffers and for two days (Wednesday and Saturday) that represent typical results for workdays and weekends. Each panel shows the results for three different values of the space buffer, i.e., $d = 0, 100$ m and 200 m. The results for the remaining days of week are displayed in the Appendix A as Figures A6–A9.

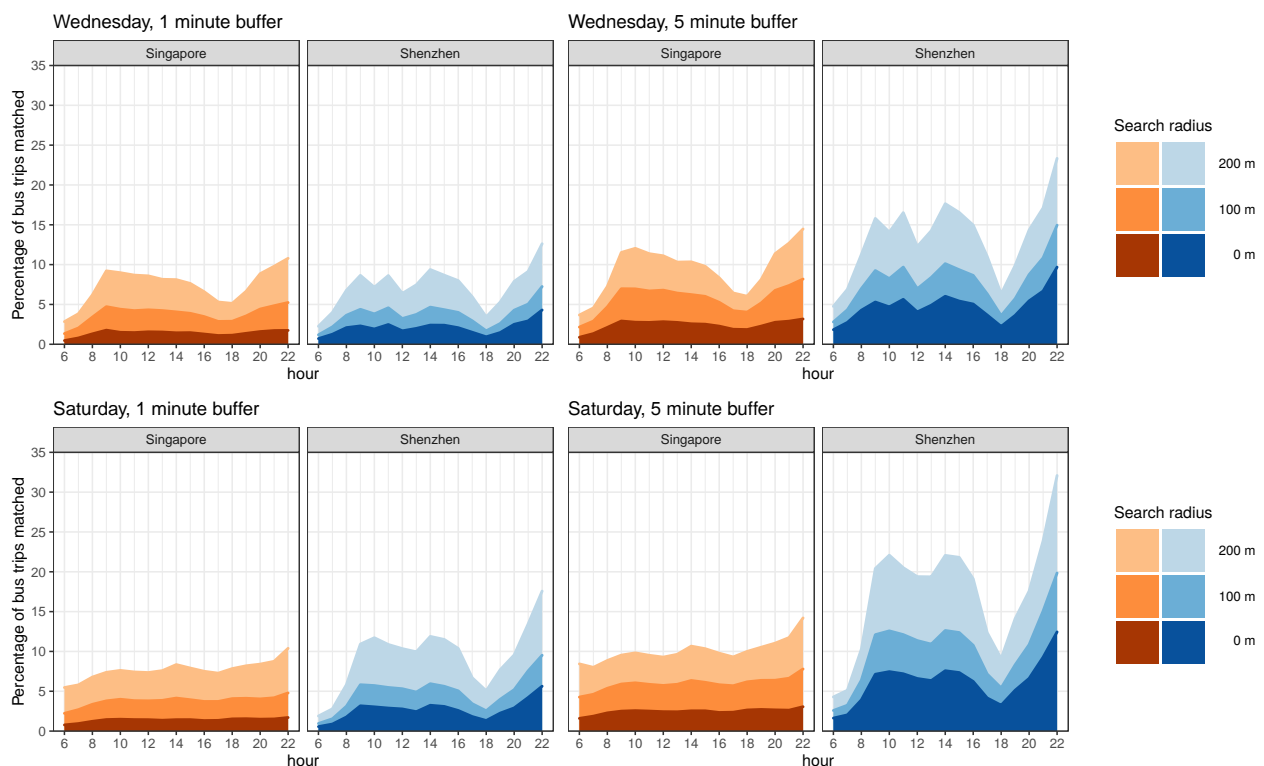


Figure 2. Ratio of bus trips that could be served as a shared trip with a taxi passenger. The results are displayed for Shenzhen and Singapore respectively in the left and right panels. The top row shows results Wednesday and the bottom row shows results for Saturday. Results for the rest of the week are presented in the Appendix A, in Figures A6 and A8. Figures on the left correspond to a time buffer of 1 min and figures on the right correspond to a time buffer of 5 min.

As expected, the percentage of bus trips matched goes up as we increase the space and time buffers. For example, if we set the time buffer to 1 min, the percentage of matched bus trips on Wednesday for Singapore for the radius of 200 m is on average a bit less than

10% in morning hours (i.e., between 9 to 12 a.m.), drops to 5% around 6 p.m. and then goes up to a bit more than 10% in the late night (i.e., around 10 p.m.). Similar, but slightly lower percentages of matched trips could be also observed for Shenzhen. However, there is a bit of different pattern during the day with two drops around 10 a.m. and midday and a much sharper drop around 6 p.m. The reason for this is the drop of the total number of taxi trips during those periods as shown in Figure A3 in the Appendix A. Namely, whereas the total number of taxi trips in Singapore stays more stable from 10 a.m. to 1 p.m., in Shenzhen there are two drops at 10 a.m. and midday. Given that the number of bus trips during the same time does not change that much either, chances for the bus passengers to share a ride with the taxi passengers are lower around 10 a.m. and 12 p.m.

For Saturday, the percentage of matched bus tips for Singapore is slightly lower than 10% and is flatter during the day, with no drop around 6 p.m. and also rising in late night hours. In Shenzhen, the percentage of bus trips matched on Saturday is slightly larger than 10%, with two drops around 1 p.m. and 6 p.m. and a large spike in late night hours. The reason for this kind of behavior is that there is a little of difference in bus distribution in Shenzhen between Wednesday and Saturday, while the number of taxi trips on Saturday is on average larger than the one on Wednesday. The same patterns are observed for 5 min time buffers, but as expected with a higher percentage of bus trips matched. However, the percentage of bus trips matched go much higher for Shenzhen than for Singapore, around 15% on Wednesday and 20% on Saturday for 200 m search radius.

