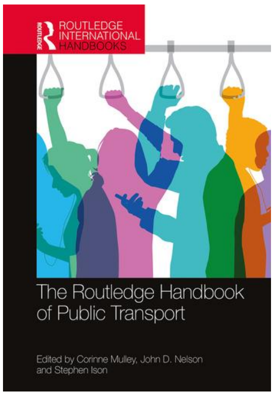

The Routledge Handbook of Public Transport

Corinne Mulley, John D. Nelson, Stephen Ison

Smart card data and its use in public transport research

Publication details
https://test.routledgehandbooks.com/doi/10.4324/9780367816698-38
Chinh Q. Ho
Published online on: 13 May 2021

33

SMART CARD DATA AND ITS USE IN PUBLIC TRANSPORT RESEARCH

An overview

Chinh Q. Ho

Introduction

Smart card technology has been used by many public transport agencies to collect fares automatically, and thus smart card systems are often referred to as automatic fare collection (AFC) systems. This technology is not new and can be traced back to the late 1960s (Shelfer & Procaccino, 2002), although its first use in public transport systems to replace existing paper-based tickets began in the late 1990s with South Korea's UPASS and Hong Kong's Octopus card (Ellison et al., 2017). While smart cards are used in the public transport domain with the main purpose of collecting fares, they also produce a large amount of data on passenger movement activities such as where and when travellers check in to (or tap on) and/or check out from (or tap off) the transport network. The data registered to a smart card system, which follows a certain specification, such as the ITSO Specification (see www.itso.org.uk/), carry valuable information, providing the evidence for innovations in the transport industry. This is because these data, referred to as smart card data hereafter, can be used for many purposes, ranging from studying day-to-day travel behaviour of the passengers, to assessing the efficiency/performance of a transport network, to forecasting travel demand in responses to some changes to the network, such as the addition of new routes or changes to service timetable. This chapter provides an overview of the use of smart card data in transport fields to set the scene for the series of case study chapters that follow.

The use of smart card data in public transport research

Smart card data have been exploited by researchers from various angles to identify opportunities for marketing, customer services, network optimisation, network planning, service adjustment and even enriching other datasets, such as household travel surveys typically found in many countries/cities around the world, or, more recently, mobile tower data. An advanced search on Scopus (March 2020) with two keywords "smart*card data" AND "transport*" returns 5,552 articles in English. Thus, smart card data underpin a large body of research in transport, and it is necessary to adopt some taxonomy to classify the literature into different categories for this overview chapter. There are at least two comprehensive reviews of the use of smart card data in transport, namely Pelletier et al. (2011) and Welch and Widita (2019). Following the former study, this chapter classifies recent studies that use smart card data into three groups using the strategic, tactical and operational (STO) framework that emerged in 1995 (van de Velde, 1999) from the international conference series on Competition and Ownership in Land Passenger Transport (a.k.a. Thredbo series). This STO framework is selected as a meaningful way to classify studies that use smart card data based on the time horizon of their contributions being long term (strategic), medium term (tactical) or short term (operational).

Research with strategic contributions

Generally, studies in this group use smart card data to address long-term planning by first understanding travel patterns and then performing demand forecasts for the purposes of planning, monitoring, marketing or retaining customers. Studies that use smart card data to obtain a better understanding of travel patterns are many, with example analyses including types of travellers such as regular vs. irregular users and their habits (Agard et al., 2006; He et al., 2020; Ortega-Tong, 2013). Insights gained from such analysis can be used in targeted marketing campaigns to retain customers in particular groups (see, for example, Bagchi & White, 2005) or to make long-term adjustment to services by adapting the network to user needs (Chu & Chapleau, 2008; Park et al., 2008; Utsunomiya et al., 2006). The latter group includes studies that aim to estimate origin-destination matrices from the smart-card data for long-term travel demand forecasts (Alsger, Assemi et al., 2016; Alsger, Mesbah et al., 2015; Luo et al., 2017).

