

Empirical Investigation of Day-To-Day Variability In Driver Travel Behaviour

The “See you next Wednesday” Effect

Richard G. Clegg (richard@richardclegg.org)

Department of Mathematics,
University of York,
YO10 5DD

Universities Transport Studies Group, Bristol, January 2005

Outline

- 1 Background
 - Day-to-Day Driver Variability
 - Recurrence Rate Definition
 - Previous Research
- 2 York Studies
 - Data Collection Methodology
 - Lendal Bridge Study
 - Fishergate Study
- 3 Data Analysis
 - Raw Matching
 - Recurrence Rates
 - Model Fitting
- 4 Conclusions and Further Work

Day-to-Day Driver Variability

- The topic is that of ambient variability in road traffic — how does traffic change between days when the network remains constant?
- Is the traffic on one day the same as the traffic the next day?
- Is a rush-hour composed of the same vehicles travelling day after day?

Day-to-Day Driver Variability

- The topic is that of ambient variability in road traffic — how does traffic change between days when the network remains constant?
- Is the traffic on one day the same as the traffic the next day?
- Is a rush-hour composed of the same vehicles travelling day after day?
- Obviously not or I wouldn't be asking these questions.
- But how does the composition change as day follows day.
- What factors affect the variability of rush hour.

Recurrence Rate Definition

Given two time periods $T_1 = (S_1, E_1)$ and $T_2 = (S_2, E_2)$, the *recurrence rate* $R(T_1, T_2)$ is given by

Recurrence Rate

$$R(T_1, T_2) = \frac{\# \text{ vehicles seen in } T_1 \text{ and } T_2}{\# \text{ vehicles seen in } T_1}.$$

Note that usually T_1 and T_2 are disjoint (for example, the rush hour on different days). Note also that

$$R(T_1, T_2) \leq R(T_1, T'_2) \quad \text{if} \quad T_2 \subseteq T'_2.$$

That is, the recurrence rate will usually be increased if the second time period examined is made larger.

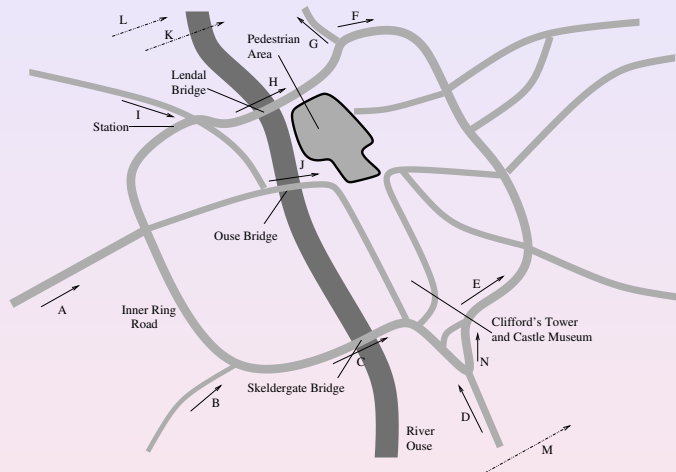
Previous Research

- “The percentage of individuals in each sample who exhibit the same characteristic across all days...is extremely small... [often] zero” (Pendyala 2003)
- “...we are led to reject the view that travel is highly routinized in the restricted sense that every weekday is assumed to look much like every other weekday.” (Huff and Hanson 1986).
- A study of recurrence rates between two days in Leeds is given by (Bonsall et al. 1984). The recurrence rate is less than 60% but depends strongly on the exact measure used.
- Reviews are given by (Batley and Clegg 2001) and (Clegg 2004).

