Mousumi Bagchi
Senior Consultant,
Steer Davies Gleave,
London, UK

Peter White
Professor of Public Transport Systems,
Transport Studies Group,
University of Westminster, London, UK

ice

What role for smart-card data from bus systems?

M. Bagchi and P. R. White

This paper examines whether data, generated from smart cards used for bus travel, can be put forward as a replacement for, or a complement to, existing transport data sources. Smart-card data possess certain advantages over existing bus ticket machine data and some sample data sources, allowing them to be used for a range of analysis applications that transport service providers may previously have been unable to or found difficult to undertake. To this end, as a new transport data source, the paper firstly reviews the nature of smart-card data. The paper then goes on to examine the impact of smart-card data in relation to two case studies—one concerning its impact on the data collection process and one looking at the impact on travel behaviour analysis.

1. INTRODUCTION

Smart-card data, generated from the use of smart cards for bus travel, constitute a new transport data source that can be used by transport service providers for a range of applications.

Smart cards are similar in look and size to credit cards and can be used in place of traditional fare media such as paper tickets and magnetic stripe cards. Each smart card can be identified by a unique serial number. The cards can be registered to a given individual, or they can be anonymous. On the cards can be placed electronically a range of fare payment products, such as travel passes and other tickets. Monetary value can also be stored on the card ('stored value') and debited as and when trips are made.

Every time a smart card is used, details of the trips being made with those cards are recorded. Overall, when considering these features of smart cards in comparison to existing bus electronic ticket machine data and certain other existing sample data sources, it means that transport service providers will

(a) have access to much *larger volumes* of personal travel data than it is possible to obtain through existing data sources
(b) be able to *link* those data to the individual card and/or to the individual traveller
(c) have access to *continuous* trip data covering longer periods of time than it is possible to obtain using existing transport data sources
(d) know who a larger proportion of their customers are.

Using smart-card technology, it is also possible to record the stage-based destination of a bus trip as a passenger disembarks from a bus, and therefore obtain automatically both origin and destination information on trips made. With smart-card data, transport service providers can 'construct' the trips that people make over the course of the day or longer, and examine travel behaviours that have been difficult to examine because of the deficiencies of existing data sources.

The perceived potential of smart-card data in these respects was one of the reasons behind the drive by local transport authorities in the mid-1990s to implement some bus-based smart-card systems in the UK. Local transport authorities require passenger trip data for a range of applications, including calculating reimbursements to transport operators for carrying concessionary travellers (typically the elderly, disabled and scholars), operator performance monitoring, demand modelling, general travel behaviour analysis and so forth. Transport operators too need better information with which to plan the running of their services.

Today, in the UK, there are over 15 bus-based smart-card schemes operating. Nearly half a million smart cards are in circulation across the UK. Local transport authorities manage the majority of these schemes, but some of the largest are commercial bus-operator-led schemes such as those run by First in Bradford, Nottingham City Transport and Lothian Buses.

While there are clear advantages of smart-card data as described above, this paper examines whether smart-card data can actually be put forward as a replacement for, or a complement to, existing transport data sources.

There are two key points to note regarding definitions in this paper. The first is that *replacement* is relative to the analysis or the process being examined and the purpose for which the analysis or process is being undertaken. For example, we may find that smart-card data can replace an existing data source for a particular analysis but not another. Secondly, when *existing transport data* sources are referred to, this relates to the type of transport data source most commonly used for a particular process or analysis.

To examine these issues in more detail, this paper considers the following three key elements.

(a) A review of the *nature* of smart-card data and what this tells us about the *replacement* versus *complement* scenario.

(b) An examination of the impact of the data on the concessionary travel reimbursement data collection *process*.

(c) An *analysis* of sample smart-card data.

To this end, the remainder of this paper is structured as follows.

- Section 2 provides an overview of smart-card technology and the capabilities the technology possesses for data collection.
- Section 3 looks at how data are captured through smart-card use, and what specific items of information can and cannot be collected through smart-card use
- Sections 4 and 5 then present case studies of smart-card data collection and use—one theoretical, and one an application of smart-card data. Section 4 is a case study of the theoretical impact of smart-card data on the concessionary travel reimbursement data collection process. Section 5 is a case study of the application of smart-card data for travel behaviour analysis.
- Section 6 summarises the conclusions of the paper.