Although not directly contributing to strategic planning, smart card data are used to improve the accuracy of traditional survey data such as household travel surveys (HTSs), a critical input for many, if not all, strategic travel models used by transport authorities around the world. Although published work that uses smart card data to enrich HTS data is limited (Trépanier et al., 2009), this is now regarded as the state of the practice by many transport authorities, such as Transport for NSW and Transport for London (Transport Systems Catapult, 2017), which use smart card data to modernise their HTS data to obtain a representative sample for their strategic travel models.

Research with tactical contributions

Research in this group analyses smart card data to obtain insights on public transport service levels, which in turn forms the basis for various improvements to service and network planning. A typical example study in this group relates to the temporal and spatial variations in ridership (i.e., patronage variation across days of the week and across routes within the same network). The tactical aims of these studies are to adjust public transport services according to the travel demand. Example studies are Utsunomiya et al. (2006) and Trépanier et al. (2009), who respectively found that ridership varied substantially across weekdays and routes while scheduled services remained the same. This forms the basis for their tactical suggestions of using different timetables for these days and routes.

One of the most frequent topics of research at the tactical level relates to the identification of transfer locations and destinations from the smart card data. This line of research emerged in the early days of smart card data as a result of many AFC systems that did not require the users to tap off when they use public transport services. A typical example system is Transantiago, which serves Santiago, the capital of Chile, and uses a flat fare where smart card users only need to tap on (Gschwender et al., 2016). In such a system, smart card data do not include information on alighting station or time. Consequently, a prerequisite task in making use of the smart card data is to infer the trip destinations. Different methods have been proposed in the literature for this, and the problem of destination estimation for smart card data is so common that this topic even has a comprehensive review (Singh et al., 2017). One example algorithm was proposed by Trépanier et al. (2009), who inferred the destination for each trip by finding the boarding point (i.e., origin) of the next trip or comparing the trip under investigation with similar trips from the same cardholder recorded in the smart card database. Another example algorithm is described in Munizaga et al. (2010) who relies on two basic assumptions to derive what is now known as the trip-chaining model for the estimation of trip destinations. The first assumption states that the alighting point is most likely the boarding point of the next trip. The second assumption requires that the last alighting point of the day is most likely the same as the origin of the first trip undertaken on that day, implicitly assuming that everyone returns to their home by the end of the day. Alternative models used in the literature include probability models and machine learning with deep learning (see Singh et al., 2017 for details).

More recent tactical research on smart card data appear to tackle a more complex methodological problem, namely which particular service (i.e. itinerary) a passenger may follow, given their tag-on and tag-off locations and times. This problem arises in many AFC systems because the smart card check-points are typically located at the stations or on the platforms instead of inside the vehicles, and thus all journeys registered to the smart card database do not include information on the service/trip identifier (such as trip number or train line) the passenger used. A few assignment models have been developed in the literature (Hong et al., 2016; Hörcher et al., 2017; Kusakabe et al., 2010; Sun et al., 2015; Zhu, 2014; Zhu et al., 2017) to assign passengers to trains (see also Chapter 34). Once this issue has been solved, valuable insights into passenger movements can be obtained and important statistics can be derived for tactical planning purposes such as crowding levels for each train and station (from which network constraints can be identified), number of transfers, average waiting time at each station (based on which schedules can be adjusted) and reliability of travel time (train punctuality and service variability). Typical tactical applications of such studies include monitoring the crowding level over time; assessing the impact of major timetable change, either ex ante or ex post, on customer experience in terms of journey time, transfers and waiting time; planning of service variations (mix of local and express services on the same route to cater for variation in travel demand along the route; optimising passenger experience given budget constraints (e.g., rolling stock, network capacity or simply subsidy level) or upgrading stations to cater to increasing demand.