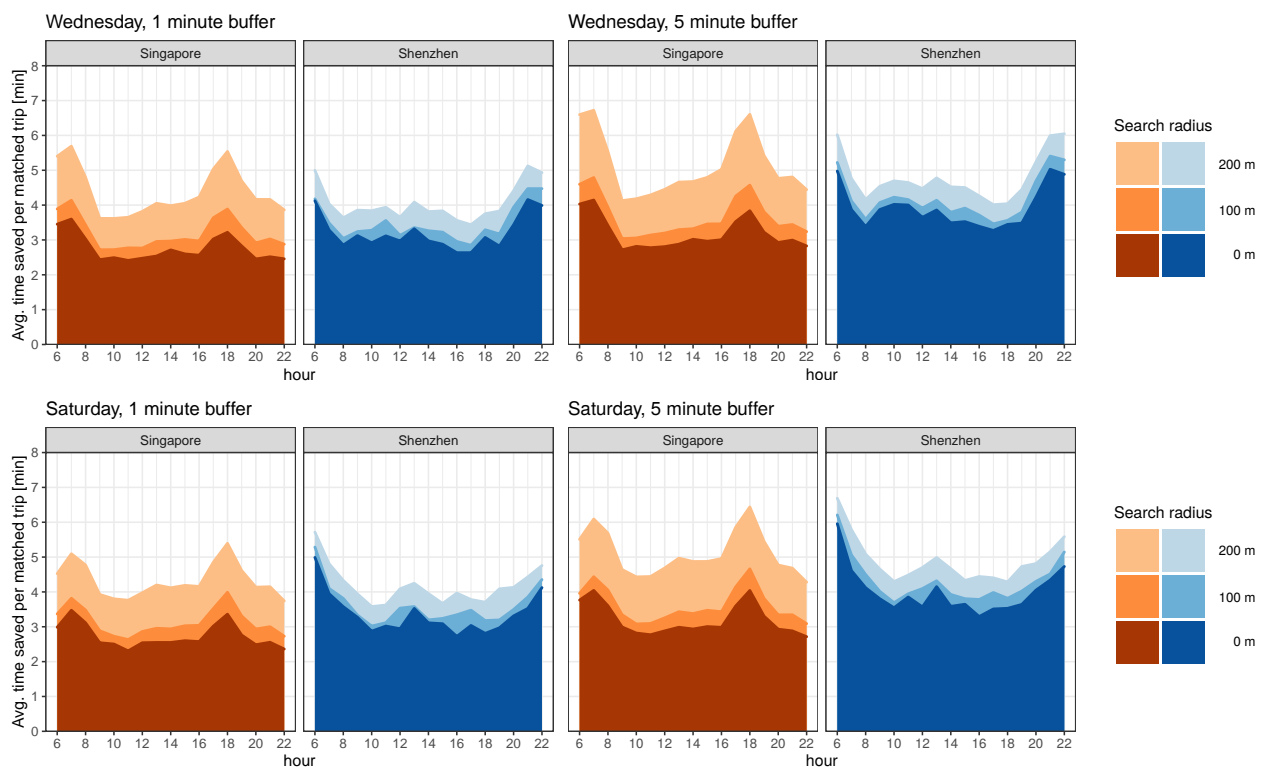


Figure 3. Average time saved by the bus passengers who are able to be served in a shared trip. Results are displayed for Shenzhen and Singapore respectively in the left and right panels. The top row shows results Wednesday and the bottom row shows results for Saturday. Results for the rest of the week are presented in the Appendix A, in Figures A7 and A9. Figures on the left correspond to a time buffer of 1 min and figures on the right correspond to a time buffer of 5 min.

When looking at the average absolute travel time saved for Singapore and Shenzhen (as illustrated in Figure 3), what we can see is that the bus passengers in Singapore who get matched with the taxi riders can save a bit more than 4 min per their trip for 1 min time buffer and 200 m space buffer. At the same time, the time savings in Shenzhen (with the same parameters) on average are a bit less than 4 min. The time savings with 5 min

time radius and 200 m space buffer can go over 5 min in the case of Singapore and over 4 min in the case of Shenzhen. This also indicates that the increased sharing opportunities contribute to more travel time savings even without including the effect of potentially starting trips earlier that is possible with a larger t_B time buffer. Regarding time patterns, in Singapore we can see larger time savings around 6 a.m. and 6 p.m. for both Wednesday and Saturday, while the pattern for Shenzhen is flatter and also does not show a lot of variety between a workday and weekend. This is an interesting finding as when comparing the average bus trip lengths of Singapore and Shenzhen (as shown in Figure A5), we can see that on average, bus trips in Singapore are shorter (i.e., mean for Singapore is around 10 min and for Shenzhen 14 min). This is understandable given that the area of Shenzhen is significantly larger. Consequently, when putting time savings into a perspective, this means that on average, absolute travel time saved in Singapore is up to 50% of mean bus duration and around 30% for Shenzhen.

5. Discussion

Until very recently, private and public transportation have been two systems that were very much separated [1,3]. However, with the emergence of Transportation Network Companies (TNCs) and the Mobility-as-a-Service (MaaS) concepts, those lines are becoming unclear [35,37,46]. This will be even more evident once Autonomous Vehicles (AVs) hit the roads bringing further disruptive changes to urban mobility [11,13,16]. If drivers' costs are removed from the equation, reduced fares will likely attract further passengers [14,15]; this in turn can result in negative externalities, such as increased congestion and energy use. This way, governments and operators will need to present new, innovate mobility concepts that allow exploiting the shareability potential that is present in the current demand and trips, including dynamic integration of private and public transportation modes in a sustainable way.

In this paper, we thus analyzed the first step of exploiting parallels between public and private transportation using today's travel demand. Using mobility patterns recorded by taxi companies and bus operators in Singapore and Shenzhen, we investigated how passengers on public transportation could reduce their travel time if paired with already existing taxi riders. While we expect that such an arrangement will only make economic sense in a future where AVs allow drastically reduced operating costs, our results show the validity of the concept, and the need to explore such possibilities in scenario-based analysis of future transportation systems as well.

The results of our analysis showed that 10–20% of the bus passengers could be potentially matched with the taxi riders, which would contribute to on average between 4 to 6 min of savings of their travel time. These results are consistent across two cities in Asia, although there are individual differences in the temporal pattern of matching ratios. The main source of difference is explained by how the total volume of taxi and bus trips changes during the day. Namely, first we see a clear difference between the distribution of total bus trips in Singapore between a weekday and a weekend, whereas this difference is less obvious in case of Shenzhen. This possibly means that in Shenzhen, more people also work on Saturdays. Second, the total amount of taxi trips on Saturday in Shenzhen is larger than on Wednesday, which is not true for Singapore.

In conclusion, our analysis shows that there is a practical potential for partial integration of public and private transportation even under the current conditions. This is an important first step when envisioning a future where AV technology allows a variety of novel transportation service types [21,27,38,39]. Researchers and policymakers can build on our work when considering a scenario-based analysis of future mobility to evaluate policy choices [13,32]; including dynamic, opportunistic sharing of rides between public and private transportation users could increase flexibility and passenger convenience without additional externalities, and thus can be an important component of any integrated mobility offering of the future.

Extensions of the current work could look into incorporating predicted demand in presumed future scenarios of regulatory and pricing environment and mode choice models that incorporate new mobility modes [47]. This will allow a more comprehensive evaluation of benefits, and inform about strategies that governments can employ to ensure sustainable urban transportation in the age of autonomous driving. Furthermore, incorporating a more detailed model of passenger choices can account for how differences in trip purpose affect the value of time and willingness to take part of such shared transportation options [47]; this in turn will inform on how the implementation of opportunistic ride-sharing can be optimized to cater more closely to the subset of public transportation users who are most eager to use it, and where maximal benefits can be realized [1,39].

