



Use of vehicle test data and changes in mileage patterns over time

James Thomas and R. Eddie Wilson

j.thomas@bristol.ac.uk

RE.Wilson@bristol.ac.uk

22nd June 2023

UK vehicle safety testing regime

VT20 MOT Test Certificate
Vehicle & Operator Services Agency

Note: If you have doubts as to whether this certificate is valid, please use the service described in note 3 overleaf to check.

MOT test number	Make	Odometer reading
761710136293	VAUXHALL	105420 Miles
Registration mark	Model	Test class
T203UNP	ASTRA	IV
Vehicle identification or chassis number	Colour	Approximate year of first use
W0LO7GF35X8091395	WHITE	1999
Expiry date	Issue date/time	Fuel type
AUGUST 25th 2007 (ZERO SEVEN)	AUGUST 18th 2006 (ZERO SIX) 13:30	Petrol
Authentication number	Design gross weight (goods vehicles)	kg
084907914489518556410227		
For all vehicles with more than 8 passenger seats	Advisory notice issued	NO
Seat belt installation checked this test	Test station number	80572
N/A		
Number of seat belts fitted at time of installation check	Previous installation check date	N/A
N/A		
Issue's name in CAPITALS	Signature of issuer	
D. S. BRYANT		
Warning: A test certificate is not evidence that the vehicle is in a satisfactory condition for use on the road.		
Check carefully that the above details are correct. Do not accept a certificate which has been altered.		
Reg Mark	Make	VTS Number
T203UNP	VAUXHALL	80572
MOT Expiry		
AUGUST 25th 2007 (ZERO SEVEN)		
Inspection Author		
HANHAM MOTOR COMPANY		
126 BRYANTS HILL		
ST GEORGE		
BRISTOL		
BS5 8RJ		

Odometer reading
105420 Miles

- ▶ 'Ministry of Transport' or MOT test is the UK's annual safety inspection for all road vehicles older than 3 years
- ▶ Since 2005: results have been captured and stored digitally
- ▶ Since 2010: data for GB published online by Department for Transport (spanning back to 2005)
- ▶ **Key interest:** the *odometer reading* recorded at each test

Sample of the data

test_id,	vehicle_id,	test_date,	test_mileage,	test_class_id,	test_type,	test_result,	postcode_area,	make,	model,	colour,	fuel_type,	cylinder_capacity,	first_use_date,
31453589,	108224,	2019-05-08,	123096,	4,	RT,	P,	KA,	HONDA,	CIVIC,	GREY,	PE,	1998,	2004-03-17,
469430857,	108224,	2020-11-16,	131593,	4,	NT,	F,	KA,	HONDA,	CIVIC,	GREY,	PE,	1998,	2004-03-17,
331854923,	108224,	2020-11-27,	131718,	4,	RT,	P,	KA,	HONDA,	CIVIC,	GREY,	PE,	1998,	2004-03-17,
321768309,	108224,	2021-12-02,	138490,	4,	NT,	P,	TD,	HONDA,	CIVIC,	GREY,	PE,	1998,	2004-03-17,
851562230,	110222,	2006-02-16,	81165,	4,	NT,	P,	DA,	VAUXHALL,	ASTRA,	RED,	PE,	1389,	1997-08-08,
894210716,	110222,	2006-12-13,	90016,	4,	NT,	P,	TN,	VAUXHALL,	ASTRA,	RED,	PE,	1389,	1997-08-08,
936859202,	110222,	2008-02-07,	96957,	4,	NT,	P,	TN,	VAUXHALL,	ASTRA,	RED,	PE,	1389,	1997-08-08,
979507688,	110222,	2009-02-06,	104701,	4,	NT,	P,	TN,	VAUXHALL,	ASTRA,	RED,	PE,	1389,	1997-08-08,
69512266,	110222,	2010-01-29,	108972,	4,	NT,	P,	TN,	VAUXHALL,	ASTRA,	RED,	PE,	1389,	1997-08-08,
1719080396,	111228,	2013-09-10,	25475,	4,	NT,	F,	B,	FORD,	FIESTA,	SILVER,	PE,	1242,	2010-09-27,
1761728882,	111228,	2013-09-10,	25475,	4,	RT,	P,	B,	FORD,	FIESTA,	SILVER,	PE,	1242,	2010-09-27,
1187841506,	111228,	2014-09-11,	29064,	4,	NT,	PRS,	B,	FORD,	FIESTA,	SILVER,	PE,	1242,	2010-09-27,
1500011509,	111228,	2015-09-22,	33333,	4,	NT,	F,	B,	FORD,	FIESTA,	SILVER,	PE,	1242,	2010-09-27,
1918972575,	111228,	2015-09-22,	33353,	4,	RT,	P,	B,	FORD,	FIESTA,	SILVER,	PE,	1242,	2010-09-27,

