Analysis of data on public transport use from smartcards – update from December 2007 seminar

Peter White,
University of Westminster,
Helene Bataille and Richard Wells
Transport for London

Presentation to ITS(UK) Passenger Information Interest Group seminar, 26 November 2008
Data from smartcards

- Large-scale smartcard systems (such as Oyster in London) now cover a large proportion of all public transport trips
- As indicated at December 2007 seminar, potential for:
  - better aggregate demand measurement
  - determining trip rate per cardholder (e.g. Travelcard, Freedom Pass) and variation around this.
  - inferring linked trips via interchange
  - use of anonymised IDs to sample individual cards
Data from smartcards

- Temporal data quality may be better than spatial
- Some limitations, e.g. in London case, current absence of Oyster PAYG and validators from most TOC stations, and only boardings (not exits) recorded on buses
- Data issue may be separate from technology as such, e.g. magnetic card systems record similar data in some respects, and new technology may be used (e.g. mobile phones)
Main examples covered last year

- Pioneering bus-based systems in Southport and Bradford – derivation of trip rates and inference of linked trips
- Use of Oyster as Freedom Pass in London – variations in trip rates, inference of linked trips (bus and underground)
- Use of Oyster as Travelcard and PAYG on a TOC within London – substantial use of PAYG by commuters as well as occasional users
Further work in 2008

- Undertaken as dissertations on MSc Transport Planning and Management, University of Westminster in by students working for TfL
- Use of smartcard data as a ‘panel’ survey (Richard Wells)
- Patterns of bus use, and scope for promoting smarter travel choices (Helene Bataille)
Smartcard ‘panel’ data

- ‘Panel’ (or longitudinal) surveys cover same individual respondents over a specified period (e.g. annually), showing how individual (as well as aggregate) behaviour changes, e.g. before and after a particular system change.
- Problems of cost, feasible duration of period covered by a travel dairy, respondent ‘fatigue’ and attrition rates.
- Smartcards offer scope for automatically collecting such data, by identifying sample of cards used on part of a network prior to change, and trip patterns on same cards (using consistent anonymised IDs) after the change.
East London Line case study

• Traditional route (Whitechapel – Surrey Quays – New Cross/New Cross Gate) closed December 2007 for rebuilding as part of extended line

• Mix of ‘local’ traffic and trips interchanging to other routes, especially for central London (e.g. via Canada Water onto JLE)

• Replacement buses provided by TfL along line of route (but not cross-River section)
East London Line (cont)

• Sample identified of cards used in period before closure, for trips entering/exiting stations unique to the ELL at least three days in one week, for whole or part of journey

• 10,360 journeys identified from this sample of cards, broken down by mode(s) used where detected by Oyster readers, i.e. not only those journeys on ELL

• Classified by mode – LUL (underground); LTB (buses); DLR (Docklands) and NR (National Rail TOCs)
Pre-closure pattern

East London Line sample – Modal Breakdown Pre-Closure (based on 10,360 journeys)
Post-closure pattern

East London Line sample – Modal Breakdown Post-Closure (based on 10,962 journeys)

- DLR (1%)
- LTB (59%)
- LUL (33%)
- LUL/DLR (3%)
- LUL/NR (1%)
- LUL/NR/DLR (1%)
- NR (2%)
East London Line

• Rerouting of journeys can be inferred, e.g. frequent user travelling from an ELL station prior to closure to a central London station via underground interchange, then using a bus route nearby after closure

• ‘After’ comparison shows an increase in total trips by same sample to 10,988, but a fall in purely LUL journeys from 40% to 33% of trips, bus rising from 51% to 59% (NB bus boardings counted as separate journeys)

• Consistent with evidence of diversion on to existing radial bus routes (notably services 100 and 188) rather than replacement services directly parallel with former ELL route
Panel survey scope

• Advantages of smartcard ‘panel’ data include automatic data collection over a long period, and potentially large sample size.