Lastly, extensions of the current work could consider other form factors, such as medium sized (i.e., 6–10 passenger) vehicles as alternatives to both taxis and buses and evaluate the total impact in terms of congestion, energy use, passenger comfort and convenience. We note that previous work in this area was limited to using taxi trips as an estimate of demand [21,47]; our results show the importance of including the *complete* picture of urban transportation demand, i.e., both public and private transportation users. Furthermore, while our work shows a potential for matching trips, any such service will face challenges in implementing user interaction solutions that can be conveniently used without excluding groups of users, e.g. those who do not use a smartphone. This means that investigating new user interaction concepts for the fluid transportation services of tomorrow will become increasingly important as more and more optimization opportunities in on-demand transportation are realized.

Author Contributions: Conceptualization, I.B. and D.K.; methodology, I.B. and D.K.; data analysis, I.B., D.K., W.T. and K.M.; validation, I.B. and D.K.; investigation, I.B. and D.K.; data curation, I.B., D.K., D.K. and W.T.; writing—original draft preparation, I.B. and D.K.; writing—review and editing, I.B., D.K. and P.S.; supervision, C.R. and P.S.; project administration, P.S.; funding acquisition, C.R. and W.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Singapore Ministry of National Development and the National Research Foundation, Prime Minister’s Office, under the Singapore-MIT Alliance for Research and Technology (SMART) programme. We also thank RATP, Dover Corporation, Allianz, Teck Resources, Lab Campus, Anas S.p.A., Ford, ENEL Foundation, the Amsterdam Institute for Advanced Metropolitan Solutions, the cities of Laval, Curitiba, Stockholm and Amsterdam, and all of the members of the MIT Senseable City Laboratory Consortium for supporting this research. Wei Tu acknowledges the funding of Nature Science Foundation of China (No. 4207010598).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Supplementary Figures

Figure A1 shows the number of taxi trips in Singapore for each of 77 consecutive days recorded in our dataset. Figure A3 shows distribution of the number of taxi and bus trips per hour on Wednesday in Singapore and Shenzhen, while Figure A4 presents the same distributions, but for Saturday. Figure A5 shows distribution of bus trip duration per hour on Wednesday and Saturday in Singapore and Shenzhen. Figures A6 and A8 show the percentage of shareable trips for the rest of the weekdays that were not included in Figure 2 in the main text (i.e., Monday, Tuesday, Thursday, Friday and Sunday). Figures A7 and A9 show the average time savings for bus passengers on these days.

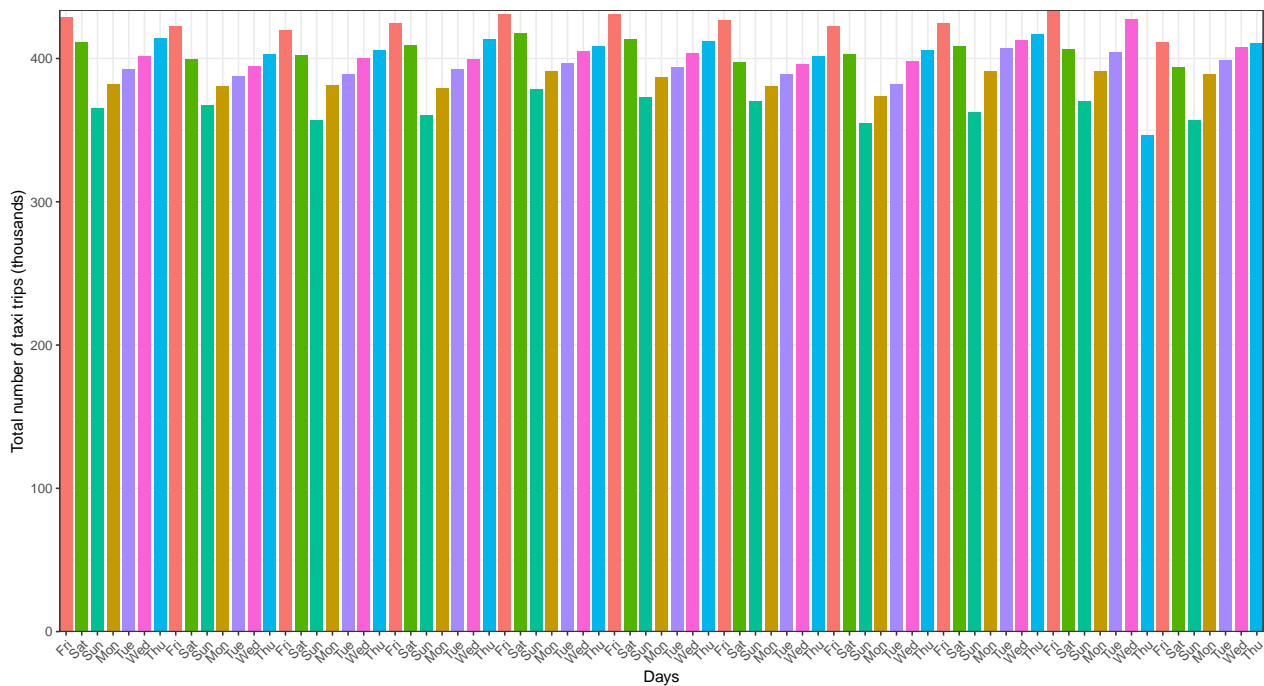


Figure A1. The number of taxi trips in Singapore per day.

Appendix A.2. Bus Trip Processing for Shenzhen

For Shenzhen, we fused bus GPS trajectories and Smart Card Data (SCD) to infer bus trips destinations. We inferred the alighting location and time by considering the spatial-temporal regularity of the SCD user. Firstly, the boarding location of a smart card user was inferred. The continuous bus GPS trajectory was recovered from GPS records by map matching considering the road network and the delay at the road crossing [40]. Each bus SCD record was linked to the corresponding trajectory based on the bus identification number. The recorded time was then used to interpolate the boarding location from the GPS trajectory. Following the direction of the bus route, we adjusted the boarding location to the nearest bus stop in the bus route. Then, the alighting location of the SCD record was inferred. The SCD user with a pair with the highly frequent boarding locations were filtered [48]. Considering the regularity of commuters, the following highly frequently boarding location was recognized as the alighting location of the preceding SCD record. The alighting time was interpolated according to the corresponding continuous GPS trajectory.