- ▶ Pseudonymised vehicle_ID links tests on the same vehicle
- ▶ Metadata includes postcode_area (location) and vehicle attributes

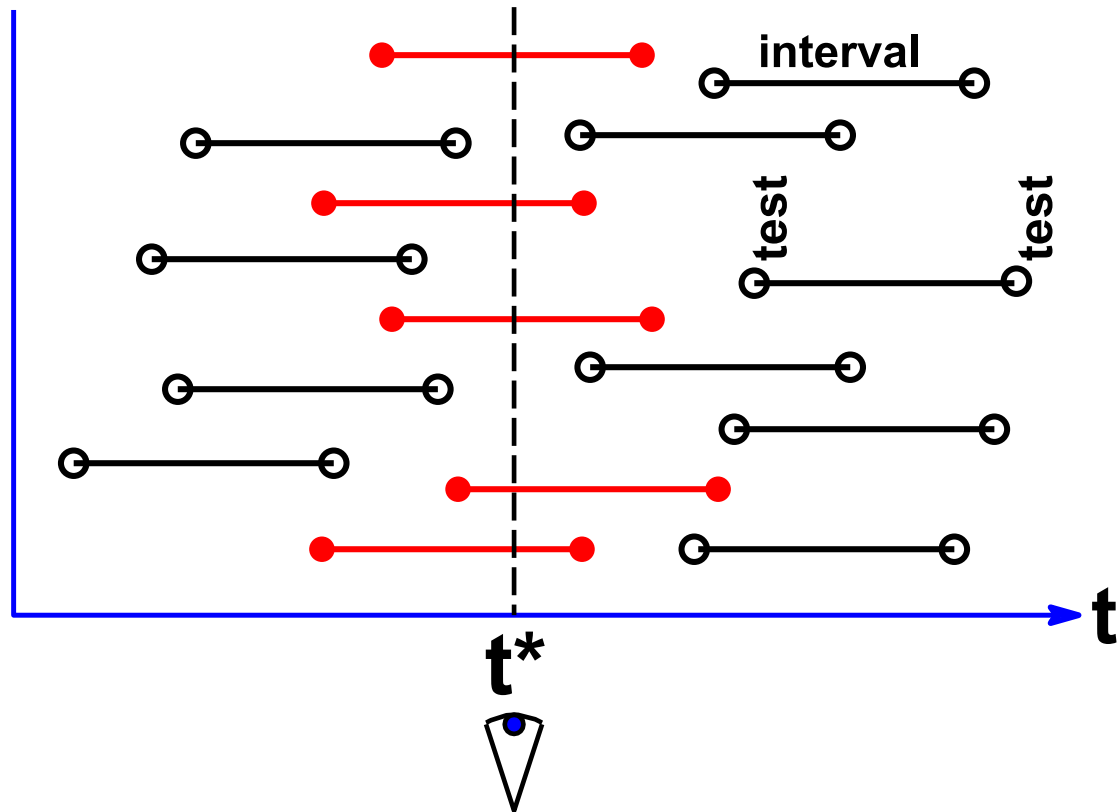
Basic analysis object: an interval and its attributes

- ▶ Re-arrange blocks of same-vehicle data into consecutive pairs of tests:

Interval	First test			Second test		
	date ₁	odometer ₁	place ₁	date ₂	odometer ₂	place ₂
1	2005-08-26	99 777	BS	2006-08-18	105 420	BS
2	2006-08-18	105 420	BS	2007-08-13	113 709	BS
3	2007-08-13	113 709	BS	2008-08-11	123 259	BS
4	2008-08-11	123 259	BS	2008-08-11	123 259	BS
5	2008-08-11	123 259	BS	2009-08-05	132 299	BS