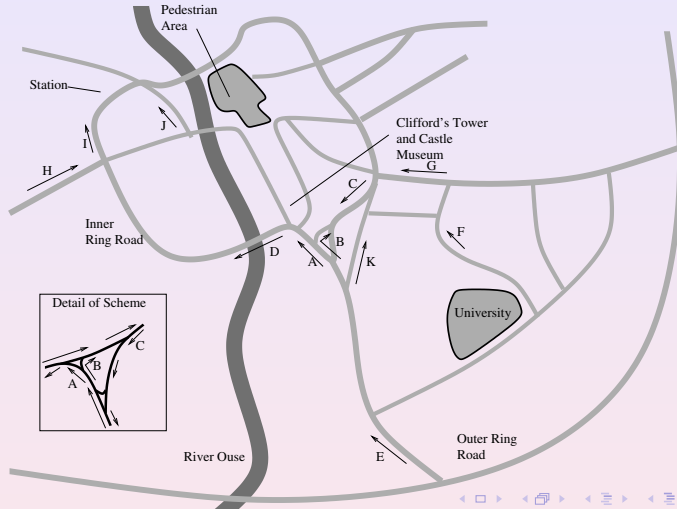
Data Collection Methodology

- Two studies took place in the city of York: Lendal Bridge and Fishergate.
- Partial licence plate data was collected at several sites for several weeks.
- The studies were centred around investigating interventions on the network.
- Licence plate data was collected manually (in almost all cases by audio tape).
- All the data here is publicly available to researchers.

Lendal Map



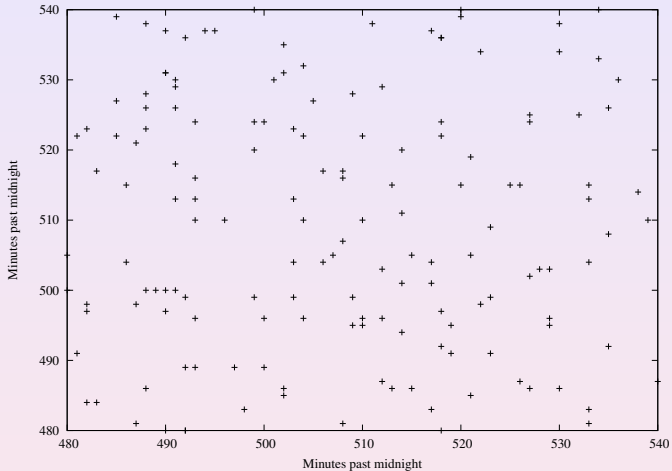
Fishergate Map



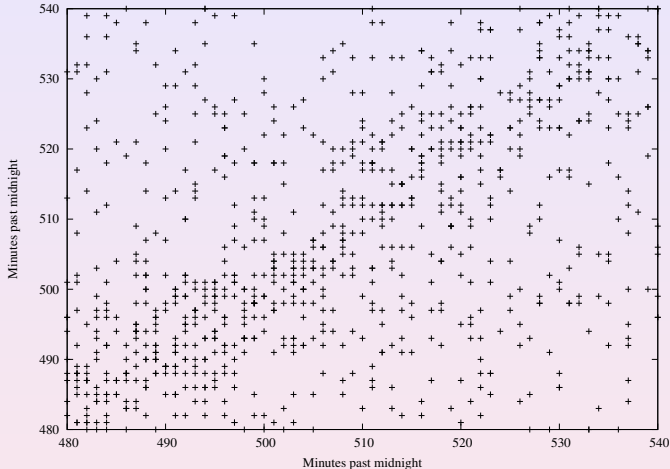
Fishergate Description

Day	Survey	Comment
25th June 2001	Before survey	Partial closure occurred at 9:15 and should not affect the data collected. First day of partial closure.
26th June 2001	Before survey	
27th June 2001	Before survey	
28th June 2001	Before survey	
29th June 2001	Before survey	
2nd July 2001	Before survey	
3rd July 2001	During survey	
4th July 2001	During survey	
5th July 2001	During survey	
6th July 2001	During survey	
11th July 2001	During survey	
12th July 2001	During survey	
13th July 2001	After survey	
16th July 2001	During Survey	Road works put back in place.

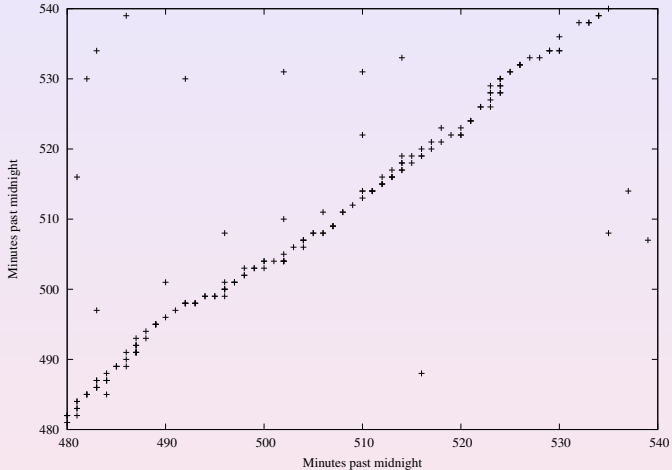
“Empty” match Lendal Sites L and M 28/6/00