Studying how sensitive public transport users are to travel time and cost (i.e., fares) using smart card data is another important topic. Examples include work by de Grange et al. (2013) and that of the Independent Pricing and Regulatory Tribunal New South Wales (Hensher and Ho, 2020). The latter work used smart card data at the aggregate level to estimate both direct and cross elasticities of demand with respect to time and fare for various transport modes. The elasticities obtained are used to made adjustment to the existing pricing of public transport service, such as setting up the maximum fares that deliver fairer and more equitable outcomes.

There are many more research topics that add value to tactical planning (Kieu et al., 2015; Morency et al., 2006, 2007; Seaborn et al., 2009). However, these studies are, in essence, a variation or a combination of the topics covered previously. While techniques and methods used may differ, these studies share the common focus of classifying passengers and/or passenger journeys to better understand travel patterns instead of offering direct implications for planning purposes.

Research with operational contributions

At the operational level, smart card data have been used to calculate detailed performance indicators on a public transport network, such as weekly or daily ridership, service punctuality (e.g. percent services running on time), actual station-to-station journey times, average travel speed by corridor and average boarding/dwelling time at each stop or each corridor/route (Morency et al., 2007; Reddy et al., 2009). Obtaining operational statistics is considered the standard of the practice by many operators and transport authorities, of which Transport for NSW and Transport for London are examples (Transport for London, 2020; Transport for NSW, 2020). These statistics are useful for operational planning under both scheduled (i.e., normal) and disrupted networks. Under the normal undisrupted services, the number of passengers observed at each stop by time of day is useful for determining appropriate infrastructure such as shelter, platform size and lighting system at every stop (Gschwender et al., 2016). Similarly, the boarding and required dwelling time at each stop is critical for the design of express bus and train services which aim to serve key stations/stops by skipping less important stops along the routes (i.e., stop and skip pattern design) (see also Chapter 38).

Under disrupted network scenarios, smart card data could be and have been used to automatically identify and compensate passengers. For example, Transport for London uses smart card data to compensate impacted users if their journeys are delayed for more than a certain threshold, with the amount of compensation depending on the extent to which the service is delayed, starting from 15 minutes. Passenger flows established from smart card data can also be used to devise operational plans should the transport network experience major disruptions. An example application for this operational planning using Singapore as a case study can be found in Jin et al. (2013).

Recognising the values and limitations of smart card data

Before turning to the series of case studies, it is worth discussing the potential of smart card data for transport research by recognising their values and limitations through the types of information the data conveys, either by themselves or in conjunction with other datasets. Note the distinction between data and information; the latter is insight gained from analysing the former. Smart card data may contain many fields, including sociodemographics of the card holders (see also Chapter 37) and the type of card they use (e.g. senior, student, adult and junior). Amongst these data fields, departure station, entry time at gate (or departure time), arrival station and exit time at gate (or arrival time) – referred to as the *quadruple* of the smart card data – are available in a majority of smart card ticketing systems. There is a consensus amongst researchers that the quadruple is a minimal set of data required for a precise estimation of passenger flows (Hong et al., 2016). However, some smart card systems (e.g., Transantiago) do not have the exit station and exit time, and thus various methods have been developed and implemented in the literature to source the missing data fields of the quadruple (see the "Research with Tactical Contributions" section previously).

There is a subtle but important difference between entry time and departure time and between exit time and arrival time. The latter of each pair are included in the smart card systems if card readers are deployed inside the vehicles (e.g., buses, trains), while the former would be part of the smart card data if card readers are installed on the platforms or the stations. This is effectively driven by the practical considerations of where the tap on and tap off take place for a smooth operation (i.e., inside the vehicles or outside the vehicles), with very little regard for potential use of the smart card data in analysis. Indeed, smart card data are a by-product of an AFC system, which is designed with the main purpose of collecting fares automatically.