- ▶ Calculate average distance driven per day: *mileage rate*
- ▶ To which can be linked vehicle-specific attributes:
VAUXHALL, ASTRA LS 8V, WHITE, P (fuel), 1598 (cc), 1999 (year)
- ▶ These mileage rates are (more or less) complete across the vehicle population — even after cleaning

Population level statistics: *straddling rate* $\bar{r}(t)$



- ▶ Select all intervals that *straddle* a given *observation date* t^*
- ▶ Each interval yields an average (per vehicle) rate r_i

- ▶ *Straddling rate* $\bar{r}(t^*)$ is then defined by **average average**:

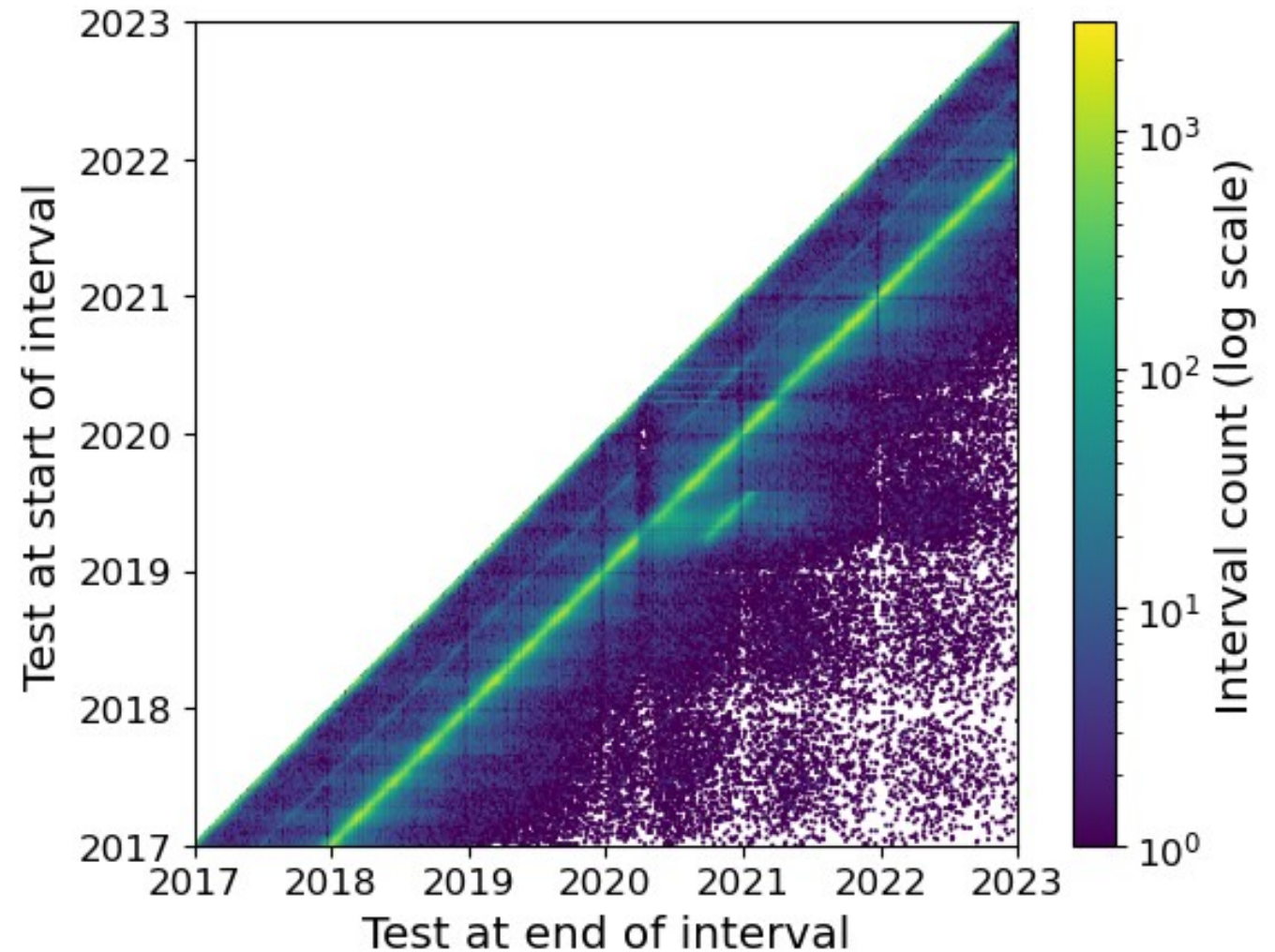
$$\bar{r}(t^*) = \frac{1}{N} \sum_{i=1}^N r_i$$