Between days match Lendal Site M 6/9/00 7/9/00



Single day match Lendal Sites I and J 28/6/00



Lendal Site L 8:00 – 9:00 recurrence rate

	27/6/00	28/6/00	6/9/00	7/9/00	8/9/00
27/6/00	107.6 (99.7)	47.3 (39.0)	24.3 (17.8)	24.6 (18.1)	25.2 (18.8)
28/6/00	44.5 (36.7)	109.8 (101.5)	25.3 (18.7)	25.1 (18.6)	22.0 (15.6)
6/9/00	29.1 (21.3)	32.2 (23.9)	105.1 (98.6)	45.4 (38.9)	39.9 (33.6)
7/9/00	29.6 (21.8)	32.1 (23.7)	45.6 (39.0)	107.9 (101.4)	43.2 (36.9)
8/9/00	31.0 (23.2)	28.8 (20.5)	41.1 (34.5)	44.3 (37.8)	105.5 (99.2)
11/9/00	29.7 (21.9)	29.7 (21.4)	37.4 (30.8)	37.9 (31.4)	37.2 (30.9)
13/9/00	26.8 (19.0)	30.6 (22.3)	34.0 (27.5)	33.6 (27.1)	34.3 (27.9)
27/9/00	31.4 (23.6)	32.7 (24.4)	34.8 (28.2)	32.5 (26.0)	32.4 (26.0)
18/10/00	25.9 (18.1)	30.7 (22.3)	29.1 (22.6)	30.9 (24.4)	29.4 (23.0)
	11/9/00	13/9/00	27/9/00	18/10/00	
27/6/00	25.5 (18.8)	20.2 (14.3)	26.8 (20.1)	22.2 (15.5)	
28/6/00	23.9 (17.2)	21.7 (15.8)	26.2 (19.6)	24.7 (18.0)	
6/9/00	38.3 (31.6)	30.7 (24.8)	35.5 (28.8)	29.9 (23.2)	
7/9/00	39.1 (32.4)	30.5 (24.6)	33.3 (26.6)	31.8 (25.1)	
8/9/00	39.3 (32.6)	31.9 (26.0)	34.0 (27.3)	31.0 (24.3)	
11/9/00	108.4 (101.6)	36.0 (30.1)	37.5 (30.8)	30.3 (23.6)	
13/9/00	40.9 (34.2)	110.0 (104.1)	34.6 (27.9)	32.8 (26.1)	
27/9/00	37.6 (30.9)	30.6 (24.7)	106.7 (100.0)	39.2 (32.5)	
18/10/00	30.3 (23.6)	28.9 (23.0)	39.0 (32.3)	104.1 (97.4)	