2. SMART-CARD TECHNOLOGY

The way smart cards are designed, the way they interface with the machines that read them, and the capabilities they possess to process information all have implications for how the cards can be used and the benefits they confer in terms of the quality of data generated from their use.

The term 'smart card' has been used to describe a range of card classifications and technologies. The microchips embedded in the smart cards can be computer chips, capable of both storing and processing information, or memory chips, which are capable only of accessing data already stored on the card.

The manner in which a smart card can be used depends on how the chip in the card interfaces with the card reader machines (e.g. a bus ticket machine). The chips on the card can have either a *contact* or a *contactless* interface. With contact smart cards, the chip is connected to the surface of the card. In order to be used, these cards have to come into contact with the device they are required to communicate with. These devices are known as 'read/ write' devices or 'readers' or 'terminals'.

Contactless cards, theoretically, do not have to come into physical contact with the device they are going to communicate with. Power is supplied to the card when it comes near the reader and within its magnetic field. Contactless cards that communicate within a range of 10 cm and conform to the international card standard ISO 14443[1] are known as 'proximity' cards. Those that can communicate at distances of up to 70 cm and conform to ISO 15693[2] are known as 'vicinity' cards.

Smart cards have to be programmed with different types and amounts of memory for storage of information. If a card has only a memory chip then it is known as a contactless memory card. If it has a memory and microchip or a microchip only then it is generically known as a microprocessor card. The amount of memory capacity will primarily dictate how many applications can be placed on the card.

Hybrid or dual-interface cards refer to smart cards that can support both the contact and contactless interface. Cards are described as hybrid when the independent contact and contactless technologies share a single card and do not communicate with one another.[3]

On the other hand, dual-interface cards (also referred to as combi cards) have a single chip that can communicate with the smart-card readers and other terminals using the contact or contactless interface. Dual-interface cards are cheaper in cost terms than hybrid cards. Also, because they have a single and integrated platform for contact and contactless applications, they may prove more popular than hybrid cards for multi-application schemes facilitating cooperation in multi-application projects across industrial sectors.[4]

3. USING SMART CARDS FOR BUS TRAVEL

3.1. The capture of data

When a person uses a smart card for travel, they place their card on or near a card-reading device. On a bus, this device can be part of the electronic ticket machine near the bus driver, although card-reading devices may be found at other places, for example at retail outlets where a person may 'top up' their card with stored value (these are collectively referred to as 'point-of-service terminals').

There are two types of transaction that can be made, and the information recorded from these transactions constitutes smart-card data. The first type is the journey transaction. For example, if a card has a travel pass on it, and a person boards a bus and places it on or near the reader, then certain details of the trip are recorded, including the time and date of boarding, the bus stage boarded, the bus route boarded, along with other details (the types of variables and attributes of bus smart-card data are discussed in detail in section 3.2). These data comprise personal travel data—that is, information on when and where a person has travelled using their card.

The second type of transaction is the monetary transaction. This is when the card is used to top up stored value, or renew a travel pass or purchase another type of ticket for placement on the card. In such circumstances, a record is made of the details of the purchase, including price, place of purchase (in code form), type of product and so forth.

The information captured by the card-reading devices is usually downloaded on a daily basis to a central repository that can be interrogated for data analysis.

The capture of data through smart-card use is an automated process with a reduced role for the subject of data collection, namely the transport user. This is in contrast to the collection of personal travel data through, for example, travel diaries, where the transport user has to recall from memory the details of the trips they have made over a defined period of time (e.g. seven days).

With smart cards, the role of the human is only in initiating the recording of the data, for example by a card user presenting their card to the reading device. The variables of information collected are pre-defined and the variables are populated depending on the nature of the transaction. This redefined role for the individual about whom data are being captured in the collection process

suggests that the potential for error in the recording of smart-card data is less in comparison to trip data collected through traditional sample surveys such as travel diary surveys.