- ▶ For annual statistics, choose $t^* = 2017-07-01, 2018-07-01$, etc.
- ▶ But $\bar{r}(t^*)$ actually incorporates distance driven over the two year span:

$$t^* - 1 \leq t < t^* + 1$$

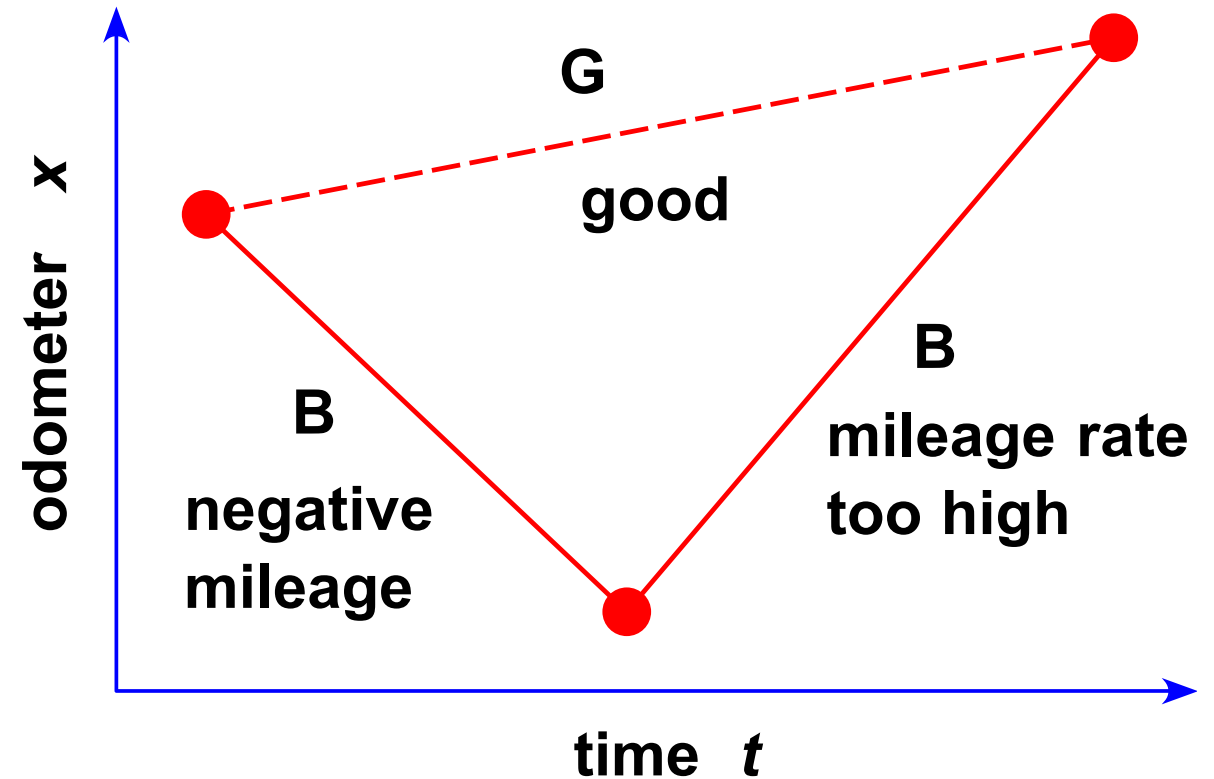
Data cleaning: examining interval duration

- ▶ Majority of intervals 1 year duration — annual testing regime
- ▶ Some intervals are very short (only days) — retests after MOT failure
- ▶ During COVID, the vehicle testing regime was relaxed from 12 to 18 months