Appendix A.3. Bus Trip Generation for Singapore

The bus dataset includes hourly counts of trips between any two bus stops in Singapore over the course of one month, aggregated separately for weekdays and weekends. As a first step, we calculated the average hourly counts for each day, by dividing the total numbers with the number of weekdays and weekend days (including public holidays), i.e., 22 and 9 respectively. Next, we generated realizations of the actual number of bus trips between each bus stop pair by sampling from a Poisson distribution with the mean given by the average counts. We repeated this process for each bus stop pair and for each hour of the day when buses are operating. We then distributed trip start times in the one-hour intervals based on a process where a given number of individually distributed bus departures were assumed for each bus stop in each hour. To achieve this, we counted the total hourly passenger counts for each bus stop, denoting by N_{ij} the number of passengers boarding

a bus at the i th stop of the j th hour. We then assumed that the expected number of bus departures in a stop was related to the number of passenger boarding in the following way:

$$B_{ij} = \begin{cases} N_{ij}^\alpha & \text{if } N_{ij} < N^* \\ B^* + (N_{ij} - N^*)/b_0 & \text{if } N_{ij} \geq N^* \end{cases} \quad (\text{A1})$$

Note that we defined the relation in Equation (A1) based on our previous work with the bus travel data. In Equation (A1), we used $B^* \equiv (N^*)^\alpha$ and the numerical parameters were $\alpha = 0.7$, $N^* = 200$ and $b_0 = 40$. We displayed the relationship between B_{ij} and N_{ij} in Figure A2. Using this relationship, we assigned B_{ij} as the expected number of bus departures for every bus stop and hour and then generated an actual number of departures from a Poisson distribution with B_{ij} mean and excluding the case when this random choice would give a zero value. As the last step, we distributed departure times within the one-hour interval among the bus departures using an exponential distribution and normalizing the total elapsed time and assigned each passenger randomly among the buses, using the departure time of the selected bus as the trip start time.

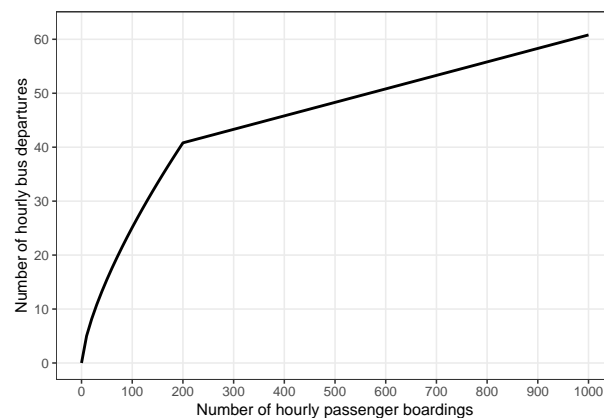


Figure A2. Empirically motivated relationship between the hourly number of bus passengers in a stop and the expected number of bus departures, i.e., the relationship described in Equation (A1).

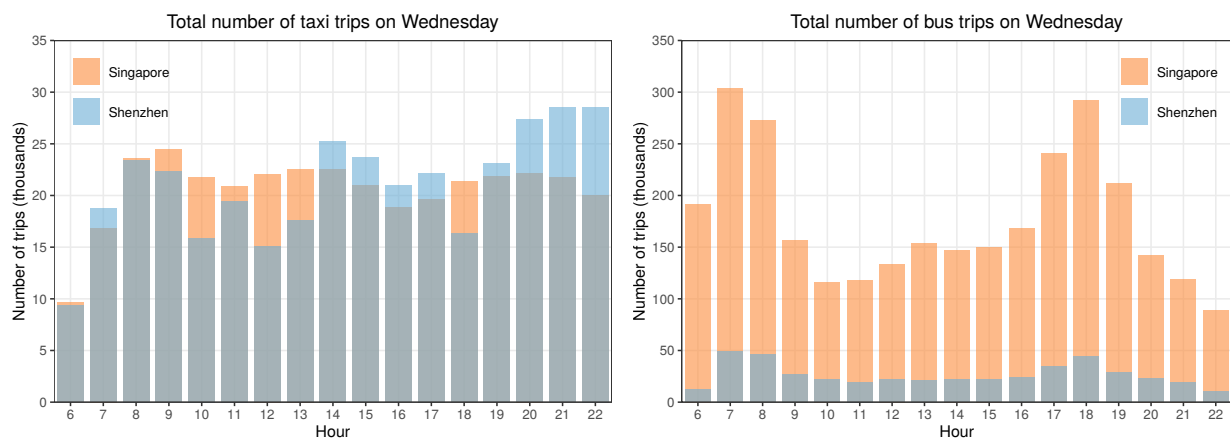


Figure A3. Comparing the total number of taxi and bus trips on Wednesday in two cities.

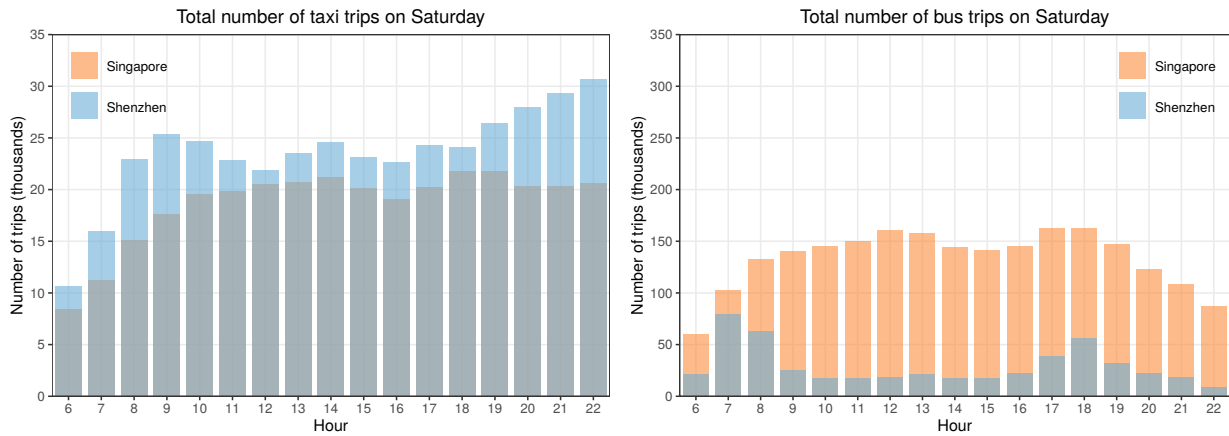


Figure A4. Comparing the total number of taxi and bus trips on Saturday in two cities.