Quality and utility of smart-card data, however, are relative to a range of factors that offset some of the benefits of the reduced role of the transport user in the collection and recording of the data. This is a theme revisited at many times in this paper. Quality factors are also considered in detail in Bagchi.[5]

3.2. What information can be collected from smart cards?

3.2.1. Overall collection of data.

(*a*) *Linking of data.* As previously mentioned, each time a smart card is used, details of the trip that is being made with that card will be recorded. Of central significance is that these trip data can be attributed to a given card, as each card possesses a unique serial number, and additionally, to the individual, if the cardholder's name and address details are known or the card is personalised with their photo. The transport service provider will be able to establish which cards have been used, a facility that is not possible with existing non-electronic travel tickets.

(*b*) *Volume and scope of the data.* Transport service providers will have access to larger volumes of passenger data than it is possible to obtain through existing data sources. The extent to which data are collected on the full range of a service provider's customers will however depend on the take-up of smart cards. There will be a proportion of passengers who will not take up the smart cards and will continue to pay cash fares only. As such, their trips will not be recorded and the service providers may not necessarily know who these customers are in the same way that they will be able to know who their smart-card customers are if they have their name and address details.

(*c*) *Continuous information.* Theoretically, all trips made using a particular card within the allowed area of use and on participating transport services, will normally be recorded. However, continuity will be broken if an individual does not use the card to make a bus trip (e.g. forgets) or if they consciously alternate between paying cash fares and using the smart card. Consequently, these situations, together with others that may arise (e.g. card and/or system failure), will mean that some trips made by an individual will not be recorded. However, because the capture of data will be continuous, it will cover longer periods of time than data collected from traditional sample surveys, which are often undertaken to capture trip data by given individuals for a period of a day or a week at the most. This opens up the possibility for more accurate long-term analysis and forecasting.

3.2.2. Spatial information and issues with 'exit' reading. Ideally, each smart-card-based trip should record boarding and alighting information. However, in the UK, it will be difficult to collect alighting information for smart-card-based bus trips until exit reading of smart cards is in place. The alternative is for the driver to record (physically), at boarding, the alighting stage for a person's smart-card-based trip.

The technological know-how for exit reading currently exists, but its reliability in providing information to the required level still has to be proven in field tests. Exit reading requires card readers to be placed at both entry and exit points in the bus (but as is explained below, there may only be one point of entry/exit).

If proximity smart cards are used to make a trip, then the card user needs to be very close to the card reader for the card to be successfully read. If a vicinity card is used, then because of its increased reading range, the probability of the transaction being recorded is higher. At present, there are no bus-based smart-card schemes incorporating vicinity-card technology in the UK. The issues with the technology are compounded by the fact that in the UK many buses have only one door, mixing entry and exit flows. This may make it difficult to obtain good-quality exit reading as well as entry reading, as there is a larger volume of people going through a single door. Conversely, elsewhere in Europe, most buses have multiple doors, making it easier to position exit readers.

A qualitative study by Transport and Travel Research[6] of a trial of exit reading undertaken by Merseytravel (the Passenger Transport Executive for the Merseyside area) on circular bus routes in the Formby area indicated a number of factors affecting the ability to have successful exit reading. In the study, many elderly concessionary travellers were reported as saying that they thought the location of the exit reader was a hindrance to using it (it was too high), indicating that placement of the reader is important for those of less than average height or those who suffer from mobility problems. In addition, some participants stated that they simply forgot to use the exit readers. Partly this was because they were conditioned to getting on the bus and showing their pass, but not so for getting off the bus. Approximately one third of those who took part in a supplementary questionnaire-based survey about the trial said they hardly or never used the exit reader during the trial period.

In some cases it would be difficult to enforce exit reading. For example, where pensioners are entitled to free travel, they may be obliged to place their pass next to the smart-card reader on entering the bus to confirm eligibility, but it would be difficult to enforce any penalty for not doing so on exiting, as no failure to pay the appropriate 'fare' would be evident to the user.

Consequently, sample surveys to establish stage-based origin and destination for trips will still need to be undertaken to establish journey length, where local transport authorities need this information. However, for exact origins and destinations to be recorded (e.g. by postcode), direct passenger surveys will continue to be needed.