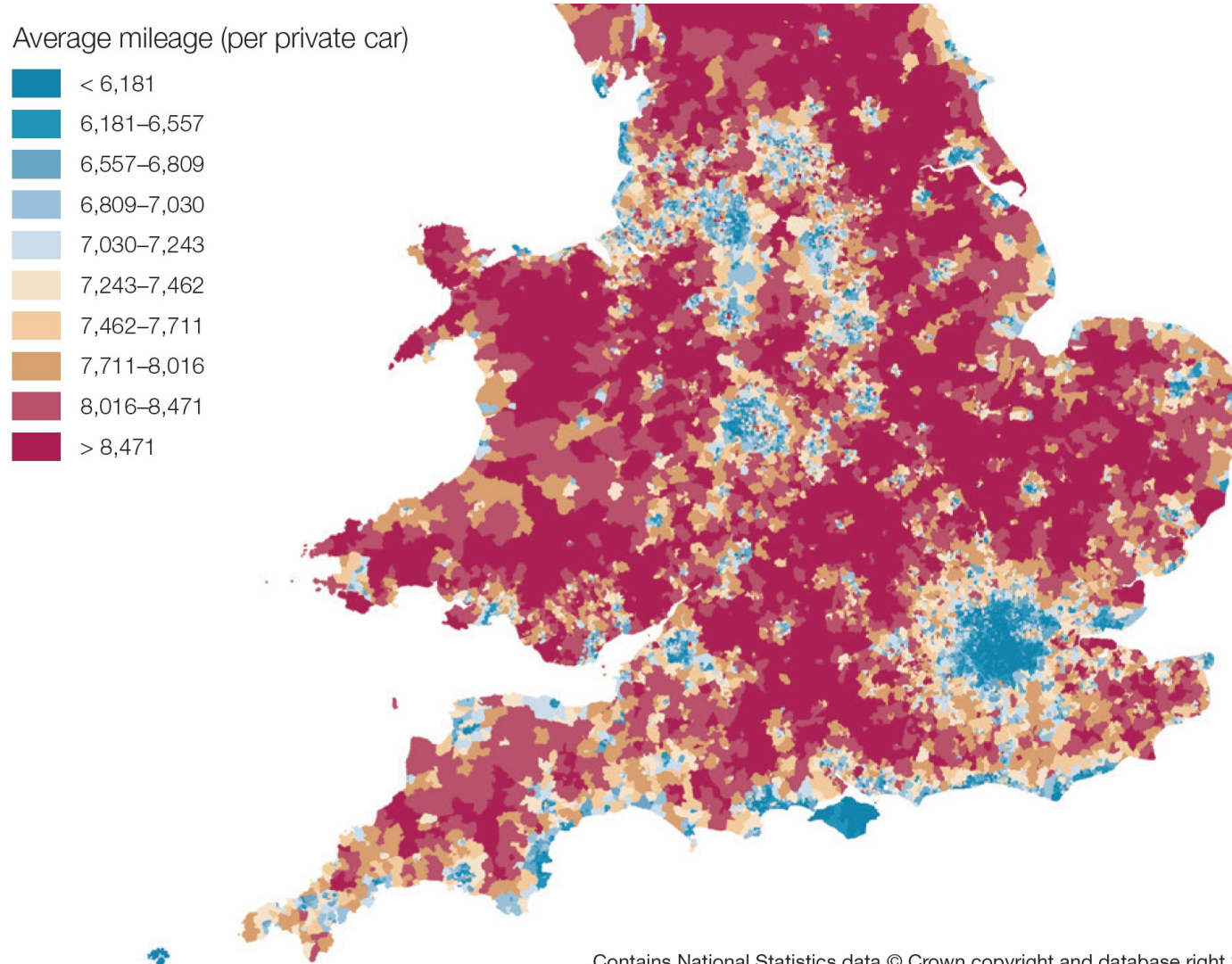


Data cleaning: when two **B**ads make a **G**ood

- ▶ Data entry is manual at the testing station
- ▶ Missing or interchanged digits are the most common
- ▶ Compute all intervals, then reject *bad intervals* with rates outside a reasonable range
 $r_i < 0$ or $r_i > 150$ (say)
- ▶ Recompute an interval spanning the middle entry