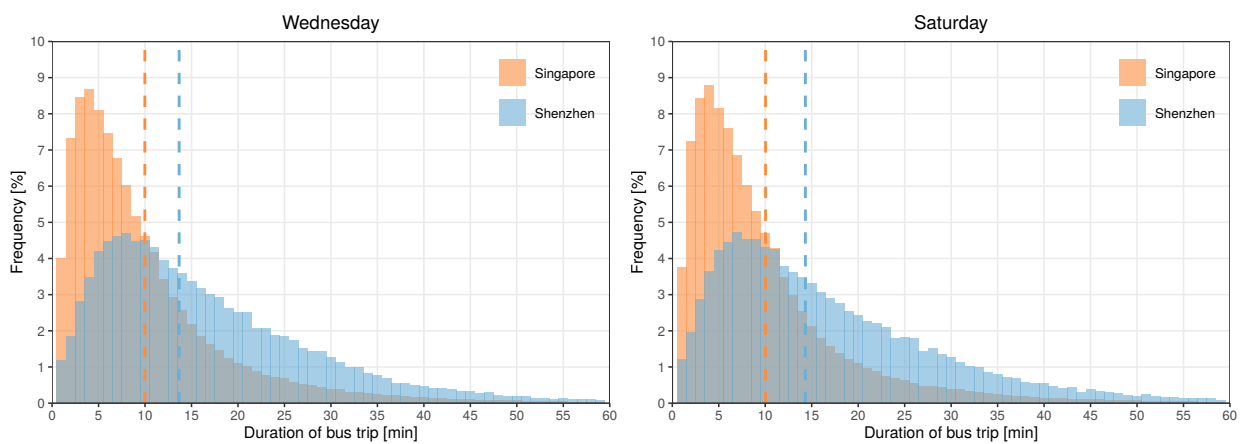


Figure A5. Comparing duration of bus trips on Wednesday and Saturday in two cities. Dashed lines display the mean of the distributions, which are 9.98 min and 10.03 min in the case of Singapore for Wednesday and Saturday respectively, and 13.67 min and 14.32 min in the case of Shenzhen.

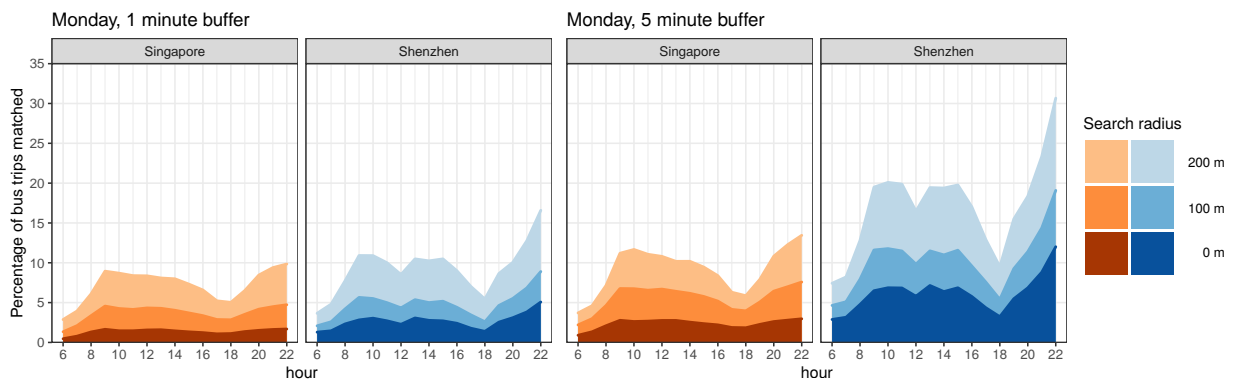


Figure A6. Cont.

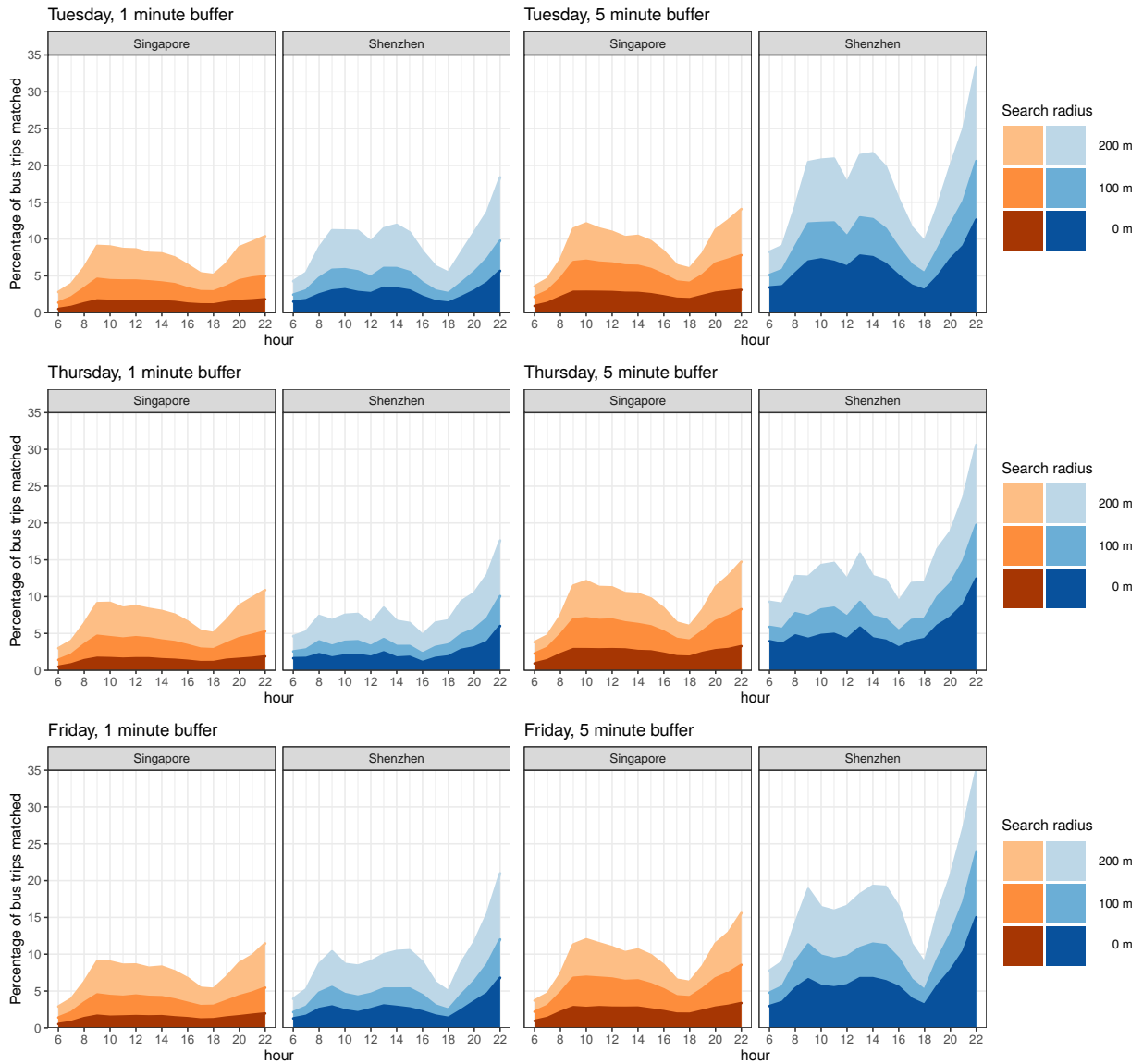


Figure A6. Percentage of bus trips shareable for weekdays.