• Disadvantages include need to make inferences about user behaviour, and understatement of trips via TOCs. It could be desirable to supplement with small direct user sample.
Use of Smart Card Data for Promoting Smarter Travel Choice and Optimising Bus Capacity
The research was divided in 2 main parts

Part 1

• Definition and characteristics of the bus users (time of travel / frequency)

Part 2

• Identification potential mode switchers based from bus trip patterns analysis.
Bus user sample

- Data for a two-week period in Spring 2008, avoiding Easter
- All bus boardings recorded by Oyster identified including Travelcard, [paid] Bus Pass, Pay as You Go (PAYG), Freedom Pass, Staff Pass and Child Bus Pass.
- These totalled 64.2 million, which corresponded to slightly over 80% of all bus boardings at that time
- Although child Oyster card use can be analysed separately, overall child use was under-represented, since it was not compulsory for children to have (or validate) an Oyster card for free travel at that time.
Bus Boarding on Weekdays

(Note that peak boardings may be understated due to some child users not having smartcards at time of data collection)
Bus Boarding on Saturdays
Bus Boarding on Sundays
Average Boarding per Card
(From sample of all Oyster bus boardings, two week period Spring 2008, avoiding Easter)

<table>
<thead>
<tr>
<th>Daily Average Boarding per Card</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunday w1</td>
<td>2.7</td>
</tr>
<tr>
<td>Monday w1</td>
<td>2.9</td>
</tr>
<tr>
<td>Tuesday w1</td>
<td>2.8</td>
</tr>
<tr>
<td>Wednesday w1</td>
<td>2.6</td>
</tr>
<tr>
<td>Thursday</td>
<td>2.8</td>
</tr>
<tr>
<td>Friday</td>
<td>2.9</td>
</tr>
<tr>
<td>Saturday</td>
<td>2.8</td>
</tr>
<tr>
<td>Sunday w2</td>
<td>2.6</td>
</tr>
<tr>
<td>Monday w2</td>
<td>2.8</td>
</tr>
<tr>
<td>Tuesday w2</td>
<td>2.8</td>
</tr>
<tr>
<td>Wednesday w2</td>
<td>2.8</td>
</tr>
<tr>
<td>Thursday w2</td>
<td>2.8</td>
</tr>
<tr>
<td>Friday w2</td>
<td>2.8</td>
</tr>
<tr>
<td>Saturday w2</td>
<td>2.8</td>
</tr>
</tbody>
</table>
Frequency of Card Use

Number of Days Cards were Validated During Week 1 on the Bus Network
(e.g. 27% of the cards validated during week 1, were validated only 1 day)
Bus Pass Users’ Profile
(Arrows indicate possible understatement of child peak trips)
Child Bus Pass User Profile

**Child Bus Pass Boardings on Weekdays**

**Number of Days Child Bus Passes were Validated During Week 1 on the Bus Network**

(E.g., 29% of the cards validated during week 1, were validated only 1 day)

- 1 day: 29%
- 2 day: 14%
- 3 day: 14%
- 4 day: 12%
- 5 day: 14%
- 6 day: 14%
- 7 day: 12%

**Boardings per Child Bus Pass on Weekdays**
PAYG Users Profile

PAYG Boardings on Weekdays

Number of Days PAYG cards were Validated During Week 1 on the Bus Network

[ E.g. 39% of the cards validated during week 1, were validated only 1 day]
Travel Card Users Profile
(on bus network)
Freedom Pass User Profile
(NB Pass for those aged 60 and over valid from 0900; for disabled users over the whole day)
Summary

Frequency of Cards Used on the Bus Network during week 1
(39% of all Child Bus passes validated during week 1 have been used only one day during that week)

Average daily boarding per ticket type

<table>
<thead>
<tr>
<th>Ticket Type</th>
<th>Weekday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus pass</td>
<td>4.5</td>
<td>4.1</td>
<td>3.7</td>
</tr>
<tr>
<td>Travelcards</td>
<td>2.8</td>
<td>2.9</td>
<td>2.7</td>
</tr>
<tr>
<td>Freedom pass</td>
<td>2.7</td>
<td>2.7</td>
<td>2.6</td>
</tr>
<tr>
<td>Staff pass</td>
<td>2.6</td>
<td>2.6</td>
<td>2.4</td>
</tr>
<tr>
<td>Child Bus pass</td>
<td>2.5</td>
<td>2.9</td>
<td>2.7</td>
</tr>
<tr>
<td>PAYG</td>
<td>2.2</td>
<td>2.2</td>
<td>2.1</td>
</tr>
</tbody>
</table>
Further research by Helene Bataille used information on interchange time coupled with bus user profiles to encourage mode switch at the busiest points of the bus network, using targeted marketing methods to promote use of soft modes (cycling / walking). Case study on short trips from route C10 to Bermondsey Station is available on request.
Acknowledgements

• Assistance of staff from Transys and TfL in provision and classification of data is gratefully acknowledged

• All inferences drawn from data are those of the authors
References


• M. Hofmann et al ‘Automated Identification of Linked Trips at Trip Level using Electronic Fare Collection Data’ TRB Annual Conference Washington DC January 2009 (forthcoming)

• Details of MSc dissertations at Westminster from P. R. White