3.2.3. Temporal and structural information. The date and time of each trip will be recorded. Details of the mode and service (e.g. bus route) will also be recorded. Information will be available on all participating services in the scheme.

3.2.4. Cardholder information. Information can be captured at the card issue stage on name, address, age and gender of the card applicant, where they are willing to provide this information. This means that service providers will know who their smart-card customers are. Presently, service providers know who their customers are to a certain extent—for example, they may have a list of all those people who have been issued with an elderly travel pass—but this information is often out of date and kept in a

manner that does not facilitate its use for data analysis. Also, while they may have this information on one group of travellers, they may not have the same level of detailed information for other customers such as period pre-paid ticket holders.

At the card registration stage, the service provider is also afforded the opportunity to capture information on variables such as household car ownership. Difficulties will exist in obtaining updates of dynamic information such as cardholder address unless there is an incentive for the cardholder to inform the service provider of any such changes.

An important consideration with regard to the capture of cardholder information and any data linked to the cardholder is privacy of information. Organisations wishing to collect and process smart-card data in the UK will be subject to rules and regulations set down by the Information Commissioner. Some of the implications of the legislative requirements are reviewed in Bagchi.[5]

3.2.5. Purchase information. Data will be collected on the type of ticket purchased with the smart card, the price of the ticket, and the place of purchase.

3.3. What information cannot be collected from smart cards?

The smart-card system will not be able to capture information regarding the purpose for which a given trip is being made, as this is information that cannot be automatically detected, although time of day might be used as a proxy.

Data cannot be obtained on the perception or satisfaction of customers with the transport service; neither can information be obtained on non-users of particular transport services and non-smart-card users. Sample surveys are traditionally undertaken to obtain such information, and will still be needed for this purpose.

4. CASE STUDY 1. PROCESS: SMART-CARD DATA AND THE CONCESSIONARY TRAVEL REIMBURSEMENT PROCESS

4.1. Background to concessionary travel reimbursement

All local transport authorities have administrative responsibility for reimbursing operators for carrying concessionary passengers where such travel arrangements exist. The Transport Act 1985 put forward the concept of revenue forgone; that is, operators should be compensated for trips generated by lower fares in estimating the net compensation required. In the Transport Act 2000, a statutory minimum concessionary fare was set down, guaranteeing half-price bus fares (i.e. half the adult full fare), for the disabled and everyone of pensionable age. In some areas of the UK, concessionary travel was introduced for the first time for the elderly and disabled following the introduction of this Act.

To date, no local authority has published results of any evaluation of the use of smart-card data in the concessionary travel reimbursement process. Given this, it is pertinent to address in theoretical terms some of the practical issues involved in the use of smart-card data for such a purpose, focusing on the *collection* of information. This section focuses on surveys that

some of the larger local transport authorities outside of London undertake for reimbursement purposes.

4.2. Sample surveys and smart-card data in the reimbursement process

In order to estimate the amount of reimbursement that is due to operators, some larger local transport authorities undertake continuous passenger surveys (usually on a four-weekly cycle), where a sample of passengers are interviewed to find out

(*a*) if they are a concessionary passenger
(*b*) the type of ticket they are using for that journey
(*c*) the fare they have paid for their journey (if any)
(*d*) where they have travelled from and where they are travelling to (origin bus stage and destination bus stage).

Passengers are also asked about the purpose for which they are making their journey and a number of other journey-related questions.

The exact reimbursement methodology deployed will differ from authority to authority. In most areas, graduated (distance-based) fares are charged to adults, hence the revenue forgone (for example due to a half-fare scheme) requires estimation of trip length. Therefore, it is imperative to obtain accurate information on the stage-based origin and destination of the bus journey.

To examine the role of smart-card data on the data collection process for concessionary travel reimbursement, it is necessary to compare the two data sources in respect of three key components

(*a*) capture process
(*b*) volume of information
(*c*) origin and destination data.