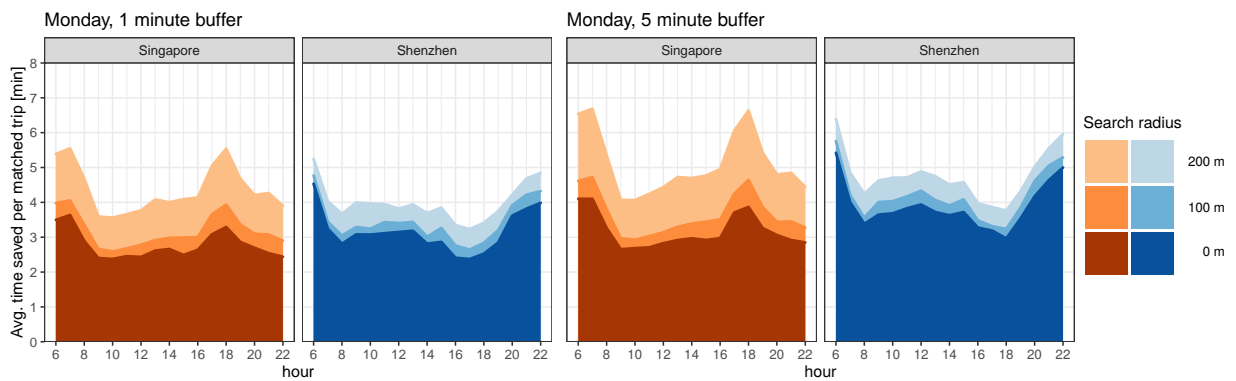


Figure A7. Cont.

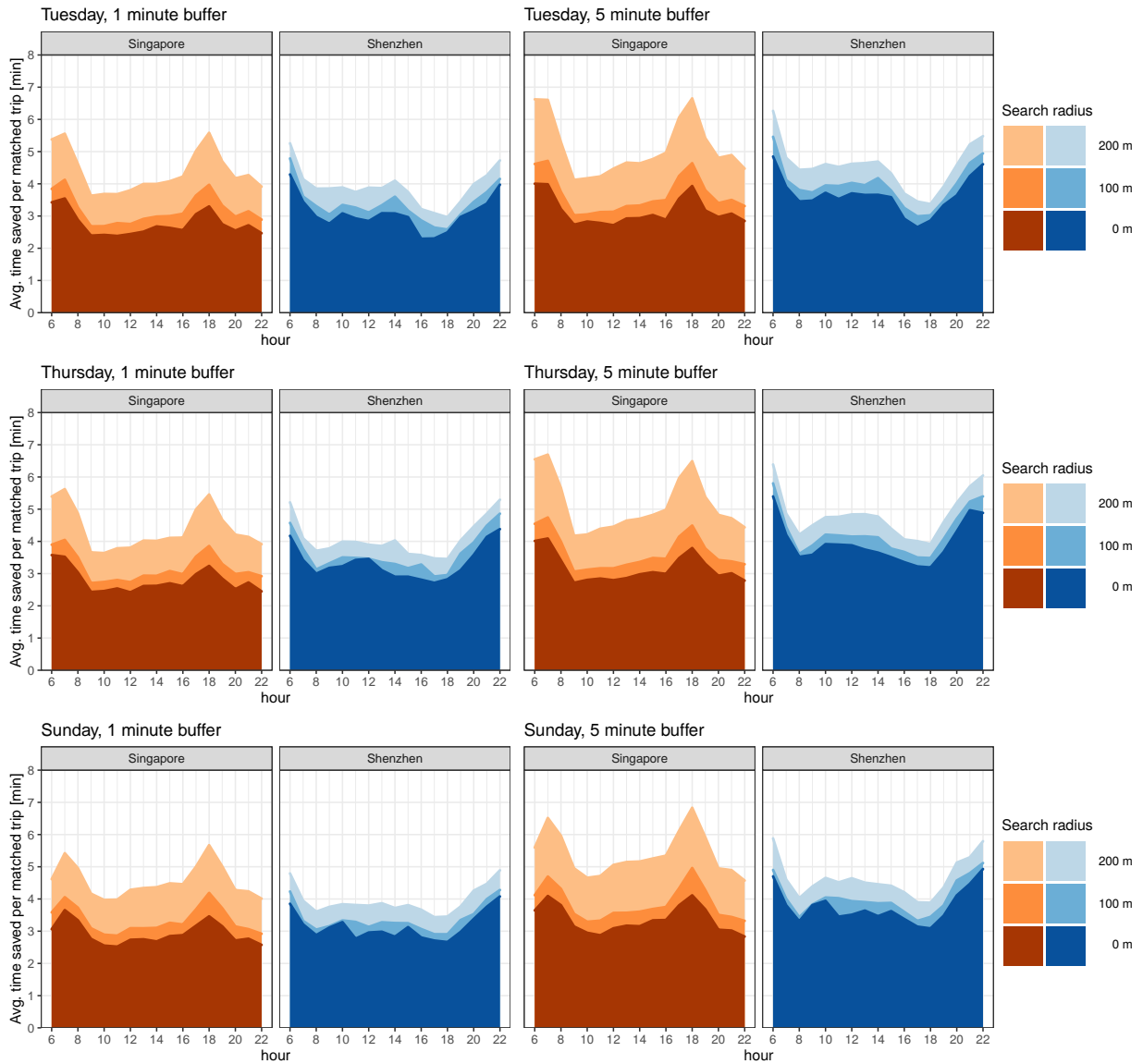


Figure A7. Average time saved for shareable trips on weekdays.