Fishergate Site A 8:00 – 9:00 recurrence rate

	25/6/01	26/6/01	27/6/01	28/6/01	29/6/01
25/6/01	116.8 (105.5)	46.1 (35.4)	45.5 (34.5)	46.1 (34.7)	43.4 (32.3)
26/6/01	48.4 (37.1)	114.9 (104.2)	52.1 (41.1)	53.4 (42.0)	44.3 (33.2)
27/6/01	46.7 (35.5)	51.1 (40.3)	118.6 (107.6)	52.8 (41.4)	45.7 (34.6)
28/6/01	45.5 (34.2)	50.3 (39.5)	50.8 (39.8)	118.3 (106.9)	47.1 (35.9)
29/6/01	43.8 (32.5)	42.7 (31.9)	44.9 (33.9)	48.0 (36.6)	115.9 (104.8)
2/7/01	48.0 (36.7)	44.0 (33.3)	44.0 (33.0)	45.9 (34.6)	43.3 (32.2)
3/7/01	44.1 (32.9)	49.9 (39.1)	46.6 (35.6)	49.1 (37.7)	43.3 (32.2)
4/7/01	40.8 (29.6)	42.5 (31.8)	49.7 (38.7)	46.5 (35.1)	43.3 (32.1)
5/7/01	42.7 (31.4)	43.1 (32.4)	44.0 (33.0)	51.1 (39.7)	40.1 (29.0)
6/7/01	38.5 (27.2)	42.1 (31.3)	41.0 (30.1)	43.6 (32.2)	44.5 (33.4)
11/7/01	39.1 (27.9)	40.6 (29.8)	46.9 (35.9)	43.4 (32.0)	37.7 (26.5)
12/7/01	37.5 (26.3)	41.1 (30.4)	41.3 (30.4)	44.7 (33.3)	38.0 (26.8)
13/7/01	35.2 (24.0)	35.3 (24.6)	37.0 (26.0)	36.7 (25.4)	40.9 (29.7)
16/7/01	43.2 (32.0)	38.8 (28.1)	41.7 (30.8)	41.0 (29.6)	37.4 (26.3)
	2/7/01	3/7/01	4/7/01	5/7/01	6/7/01
25/6/01	46.7 (35.7)	35.5 (26.4)	34.3 (24.8)	34.7 (25.5)	32.1 (22.7)
26/6/01	44.9 (34.0)	42.0 (33.0)	37.4 (28.0)	36.7 (27.6)	36.7 (27.4)
27/6/01	44.0 (33.0)	38.5 (29.4)	42.8 (33.4)	36.7 (27.6)	35.1 (25.8)
28/6/01	44.2 (33.2)	39.0 (29.9)	38.6 (29.1)	41.1 (31.9)	35.9 (26.5)
29/6/01	42.5 (31.6)	35.1 (26.1)	36.6 (27.2)	32.9 (23.8)	37.4 (28.0)
2/7/01	112.6 (101.6)	44.5 (35.4)	42.6 (33.1)	39.7 (30.5)	37.7 (28.3)
.
.

Fishergate Site A 8:20 – 8:40 vs entire sample rec. rate

	25/6/01	26/6/01	27/6/01	28/6/01	29/6/01
25/6/01	123.7 (107.7)	57.5 (42.2)	55.8 (40.3)	55.4 (39.1)	52.0 (36.2)
26/6/01	54.7 (38.8)	116.0 (100.7)	61.0 (45.5)	62.6 (46.3)	51.4 (35.5)
27/6/01	54.9 (38.9)	55.3 (40.0)	116.6 (101.1)	64.2 (47.9)	54.2 (38.4)
28/6/01	56.3 (40.3)	59.2 (43.8)	61.0 (45.5)	130.5 (114.2)	52.3 (36.4)
29/6/01	54.1 (38.1)	48.6 (33.2)	53.9 (38.4)	57.8 (41.5)	115.2 (99.4)
2/7/01	56.2 (40.2)	55.1 (39.8)	54.0 (38.5)	57.7 (41.4)	52.7 (36.9)
3/7/01	50.6 (34.7)	61.0 (45.7)	55.9 (40.5)	59.0 (42.7)	51.9 (36.1)
4/7/01	50.0 (34.0)	54.7 (39.4)	56.2 (40.7)	58.9 (42.6)	48.5 (32.7)
5/7/01	50.4 (34.4)	49.6 (34.3)	50.4 (34.9)	61.0 (44.7)	46.8 (30.9)
6/7/01	44.5 (28.5)	51.4 (36.1)	50.9 (35.4)	52.4 (36.1)	52.7 (36.9)
11/7/01	44.1 (28.2)	49.7 (34.4)	54.5 (39.0)	56.9 (40.6)	46.0 (30.2)
12/7/01	43.9 (27.9)	46.4 (31.0)	48.6 (33.1)	52.8 (36.5)	42.7 (26.9)
13/7/01	42.7 (26.8)	39.6 (24.3)	46.3 (30.8)	46.7 (30.4)	45.4 (29.5)
16/7/01	50.4 (34.4)	44.4 (29.1)	46.9 (31.4)	48.9 (32.6)	41.4 (25.6)
	2/7/01	3/7/01	4/7/01	5/7/01	6/7/01
25/6/01	58.1 (42.2)	46.9 (33.6)	44.4 (30.8)	46.9 (33.8)	45.5 (32.1)
26/6/01	54.7 (38.8)	51.4 (38.0)	44.8 (31.3)	44.1 (31.0)	43.5 (30.1)
27/6/01	54.2 (38.3)	48.3 (34.9)	49.7 (36.1)	45.6 (32.5)	43.8 (30.4)
28/6/01	51.5 (35.6)	49.5 (36.2)	47.4 (33.8)	51.7 (38.6)	44.3 (30.9)
29/6/01	51.4 (35.5)	45.3 (31.9)	44.2 (30.6)	40.2 (27.1)	45.5 (32.1)
2/7/01	117.6 (101.6)	56.4 (43.0)	52.9 (39.4)	45.3 (32.2)	47.3 (34.0)
.
.