MOT projects: 2011–2017



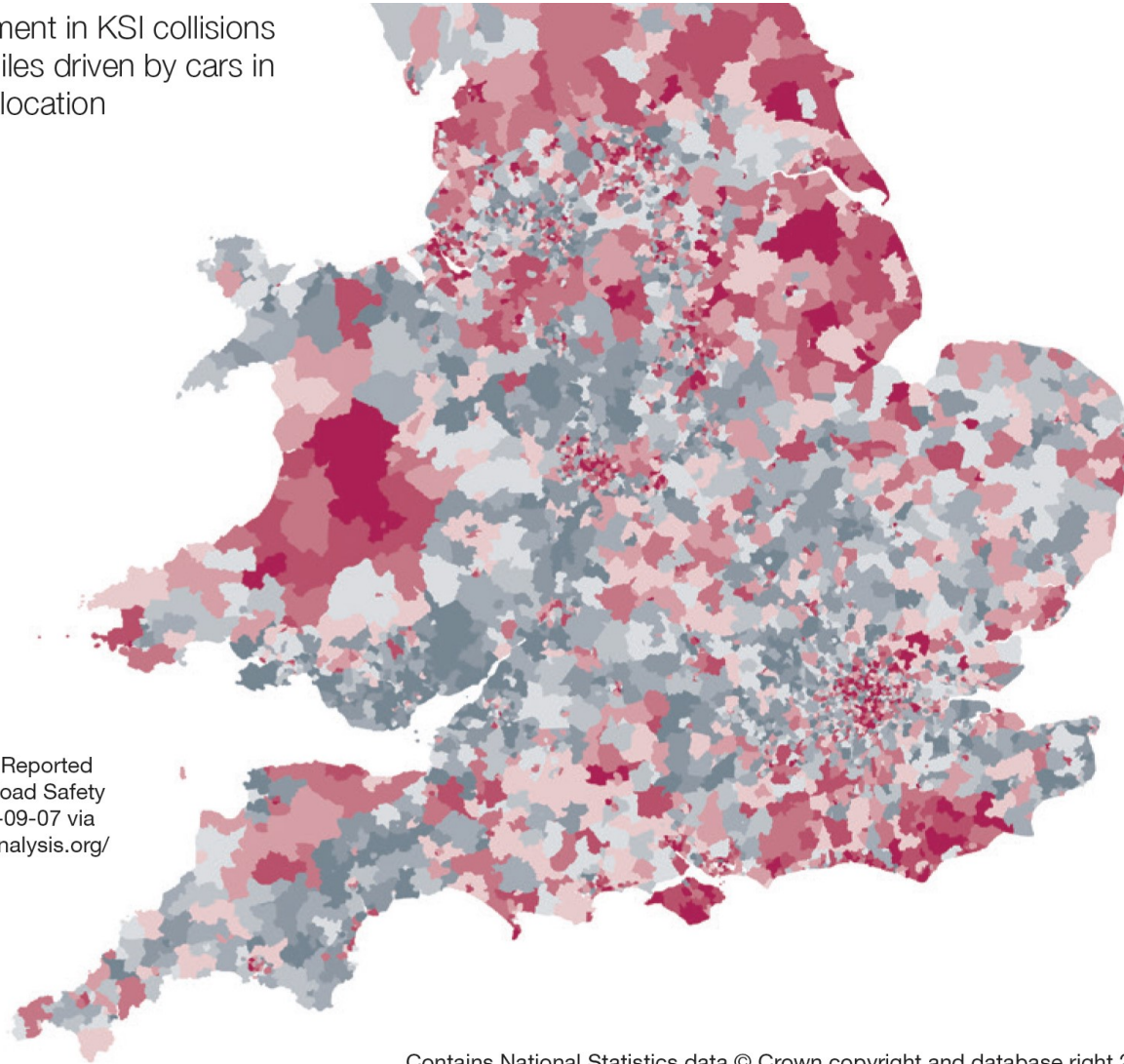
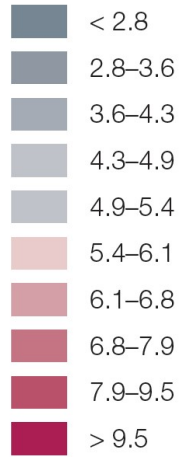
Contains National Statistics data © Crown copyright and database right 2016
Contains OS data © Crown copyright and database right 2016

- Confidential versions of dataset can be linked to *vehicle keeper* records
- Vehicles can be located to more precise geographies ('LSOA')

Cairns et al. (2017)
*MOToring Along: The lives of cars
seen through licensing and test data*

MOT projects: 2011–2017

Car driver involvement in KSI collisions
per 10,000,000 miles driven by cars in
the driver's home location