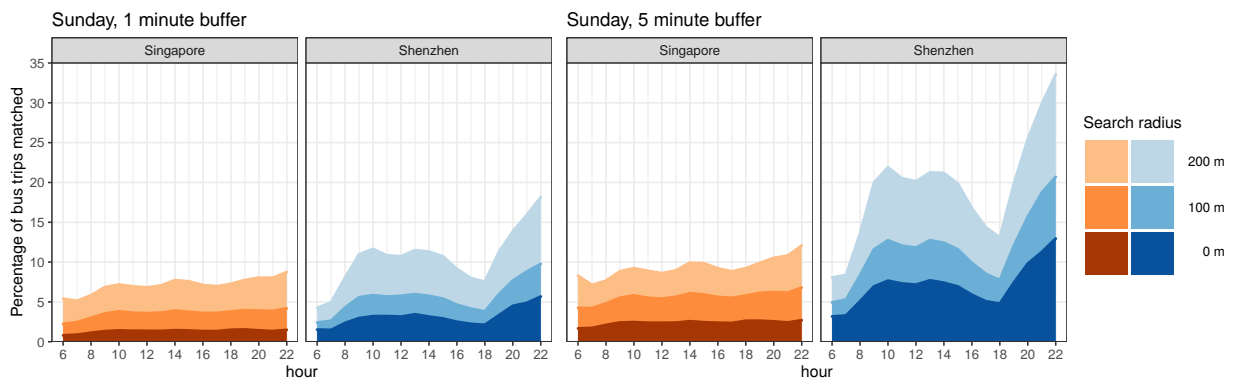


Figure A8. Percentage of bus trips shareable for Sunday.

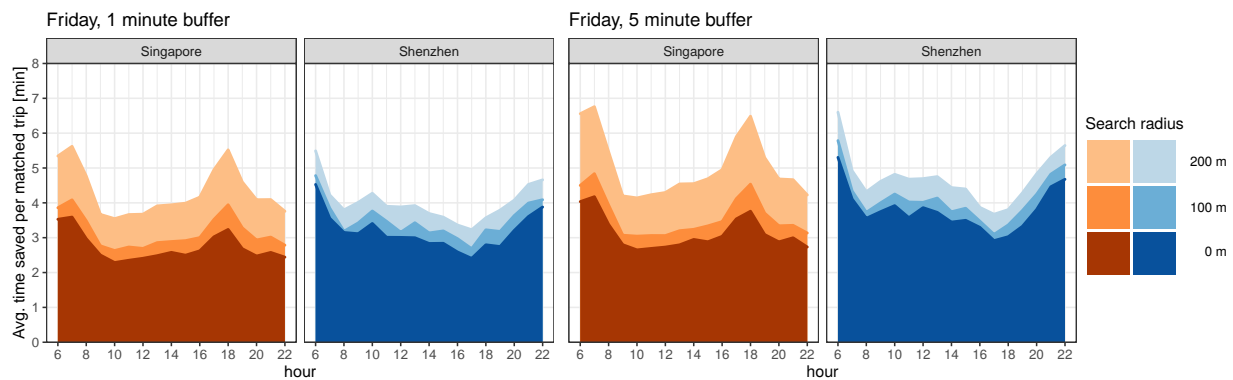


Figure A9. Average time saved for shareable trips on Sunday.

References

- Jacques, C.; Manaugh, K.; El-Geneidy, A.M. Rescuing the captive [mode] user: An alternative approach to transport market segmentation. *Transportation* **2013**, *40*, 625–645. [\[CrossRef\]](#)
- Verbavatz, V.; Barthélemy, M. Critical factors for mitigating car traffic in cities. *PLoS ONE* **2019**, *14*, e0219559. [\[CrossRef\]](#) [\[PubMed\]](#)
- Jakob, A.; Craig, J.L.; Fisher, G. Transport cost analysis: a case study of the total costs of private and public transport in Auckland. *Environ. Sci. Policy* **2006**, *9*, 55–66. [\[CrossRef\]](#)
- Lindau, L.A.; Hidalgo, D.; Facchini, D. Curitiba, the cradle of bus rapid transit. *Built Environ.* **2010**, *36*, 274–282. [\[CrossRef\]](#)
- Li, S.; Poolla, K.; Varaiya, P. Congestion Surcharge and Wage Regulation on TNCs: A Case Study for San Francisco. *arXiv* **2020**, arXiv:2003.02550.
- Anderson, D.N. “Not just a taxi”? For-profit ridesharing, driver strategies, and VMT. *Transportation* **2014**, *41*, 1099–1117. [\[CrossRef\]](#)
- Glöss, M.; McGregor, M.; Brown, B. Designing for labour: uber and the on-demand mobile workforce. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 1632–1643.
- Feigon, S.; Murphy, C. *Shared Mobility and the Transformation of Public Transit*; National Academies of Sciences, Engineering, and Medicine: Washington, DC, USA, 2016.
- Rayle, L.; Dai, D.; Chan, N.; Cervero, R.; Shaheen, S. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transp. Policy* **2016**, *45*, 168–178. [\[CrossRef\]](#)
- Hawas, Y.E.; Hassan, M.N.; Abulibdeh, A. A multi-criteria approach of assessing public transport accessibility at a strategic level. *J. Transp. Geogr.* **2016**, *57*, 19–34. [\[CrossRef\]](#)
- Fagnant, D.J.; Kockelman, K. Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transp. Res. Part A Policy Pract.* **2015**, *77*, 167–181. [\[CrossRef\]](#)
- OECD International Transport Forum. *Urban Mobility System Upgrade: How Shared Self-Driving Cars Could Change City Traffic*; Technical Report; International Transport Forum: Paris, France, 2015.
- Stead, D.; Vaddadi, B. Automated vehicles and how they may affect urban form: A review of recent scenario studies. *Cities* **2019**, *92*, 125–133. [\[CrossRef\]](#)
- Burns, L.D.; Jordan, W.C.; Scarborough, B.A. *Transforming Personal Mobility*; Technical Report; The Earth Institute, Columbia University: New York, NY, USA, 2013.
- Brownell, C.; Kornhauser, A. A Driverless Alternative Fleet Size and Cost Requirements for a Statewide Autonomous Taxi Network in New Jersey. *Transp. Res. Rec.* **2014**, *2416*, 73–81. [\[CrossRef\]](#)
- Smith, B.W. Managing Autonomous Transportation Demand. *Santa Clara Law Rev.* **2012**, *52*, 1401–1422. [\[CrossRef\]](#)
- Tachet, R.; Santi, P.; Sobolevsky, S.; Reyes-castro, L.I.; Frazzoli, E.; Helbing, D.; Ratti, C. Revisiting Street Intersections Using Slot-Based Systems. *PLoS ONE* **2016**, *11*, e0149607. [\[CrossRef\]](#) [\[PubMed\]](#)
- Greenblatt, J.B.; Saxena, S. Autonomous taxis could greatly reduce greenhouse-gas emissions of US light-duty vehicles. *Nat. Clim. Chang.* **2015**, *5*, 860–863. [\[CrossRef\]](#)
- Nealer, R.; Hendrickson, T.P. Review of Recent Lifecycle Assessments of Energy and Greenhouse Gas Emissions for Electric Vehicles. *Curr. Sustain./Renew. Energy Rep.* **2015**, *2*, 66–73. [\[CrossRef\]](#)
- Christie, D.; Koymans, A.; Chanard, T.; Lasgouttes, J.M.; Kaufmann, V. Pioneering Driverless Electric Vehicles in Europe: The City Automated Transport System (CATS). *Transp. Res. Procedia* **2016**, *13*, 30–39. [\[CrossRef\]](#)
- Alonso-Mora, J.; Samaranayake, S.; Wallar, A.; Frazzoli, E.; Rus, D. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 462–467. [\[CrossRef\]](#) [\[PubMed\]](#)
- Harper, C.D.; Hendrickson, C.T.; Mangones, S.; Samaras, C. Estimating potential increases in travel with autonomous vehicles for the non-driving, elderly and people with travel-restrictive medical conditions. *Transp. Res. Part C Emerg. Technol.* **2016**, *72*, 1–9. [\[CrossRef\]](#)