Nevertheless, the value of smart card data is quite limited when considered on its own, since the information derived directly from such data would be limited to the number of taps on and taps off (i.e., demand) at a particular station/stop, possibly segmented by time of day. To make it more useful, analysts have to bring in more data from other sources, such as service timetable data, automatic vehicle location data, census data, household travel survey data and meteorology data, to name just a few. Along the process of marrying the smart card data with secondary data lie various assumptions. For example, to estimate the waiting time by combing the smart card data with the timetable data, it is necessary to assume that services are punctual to the timetable if automatic vehicle location data (or GPS data of the vehicles) are not available. Similarly, to predict where transfers would occur in a closed-system (i.e., a system where passengers only leave trace at the entry and exit points) using some kind of passengers-to-train assignment models, some rules/assumptions must be used. These could be simple, with equal probabilities for each possible transfer locations, or more sophisticated, with some behavioural theory such as utility maximisation and Bayesian theory. The point is that there is no smart card system that provides all the data required that one might need, and thus all sorts of assumptions are needed.

In addition, smart card data provide information on the behaviour of public transport users while offering no insights for non-users. Thus, the use of data sources to enrich smart card data is inevitable, especially for obtaining a better understanding of travel demand of public transport users but also that of non-users; both are necessary for effective planning and modelling work. As far as smart card data are concerned, non-users include public transport non-users and those who opt to use alternative methods of payment, such as paper-tickets, credit/debit cards and recently mobile phones. As more and more cities progressively allow contactless payment to operate alongside the smart card, the smart card is becoming part of the phone, and this raises a lot of interesting possibilities for seamless mobility such as Mobility as a Service (MaaS) (see also Chapter 3). This does not mean that smart card data will become useless or biased, because regardless of the contactless payment method used, transactions are registered to the smart card systems and equivalent data are captured and used for planning purposes. Even when such data were not included in the smart card system, methodological developments would be available to make smart card data useful and representative of the public transport users. What is more challenging is the identification of public transport non-users' behaviour. To this end, the use of secondary, preferably large-scale, data is necessary. Another challenge that limits the value of the smart card data is the willingness and capability of transport authorities to use the data instead of being swamped by it. It is encouraging to see more and more practical uses of smart card data by transport authorities – an example of which is found in Diab and El-Geneidy (2012).

An overview of the case study chapters

This section briefly introduces the series of case studies in this Handbook that analyse smart card data to obtain insights for planning purposes from different parts of the world. All case study chapters aim to make contributions at the tactical and operational levels. These two levels of analysis seem to dominate the recent trend in mining smart card data for planning purposes as shown by, for example, Alsger et al. (2018) inferring the purposes of trips recorded to the Brisbane's Go-Cards; Briand et al. (2017) analysing year-to-year variation in public transport users travel behaviour in Gatineau, Canada; Gordon et al. (2018) estimating passenger flows on the London network Kim et al. (2017) examining habitual behaviour of bus users in Brisbane Australia; Li et al. (2017) forecasting passenger flow under special events using Beijing smart card data and Tao et al. (2018) and Zhou et al. (2017) studying the impact of weather on bus ridership using smart card data in Brisbane Australia, and Shenzhen, China respectively.

Ho and Ho (Chapter 34) start the series of case study chapters with an investigation on crowding on platforms using smart card data from Sydney, Australia. The study combines smart card data with real-time generic transit feed specification (GTFS) data giving the status of each service (being on time, late, early or being cancelled) and journey planner data to trace individual passenger movements, up to the platform level, within the closed network of Sydney Trains, where train users only leave traces when they tap on or off at the entry and exit stations, respectively. Aggregating the passenger movements and dividing the results by platform areas, they obtain the variation of platform density to give information about crowding across time of day and across stations within the Sydney Trains networks. These results could be useful in designing station layout but also for train timetable planning and economic appraisal of initiatives aiming to improve the customer experience by increasing station and/or platform capacity.

Clifton (Chapter 35) and Yen (Chapter 36) use Sydney Opal and South-East Queensland Go-card data, respectively, to provide insights for the tactical planning of the frequency of services, particularly around the service headway of core corridors that trigger a turn-up-and-go behaviour from users. The research assesses the correlation between service headway and actual waiting time at the station, with the latter being extracted from smart-card data as the elapsed time from actual arrival time to planned departure time, assuming that all trains are punctual to the timetable.