The General Linear Model

Hypothesised GLM

The data suggests the following GLM

$$E[R] = \beta_0 + \beta_1|d| + \beta_2I_w + \beta_3I_d.$$

Where

- R is the corrected recurrence rate between the two days.
- β_i are the model parameters.
- $|d|$ is the number of days (ex. weekends) between the two.
- I_w is one if the days are in different weeks.
- I_d is one if the days are the same day of the week.

GLM fitting with highest recurrence rate

Site	β_0	β_1	β_2	β_3	R_a^2	p-value
A	42.2 (0.1%)	-0.800 (0.1%)	-3.35 (1%)	5.41 (0.1%)	0.545	1.73×10^{-15}
B	24.6 (0.1%)	-0.476 (low)	-3.40 (low)	0.0966 (low)	0.0225	0.198
C	37.6 (0.1%)	-0.510 (0.1%)	-3.00 (5%)	4.39 (0.1%)	0.386	6.79×10^{-10}
D	37.4 (0.1%)	-0.780 (0.1%)	-2.35 (10%)	5.18 (0.1%)	0.556	1.32×10^{-11}
E	53.5 (0.1%)	-0.731 (0.1%)	-5.36 (0.1%)	2.90 (1%)	0.644	$< 2.2 \times 10^{-16}$
F	34.2 (0.1%)	-0.597 (0.1%)	-2.46 (low)	2.01 (low)	0.242	5.36×10^{-6}
G	40.1 (0.1%)	-1.00 (0.1%)	-4.32 (1%)	5.11 (0.1%)	0.614	6.18×10^{-16}
H	25.5 (0.1%)	-0.360 (1%)	-1.23 (low)	3.35 (1%)	0.203	0.000180
I	35.9 (0.1%)	-0.831 (0.1%)	-3.23 (1%)	3.70 (1%)	0.540	1.361×10^{-15}
J	39.2 (0.1%)	-0.412 (1%)	-5.41 (0.1%)	2.49 (10%)	0.389	1.27×10^{-8}
K	42.7 (0.1%)	-0.719 (0.1%)	-4.87 (0.1%)	3.49 (1%)	0.636	$< 2.2 \times 10^{-16}$
All *	39.0 (0.1%)	-0.666 (0.1%)	-3.71 (0.1%)	3.72 (0.1%)	0.222	$< 2.2 \times 10^{-16}$

Conclusions and Further Work (1)

- Note that while the original study was intended
- Not only does the make up of traffic vary considerably from day to day but the recurrence rate drops off very quickly.
- The combined model shows a recurrence rate of 39% falling off at around 0.7% per day. Obviously this model can only be considered as valid for more than the few weeks being considered.
- Days which are in different weeks differ suffer a reduced recurrence rate (3.7%).
- Days which are on the same day of the week have an increased recurrence rate (3.7%).

Conclusions and Further Work (2)

- Some sites have a considerably higher recurrence rates than others.
- No pattern has yet been determined in why some sites have a higher recurrence rate or whether this is a long-term feature of the sites or a feature of the specific time surveyed.
- If we want to accurately model a traffic system we must account for how quickly the vehicles in the system are being replaced.
- At the moment this work is relying on only data from York more data sets will be analysed soon (more data is welcome).

Bibliography

- 1) This talk online at www.richardclegg.org/pubs/.
- 2) Clegg(2004) "The Statistics of Dynamic Networks" PhD Thesis. Dept of Maths. University of York. Online at: www.richardclegg.org/pubs/thesis.pdf
- 3) Clegg and Batley(2001) "Driver route Choice and Departure time Choice: The Models and the Evidence" Presented at UTSG 2001. Available online at www.richardclegg.org/pubs/utsg2001.doc
- 4) Huff and Hanson (1986). "Repetition and variability in urban travel." Geographical Analysis, 18(2) pp 97–114.
- 5) Pendyala (2003) "Measuring day-to-day variability in travel behavior using GPS data". Report for Federal Highway Administration. Available online at: www.fhwa.dot.gov/ohim/gps/index.html