30 MAST Online (2017). Reported
road casualties tool by Road Safety
Analysis Accessed 2017-09-07 via
<https://mast.roadsafetyanalysis.org/>

Contains National Statistics data © Crown copyright and database right 2016
Contains OS data © Crown copyright and database right 2016

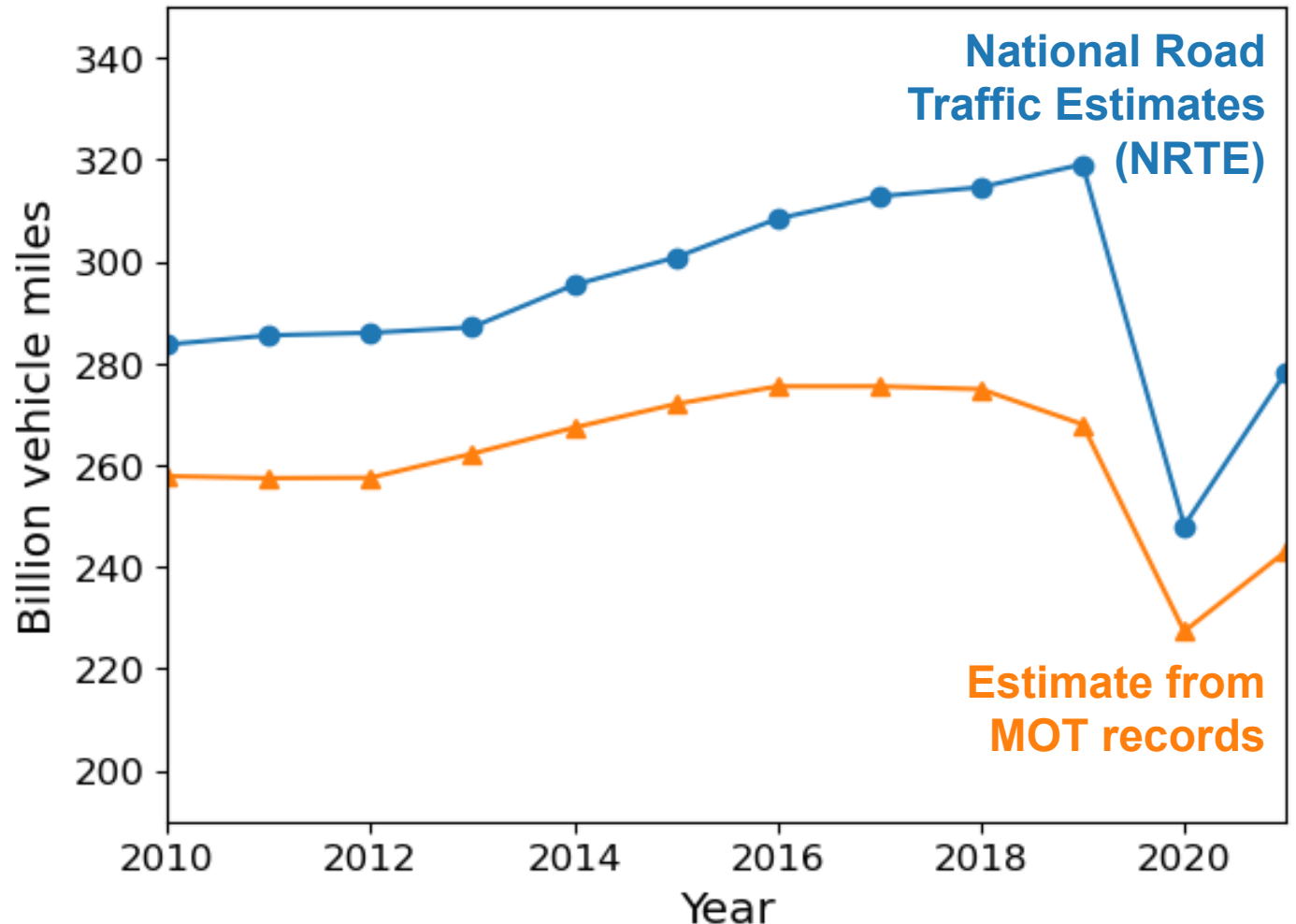
- ▶ Can also link with other datasets and statistics
- ▶ For example: poverty, deprivation, pollution, safety — accident rates

Cairns et al. (2017)