Yamamoto and Nakamura in Chapter 37 use data from the city of Shizuoka smart-card data to study public transport passengers' travel patterns in terms of trip timing and locations, segmented by the sociodemographics of the passengers. Leveraging the spatio-temporal entropy index, this study provides valuable insights for policy formulation, particularly around public transport pricing, to encourage social inclusion of the elderly. Lee and Bencekri (Chapter 38) use Seoul smart-card data to provide insights for train operations, particularly around the design of stop and skip services (i.e., express vs. all-stops services) to maximise passenger welfare such as aggregate travel time, waiting time and transfers. Trips observed in the smart card data were first classified into different purposes, which were then treated as separate segments in the simulation framework of MATSim software to decide the optimal ratio of express to all-stop services.

Utsumi, Schmoecker and Nakamura (Chapter 39), also using Shizuoka smart-card data, provide insights for bus operation by assessing the impact of reducing bus frequency on bus users' behaviour, including the user segments that are likely to change their travel behaviour in terms of departure time choice (use earlier or later bus or stop using bus) and the time they took to settle down their choices of departure time if they continue using bus after the timetable change.

While these case studies show how smart card data are used in public transport planning and network design, they by no means present a full collection of how smart card data could be used in transport research and practice, particularly in the strategic context. It is also worth noting that while smart card data have a wide range of application in research that is not necessarily limited to the transport domain (such as using smart card data for profiling customers for marketing purposes), the way in which it is used is more reasonable and defensible than other use cases. An example of the latter would be the (mis)use of smart card data that do not include personal information to infer how passengers access and egress public transport services, assuming the true origin and destination of each trip. The point here is that both the strengths and weaknesses of smart card data need to be recognised and acknowledged in order to make a convincing case for its usefulness.

References

Agard, B., Morency, C., & Trépanier, M. (2006). Mining public transport user behaviour from smart card data. *IFAC Proceedings Volumes*, *39*(3), 399–404.

Alsger, A., Assemi, B., Mesbah, M., & Ferreira, L. (2016). Validating and improving public transport origin–destination estimation algorithm using smart card fare data. *Transportation Research Part C: Emerging Technologies*, *68*, 490–506. doi:10.1016/j.trc.2016.05.004

Alsger, A. A., Mesbah, M., Ferreira, L., & Safi, H. (2015). Use of smart card fare data to estimate public transport origin–destination matrix. *Transportation Research Record*, *2535*(1), 88–96. doi:10.3141/2535-10

Alsger, A., Tavassoli, A., Mesbah, M., Ferreira, L., & Hickman, M. (2018). Public transport trip purpose inference using smart card fare data. *Transportation Research Part C: Emerging Technologies*, *87*, 123–137. doi:10.1016/j.trc.2017.12.016

Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data. *Transport Policy*, *12*(5), 464–474.

Briand, A.-S., Côme, E., Trépanier, M., & Oukhellou, L. (2017). Analyzing year-to-year changes in public transport passenger behaviour using smart card data. *Transportation Research Part C: Emerging Technologies*, *79*, 274–289. doi:10.1016/j.trc.2017.03.021

Chu, K. K. A., & Chapleau, R. (2008). Enriching archived smart card transaction data for transit demand modeling. *Transportation Research Record*, *2063*(1), 63–72. doi:10.3141/2063-08

de Grange, L., González, F., Muñoz, J. C., & Troncoso, R. (2013). Aggregate estimation of the price elasticity of demand for public transport in integrated fare systems: The case of Transantiago. *Transport Policy*, *29*, 178–185. doi:10.1016/j.tranpol.2013.06.002

Diab, E. I., & El-Geneidy, A. M. (2012). Understanding the impacts of a combination of service improvement strategies on bus running time and passenger's perception. *Transportation Research Part A: Policy and Practice*, *46*(3), 614–625. doi:10.1016/j.tra.2011.11.013

Ellison, R. B., Ellison, A. B., Greaves, S. P., & Sampaio, B. (2017). Electronic ticketing systems as a mechanism for travel behaviour change? Evidence from Sydney's Opal card. *Transportation Research Part A: Policy and Practice*, *99*, 80–93.