Figure 1 presents a matrix summarising the role of smart-card data with respect to each of these components. The matrix illustrates the fact that the capture of information for reimbursement surveys can be either manual (i.e. the interviewer asks the respondent questions and notes down the answers on paper) or automated (i.e. collected using a handheld computer especially designed for this procedure).

While the actual capture of smart-card data is automated, as indicated in section 2, the bus user still has a role to play in the process because successful capture of trip data is dependent on the bus user making sure their card is successfully read on entry to and exit from the vehicle. This will also affect the potential to obtain stage-based origin and destination information on the trip being made by the smart-card user, and is particularly the case if proximity cards are being used. If there is no set-up for exit reading—that is, there is no reader at the exiting door or there is no requirement for the bus user to present their card on exit on a single-door bus—then stage-based information will not be obtained at all.

The burden of participation in the data collection process on the bus user can be reduced if vicinity cards are used. This is because, theoretically, the increased reading range increases the chance of the card being automatically read on entry and exit without the bus user having to be in close proximity to the card reader. However, even in this situation there are limitations. For

	Sample surveys	Smart-card data
Information capture	Manual and automated	Manual and automated
Trips	Sample	Potentially all
Origin and destination	Both	Varies

Fig. 1. Comparison of concessionary travel reimbursement sample surveys and smart-card data

example, if boarding and alighting is undertaken at different doors (rather than through a single door) then it is possible to obtain a number of unmatched trip records (i.e. entry record but no exit record) because exit from buses is typically less ordered than entry. There is also the problem of unintentional transactions being recorded by bus users standing/sitting near the card reader during their journey. This is just one example of where quality of data can be compromised by the type of smart-card technology being deployed and the ergonomics of bus use.

Through smart-card use, there is the potential to collect data on all bus trips made by concessionary travellers using smart-cards (census). This can only be achieved if all concessionary travellers use smart-cards and if the data capture processes are not compromised in ways described above or in other ways (e.g., card failure or reader failure), resulting in a loss of trip data or partial trip data being obtained.

4.3. Conclusions on smart-card data in reimbursement process

Given the review of the advantages and disadvantages of smart-card data in collecting information on concessionary travellers for the reimbursement process highlighted above, what does this suggest about the role of existing sample surveys for this purpose? The examination highlights that although there will be larger volumes of trip data collected, transport service providers will not be able to rely on this data source for the complete information they need. In this respect, local transport authorities may consider continuing existing sample surveys for reimbursement purposes but at a reduced level (i.e. fewer number of interviews being undertaken given greater volume of trip records already collected through smart-card use).

However, whether the sample survey continues to play the lead role or whether it plays a complementary role to smart-card data will depend on how the smart-card data are used. For example, local transport authorities may choose to use smart-card data as the basis for the reimbursement calculations, and use sample surveys to calibrate the smart-card data results. Through undertaking the sample surveys, it will be possible to

(a) check the results derived from smart-card data (e.g. from detecting trips not picked up through smart-card use)
(b) provide information on other variables that will help in the disaggregation of information for the reimbursement process (e.g. trip purpose).

Alternatively, local transport authorities may seek to continue using sample surveys as the main reimbursement tool, but use smart-card data as an occasional check on the sample survey results. This latter strategy will be more of an option during the initial period after a smart-card scheme has been introduced, where some smart-card-eligible concessionary travellers may be using their cards before others have been issued with their cards. It will also be an option if the local transport authority perceives there to be a continuing issue regarding smart-card data quality.

5. CASE STUDY 2. ANALYSING TRAVEL BEHAVIOUR—THE EXAMPLE OF BUS-TO-BUS INTERCHANGE

5.1. Introduction

This case study examines the use of smart-card data for travel behaviour analysis. Section 3 indicated how, through smart-card systems, transport service providers would

(a) have access to much larger volumes of personal travel data
(b) be able to link those data to the individual card and/or to the individual traveller
(c) have access to continuous trip data covering longer periods of time than it is possible to obtain using existing transport data sources
(d) know who a larger proportion of their customers are.

It is through the greater certainty about how, when and the proportion of trips made on their services that transport service providers will find smart-card data an attractive data source for a range of applications.