23. Daziano, R.A.; Sarrias, M.; Leard, B. Are consumers willing to pay to let cars drive for them? Analyzing response to autonomous vehicles. *Transp. Res. Part C Emerg. Technol.* **2017**, *78*, 150–164. [[CrossRef](#)]
24. Bansal, P.; Kockelman, K.M. Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies. *Transp. Res. Part A Policy Pract.* **2017**, *95*, 49–63. [[CrossRef](#)]
25. Millard-Ball, A. The autonomous vehicle parking problem. *Transp. Policy* **2019**, *75*, 99–108. [[CrossRef](#)]
26. Shen, Y.; Zhang, H.; Zhao, J. Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in Singapore. *Transp. Res. Part A Policy Pract.* **2018**, *113*, 125–136. [[CrossRef](#)]
27. Wen, J.; Chen, Y.X.; Nassir, N.; Zhao, J. Transit-oriented autonomous vehicle operation with integrated demand-supply interaction. *Transp. Res. Part C Emerg. Technol.* **2018**, *97*, 216–234. [[CrossRef](#)]
28. New York State's Congestion Surcharge. 2019. Available online: <https://www1.nyc.gov/site/tlc/about/congestion-surcharge.page> (accessed on 14 February 2021).
29. San Francisco's Transportation Network Companies Tax. 2019. Available online: <https://www.sfcta.org/funding/tnc-tax> (accessed on 14 February 2021).
30. City of Chicago Congestion Pricing. 2020. Available online: <https://www.chicago.gov/city/en/depts/bacp/supp-info/city-of-chicago-congestion-pricing.html> (accessed on 14 February 2021).
31. Salazar, M.; Rossi, F.; Schiffer, M.; Onder, C.H.; Pavone, M. On the Interaction between Autonomous Mobility-on-Demand and Public Transportation Systems. In Proceedings of the 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 2262–2269. [[CrossRef](#)]
32. Milakis, D.; Van Arem, B.; Van Wee, B. Policy and society related implications of automated driving: A review of literature and directions for future research. *J. Intell. Transp. Syst. Technol. Plan. Oper.* **2017**, *21*, 324–348. [[CrossRef](#)]
33. Hensher, D.A. Tackling road congestion—What might it look like in the future under a collaborative and connected mobility model? *Transp. Policy* **2018**, *66*, A1–A8. [[CrossRef](#)]
34. Rantasila, K. The impact of Mobility as a Service concept to land use in Finnish context. In Proceedings of the 2015 International Conference on Sustainable Mobility Applications, Renewables and Technology (SMART), Kuwait, Kuwait, 23–25 November 2015; pp. 1–7.
35. Smith, G.; Sochor, J.; Karlsson, I.M. Mobility as a Service: Development scenarios and implications for public transport. *Res. Transp. Econ.* **2018**, *69*, 592–599. [[CrossRef](#)]
36. Heikkilä, S. Mobility as a Service—A proposal for action for the public administration, Case Helsinki. Master's Thesis, Aalto University, Espoo, Finland, 2014.
37. Schikofsky, J.; Dannewald, T.; Kowald, M. Exploring motivational mechanisms behind the intention to adopt mobility as a service (MaaS): Insights from Germany. *Transp. Res. Part A Policy Pract.* **2020**, *131*, 296–312. [[CrossRef](#)]
38. Meghjani, M.; Pendleton, S.D.; Marczuk, K.A.; Eng, Y.H.; Shen, X.; Ang, M.H.J.; Rus, D. Multi-class Fleet Sizing and Mobility on Demand Service. *Complex Syst. Des. Manag. Asia* **2018**, *2018*, 37–49. [[CrossRef](#)]
39. Bian, Z.; Liu, X. Mechanism design for first-mile ridesharing based on personalized requirements part I: Theoretical analysis in generalized scenarios. *Transp. Res. Part B Methodol.* **2019**, *120*, 147–171. [[CrossRef](#)]
40. Tu, W.; Cao, R.; Yue, Y.; Zhou, B.; Li, Q.; Li, Q. Spatial variations in urban public ridership derived from GPS trajectories and smart card data. *J. Transp. Geogr.* **2018**, *69*, 45–57. [[CrossRef](#)]
41. DataMall. (Web Interface). Bus Data for January 2019 was Downloaded on 13 February 2019. Available online: <https://www.mytransport.sg/content/mytransport/home/dataMall/dynamic-data.html> (accessed on 1 November 2019) .
42. Haklay, M.; Weber, P. Openstreetmap: User-generated street maps. *IEEE Pervasive Comput.* **2008**, *7*, 12–18. [[CrossRef](#)]
43. Yang, C.; Gidófalvi, G. Fast map matching, an algorithm integrating hidden Markov model with precomputation. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 547–570. [[CrossRef](#)]
44. Edmonds, J. Paths, trees, and flowers. *Can. J. Math.* **1965**, *17*, 449–467. [[CrossRef](#)]
45. Dezso, B.; Jüttner, A.; Kovács, P. LEMON - An open source C++ graph template library. *Electron. Notes Theor. Comput. Sci.* **2011**, *264*, 23–45. [[CrossRef](#)]
46. Jin, S.T.; Kong, H.; Sui, D.Z. Uber, public transit, and urban transportation equity: A case study in new york city. *Prof. Geogr.* **2019**, *71*, 315–330. [[CrossRef](#)]
47. Oh, S.; Seshadri, R.; Le, D.T.; Zegras, P.C.; Ben-Akiva, M.E. Evaluating Automated Demand Responsive Transit Using Microsimulation. *IEEE Access* **2020**, *8*, 82551–82561. [[CrossRef](#)]
48. Zhong, C.; Manley, E.; Arisona, S.M.; Batty, M.; Schmitt, G. Measuring variability of mobility patterns from multiday smart-card data. *J. Comput. Sci.* **2015**, *9*, 125–130. [[CrossRef](#)]