Gordon, J. B., Koutsopoulos, H. N., & Wilson, N. H. M. (2018). Estimation of population origin–interchange–destination flows on multimodal transit networks. *Transportation Research Part C: Emerging Technologies*, *90*, 350–365. doi:10.1016/j.trc.2018.03.007

Gschwender, A., Munizaga, M., & Simonetti, C. (2016). Using smart card and GPS data for policy and planning: The case of Transantiago. *Research in Transportation Economics*, *59*, 242–249. doi:10.1016/j.retrec.2016.05.004

He, L., Agard, B., & Trépanier, M. (2020). A classification of public transit users with smart card data based on time series distance metrics and a hierarchical clustering method. *Transportmetrica A: Transport Science*, *16*(1), 56–75. doi:10.1080/23249935.2018.1479722

Hensher, D., & Ho. C. Q. (2020). Obtaining direct and cross-fare elasticities from opal e-data in Sydney, Australia. *Journal of Transport Economics and Policy*, *54*(4), 289–315.

Hong, S.-P., Min, Y.-H., Park, M.-J., Kim, K. M., & Oh, S. M. (2016). Precise estimation of connections of metro passengers from smart card data. *Transportation*, *43*(5), 749–769. doi:10.1007/s11116-015-9617-y

Hörcher, D., Graham, D. J., & Anderson, R. J. (2017). Crowding cost estimation with large scale smart card and vehicle location data. *Transportation Research Part B: Methodological*, *95*(Suppl. C), 105–125. doi:10.1016/j.trb.2016.10.015

Jin, J. G., Teo, K. M., & Sun, L. (2013). *Disruption response planning for an urban mass rapid transit network*. Transportation Research Board 92nd Annual Meeting.

Kieu, L. M., Bhaskar, A., & Chung, E. (2015). Passenger segmentation using smart card data. *IEEE Transactions on Intelligent Transportation Systems*, *16*(3), 1537–1548. doi:10.1109/TITS.2014.2368998

Kim, J., Corcoran, J., & Papamanolis, M. (2017). Route choice stickiness of public transport passengers: Measuring habitual bus ridership behaviour using smart card data. *Transportation Research Part C: Emerging Technologies*, *83*, 146–164. doi:10.1016/j.trc.2017.08.005

Kusakabe, T., Iryo, T., & Asakura, Y. (2010). Estimation method for railway passengers' train choice behavior with smart card transaction data. *Transportation*, *37*(5), 731–749. doi:10.1007/s11116-010-9290-0

Li, Y., Wang, X., Sun, S., Ma, X., & Lu, G. (2017). Forecasting short-term subway passenger flow under special events scenarios using multiscale radial basis function networks. *Transportation Research Part C: Emerging Technologies*, 77, 306–328. doi:10.1016/j.trc.2017.02.005

Luo, D., Cats, O., & van Lint, H. (2017). Constructing transit origin–destination matrices with spatial clustering. *Transportation Research Record*, *2652*(1), 39–49. doi:10.3141/2652-05

Morency, C., Trepanier, M., & Agard, B. (2006, September 17–20). *Analysing the variability of transit users behaviour with smart card data*. IEEE Intelligent Transportation Systems Conference.

Morency, C., Trépanier, M., & Agard, B. (2007). Measuring transit use variability with smart-card data. *Transport Policy*, *14*(3), 193–203. doi:10.1016/j.tranpol.2007.01.001

Munizaga, M., Palma, C., & Mora, P. (2010). *Public transport OD matrix estimation from smart card payment system data*. Proceedings of 12th World Conference on Transport Research, Lisbon, USB.

Ortega-Tong, M. A. (2013). *Classification of London's public transport users using smart card data*. Massachusetts Institute of Technology.