The previous sections have highlighted some of the limitations of smart-card data despite the above-stated advantages. However, to what extent can the above attributes of smart-card data help to overcome some of the deficiencies of existing transport data sources with regard to the analysis of certain travel behaviours?

One example of an analysis that has been difficult to undertake using existing data sources is the examination of bus-to-bus interchange. Currently, when a passenger boards a bus, the bus driver records their trip as a boarding on the electronic ticket machine. However, this boarding information cannot be linked to that specific passenger in any way. With smart-card data, because data can be linked to the individual or card, it is possible to 'construct' the trips an individual or groups of individuals make over the course of a day or longer to examine various travel patterns. In this way, theoretically, transport service providers can see when a card has been used to interchange between two or more buses within a defined time period, and identify the extent to which the interchange is taking place. To find this out at present, transport service providers typically undertake on-board surveys to ask passengers if they have interchanged with other services as part of their journey. The service provider can then, for example, make a service planning decision on whether a through bus service should be introduced to cater for the groups of users interchanging between certain bus services.

This and many other analyses are examples of where aggregate information is important for service planning, but in order to derive that aggregate information, it is necessary to look at individual behaviour. Smart-card data allow this benefit to be realised.

5.2. Methodology

5.2.1. Smart-card datasets. In order to examine bus-to-bus interchange, two sample smart-card datasets were analysed: one relating to a commercially operated scheme, and one relating to a trial of smart-cards for concessionary travellers.

The dataset from the commercially-run scheme was supplied by First in Bradford (FiB) who run a smart-card scheme in and around Bradford in West Yorkshire, UK. The scheme is open to all users of First in Bradford buses (they operate approximately 250 vehicles). On their smart-cards (branded 'FirstCard') can be placed FiB period bus passes and one-day travelcards. The bus passes are valid within an area known as the Rider boundary, surrounding Bradford. The card also has stored value, which can be used to purchase a range of bus tickets (including the passes mentioned). To date, approximately 40 000 cardholders have been registered. It is one of the longest-running commercial smart-card schemes in the UK.

The dataset from the trial scheme was supplied by the Passenger Transport Executive (PTE) for the Merseyside area, Merseytravel. The trial took place in the Southport area of Merseyside, where elderly concessionary travellers were allowed to use a smart-card loaded with a free bus pass to make trips on the then Southport and District Bus Company buses.

Contactless proximity cards are currently used by FiB. Merseytravel used this same card type in their smart-card trial. The FiB card is transferable—that is, it can be used by individuals other than the cardholder and does not have any photo identification or name of cardholder stamped on it. In contrast, the Southport trial smart-card was personalised with the cardholder's photo and name. Card transferability has implications for data processing because trips made by individuals other than the cardholder can be recorded through the use of a single card. As a result, it is not possible to attribute trips to a given individual, but rather to the card.

Merseytravel allows free travel on public transport for concessionary travellers across the PTE area (except for journeys between 0631 and 0929 h). In West Yorkshire, a 30 p off-peak flat fare operates for concessionary travellers (provided for by West Yorkshire PTE—this was 20 p at time of analysis).

It should be noted that the datasets might be representative of groups holding the smart-cards in the study areas, but not necessarily of the population. As a result, inferences are not drawn about the population from analysis of the samples. In the Bradford sample, it is probable that concessionary travellers especially are under-represented because the majority of the elderly opt to pay cash fares as and when they travel (20 p off-peak flat fare). As a result, the smart-card may appear attractive to high-frequency elderly concessionary travellers.

5.2.2. Data processing. Data were extracted from the offices of First in Bradford and Merseytravel (samples A). Names and address details of cardholders were *not* extracted. Due to the large quantities of data extracted, and in order to make the data analysis manageable, following validations, a random sample of approximately 10% of cards issued was drawn from both samples. All records corresponding to those cards were then extracted.

The resulting Bradford sample consisted of 3028 cards and 396 331 trip and other transaction records. The Southport sample consisted of 480 cards and 90 062 trip records (samples B). Checks were made to see if the samples were of sufficient size for the analyses being undertaken. All analyses were undertaken on the randomly drawn samples B.

5.2.3. Key issues in processing—defining the 'trip'. The UK National Travel Survey definition of a trip is a one-way journey from one activity to another, which may consist of journey stages defined as a change of mode during that one-way journey, or same-mode transfers. With smart-card data therefore, what we actually have are disaggregate boardings data—that is, each separate journey record is a 'boarding'.

In processing these data, a decision has to be made on whether to group these records or to analyse them ungrouped. Ungrouped boardings records can be referred to as 'unlinked' trips. When looking at bus-to-bus interchange, grouping is necessary, and therefore 'linked' trips are being examined. We can then build rules to process these defined linked trips. A linked trip, for the purpose of this paper, is defined as two or more boardings on different bus services by an individual on a given day within a certain time limit, as part of a one-way journey from one activity to another.

5.2.4. Key issues in processing—'rules-based' processing. Using the two samples, linked trips were inferred for two groups. These were period travelcards (Bradford only), and elderly concessionary travellers (note: for Bradford, looking at trips made using smart-card-based period cards and elderly concessionary tickets).

Within each group, linked trips were identified for each individual card for each day that card had been used. The time lapsed between successive boardings was examined in each case. This was applied over the entire set of trips undertaken by each group, rather than sampling specific time periods. It was thought that the fact that the two datasets of Southport and Bradford span different time periods (four years versus two years and two months) would not affect the comparability of the linked trip figures, as linked trips would be calculated as a proportion of all trips undertaken using the cards in each group, for each dataset. Four variables were used for linked trip identification. These were card number, date of travel, time of travel, and bus service number used.

The default transfer time limit was set at 30 min. The setting of this time limit can be seen as arbitrary. Elderly concessionary groups, for example, might undertake activities over shorter time periods than other user groups. This highlights some problems of logic in rules-based processing. For example, routine trips to the post office to collect pension payments might be the only reason a journey has been made and could typically be undertaken within a few minutes after disembarking from a bus. Consequently, longer transfer time limits for this group may pick up return trips as well as transfer trips, whereas a short transfer time limit would pick up transfer trips and a small number of genuine return trips that happened to satisfy the transfer trip criteria.

When identifying transfer trips, transfer time limits therefore need to be set with reference to the user group being examined.

Card No.	Date	Time	Boarding sequence no.	Service	Lapsed time: min	≤30 min	Different service?
xxxx	04/03/1997	13:52	1	A	0	—	—
xxxx	04/03/1997	15:08	2	B	76	No	—
xxxx	04/03/1997	15:22	3	B	14	Yes	No
xxxx	04/03/1997	15:38	4	A	16	Yes	Yes

Table 1. Processing logic exclusions

What would help to make the setting of the time limit less arbitrary is information on the purpose for which a journey was undertaken, and the length of time the journey took (if exit reading took place). We could then derive activity length, which would help to make the setting of the time limit more accurate. Unfortunately, with smart-card data we have neither piece of information.

The problems of rules-based processing can also lead to 'processing logic' exclusions. This is illustrated in Table 1. The table indicates a pattern whereby an individual has seemingly undertaken a return trip. For the outward trip, they have travelled on bus service A and then bus service B after a gap of 30 minutes. That does not register as a linked trip. They then take bus B and after a short while bus A. That registers as a linked trip.

One way of addressing these rules-based processing issues is to examine the direction of travel as indicated by the fare stage information crosschecked with operator fare stage tables. An attempt was made to look at this and incorporate it in analyses. However, this requires information on both the boarding and alighting stages of a trip, which is only recorded for single-ticket journeys. Additionally, fare stage tables and routes change frequently, which means it is difficult to obtain accurate information over the entire length of time covered by a smart-card dataset. Given these drawbacks, it was decided not to proceed with this aspect of the analysis.

5.3. Data analysis results and discussion

The results of the linked trip analysis are shown in Table 2. For the Bradford elderly concessionary group, the results show that linked trips constitute 7·6% of all trips undertaken using elderly concessionary smart-card-based tickets over the course of the Bradford scheme being in operation (to date of analysis). Of these linked trips, 98% were two boardings linked trips, with the remainder being three and four boardings linked trips.