Park, J. Y., Kim, D. J., & Lim, Y. (2008). The potential of using the smart card data to define the use of public transit in Seoul. *Transportation Research Record: Journal of the Transportation Research Board*, *2063*, 3–9.

Pelletier, M.-P., Trépanier, M., & Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, *19*(4), 557–568.

Reddy, A., Lu, A., Kumar, S., Bashmakov, V., & Rudenko, S. (2009). Entry-only automated fare-collection system data used to infer ridership, rider destinations, unlinked trips, and passenger miles. *Transportation Research Record*, *2110*(1), 128–136. doi:10.3141/2110-16

Seaborn, C., Attanucci, J., & Wilson, N. H. M. (2009). Analyzing multimodal public transport journeys in London with smart card fare payment data. *Transportation Research Record*, *2121*(1), 55–62. doi:10.3141/2121-06

Shelfer, K. M., & Procaccino, J. D. (2002). Smart card evolution. *Association for Computing Machinery. Communications of the ACM*, *45*(7), 83–83.

Singh, P., Oh, K., & Jae-Yoon, J. (2017). Flow orientation analysis for major activity regions based on smart card transit data. *ISPRS International Journal of Geo-Information*, *6*(10), 318. doi:10.3390/ijgi6100318

Sun, L., Lu, Y., Jin, J. G., Lee, D.-H., & Axhausen, K. W. (2015). An integrated Bayesian approach for passenger flow assignment in metro networks. *Transportation Research Part C: Emerging Technologies*, *52*(Suppl. C), 116–131. doi:10.1016/j.trc.2015.01.001

Tao, S., Corcoran, J., Rowe, F., & Hickman, M. (2018). To travel or not to travel: 'Weather' is the question. Modelling the effect of local weather conditions on bus ridership. *Transportation Research Part C: Emerging Technologies*, *86*, 147–167. doi:10.1016/j.trc.2017.11.005

Transport for London. (2020). *Public transport journeys by type of transport*. Retrieved March 20, 2020, from https://data.london.gov.uk/dataset/public-transport-journeys-type-transport

Transport for NSW. (2020). *Train patronage*. Retrieved March 30, 2020 from www.transport.nsw.gov.au/data-and-research/passenger-travel/train-patronage

Transport Systems Catapult. (2017, February). *Modernising the national travel survey. WP1 – Review of current and future technologies and data sources report*. Retrieved April 30, 2020, from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/641452/modernising-the-national-travel-survey.pdf

Trépanier, M., Morency, C., & Agard, B. (2009). Calculation of transit performance measures using smart-card data. *Journal of Public Transportation*, *12*(1), 79–86.

Utsunomiya, M., Attanucci, J., & Wilson, N. (2006). Potential uses of transit smart card registration and transaction data to improve transit planning. *Transportation Research Record*, *1971*(1), 118–126. doi:10.1177/0361198106197100114

van de Velde, D. M. (1999). Organisational forms and entrepreneurship in public transport: Classifying organisational forms. *Transport Policy*, *6*(3), 147–157. doi:10.1016/S0967-070X(99)00016-5

Welch, T. F., & Widita, A. (2019). Big data in public transportation: A review of sources and methods. *Transport Reviews*, *39*(6), 795–818. doi:10.1080/01441647.2019.1616849

Zhou, M., Wang, D., Li, Q., Yue, Y., Tu, W., & Cao, R. (2017). Impacts of weather on public transport ridership: Results from mining data from different sources. *Transportation Research Part C: Emerging Technologies*, *75*, 17–29. doi:10.1016/j.trc.2016.12.001

Zhu, Y. (2014). *Passenger-to-train assignment model based on automated data*. Massachusetts Institute of Technology.

Zhu, Y., Koutsopoulos, H. N., & Wilson, N. H. M. (2017). A probabilistic passenger-to-train assignment model based on automated data. *Transportation Research Part B: Methodological*, *104*, 522–542. doi:10.1016/j.trb.2017.04.012