The results suggest that elderly concessionary travellers in Southport make a smaller proportion of linked trips than elderly concessionary travellers in Bradford. One reason for this could be

because the Southport concessionary travellers were allowed to use their existing free travel pass that also allowed travel on buses and trains, at the same time as being allowed to use their trial smart-card. Consequently, some travellers may have opted to use their existing free pass, deflating the number of linked trips recorded by the smart-card system. Other possible reasons are discussed in detail in Bagchi and White.[7]

Looking at the proportion of linked trips undertaken using period travelcards, a much higher proportion is observed, at 24%, with a high boardings to actual trips ratio of 1·25. Of these, 95% were two boardings linked trips. We know from previous analysis that a small number of period travelcards are smart-card based. It was noted by White[8] that where a small number of period cards are used, a high degree of interchange may be observed among that group. A proportion of period travelcard users may use their cards for the routine journey to work, entailing a change of bus. In addition to this, the small area over which the period travelcards are valid means that average journey times may be short (compared to other areas), and that should a transfer from one bus to another be required as part of a person's journey, then it would be undertaken within a shorter time than may be the case in other areas. Hence, a transfer time limit of 30 min may be appropriate for detecting linked trips by First in Bradford period travelcard holders but not for other period travelcard holders in larger conurbations. What this additionally suggests is that in setting the transfer time limit, the opportunity to travel as illustrated by travel-to-work distance must also be considered.

6. CONCLUSIONS

This paper has revealed that the key question is what data source will play the lead role and what data sources will play the complementary role for any given process or analysis. The roles will differ according to the analysis being undertaken and the nature of the variables of information that need to be captured for that analysis.

This paper has also shown that the nature of smart-card data means that rules-based processing dominates the smart-card data analysis processes. However, in the absence of information on

Group	Dataset	No. of boardings	Inferred trips (number of linked or separate boardings)	Number of linked trips	Linked trips: %	Ratio of boardings to inferred trips
Period travelcards	Bradford	82 367	65 637	15 945	24	1·25
Elderly concessionary	Southport	90 062	86 866	3 032	3·5	1·04
Elderly concessionary	Bradford	8 086	7 504	568	7·6	1·08

Table 2. Linked trips—results

trip activity plus other quality constraints, the setting of rules can become arbitrary and can also lead to logic-based exclusions which may not have occurred using data sources where information is available on variables such as trip activity.

REFERENCES

1. INTERNATIONAL STANDARDS ORGANISATION. *Identification Cards: Contactless Integrated Circuit Cards—Proximity Cards.* ISO 14443-1–4, 2000/2001.

2. INTERNATIONAL STANDARDS ORGANISATION. *Identification Cards: Contactless Integrated Circuit Cards—Vicinity Cards.* ISO 15693-1–3, 2000/2001.

3. SMART CARD ALLIANCE. *Contactless Technology for Secure Physical Access: Technology and Standard Choices.* Smart Card Alliance, October 2002.

4. HOLM C. and BLYTHE P. Use of combi cards in automatic fare payment systems. *Traffic Engineering and Control*, 2002, **43**, No. 1, 16–20.

5. BAGCHI M. *Use of Smart Card Data from Bus Systems for Travel Behaviour Analysis, and Implications for Marketing.* Unpublished PhD thesis, University of Westminster, 2003.

6. TRANSPORT AND TRAVEL RESEARCH. *Swift Card Trial Formby—Report on Focus Group and Quantitative Research.* TTR, 2001, unpublished report for Merseytravel.

7. BAGCHI M. and WHITE P. R. Use of public transport smart card data for understanding travel behaviour. *Proceedings of the European Transport Conference*, Strasbourg, 8–10 October 2003 (CD-ROM).

8. WHITE P. R. Travelcard tickets in urban public transport. *Journal of Transport Economics and Policy*, 1981, **15**, No. 1, 17–34.

Please email, fax or post your discussion contributions to the secretary by 1 September 2004: email: kathleen.hollow@ice.org.uk; fax: +44 (0)20 7799 1325; or post to Kathleen Hollow, Journals Department, Institution of Civil Engineers, 1–7 Great George Street, London SW1P 3AA.