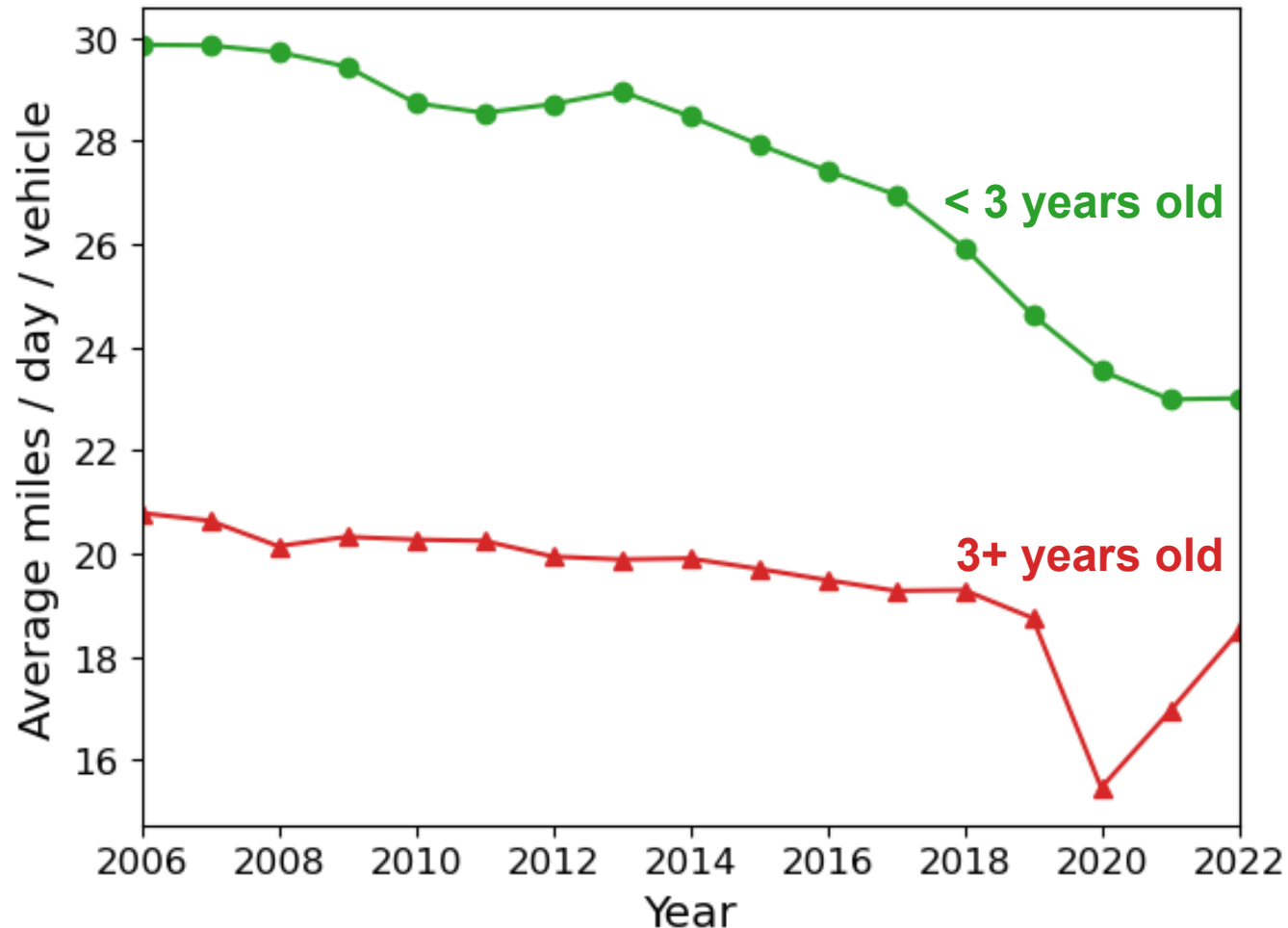
*MOToring Along: The lives of cars
seen through licensing and test data*

Reproducing official vehicle usage statistics

- ▶ The UK's **National Road Traffic Estimates (NTRE)** can be reproduced using **MOT records**
- ▶ Consistent with NRTE however regularly undershoots & smoother than official figures
- ▶ Pre-COVID, total distance driven was increasing — due to increasing vehicle population (older vehicles not retiring)

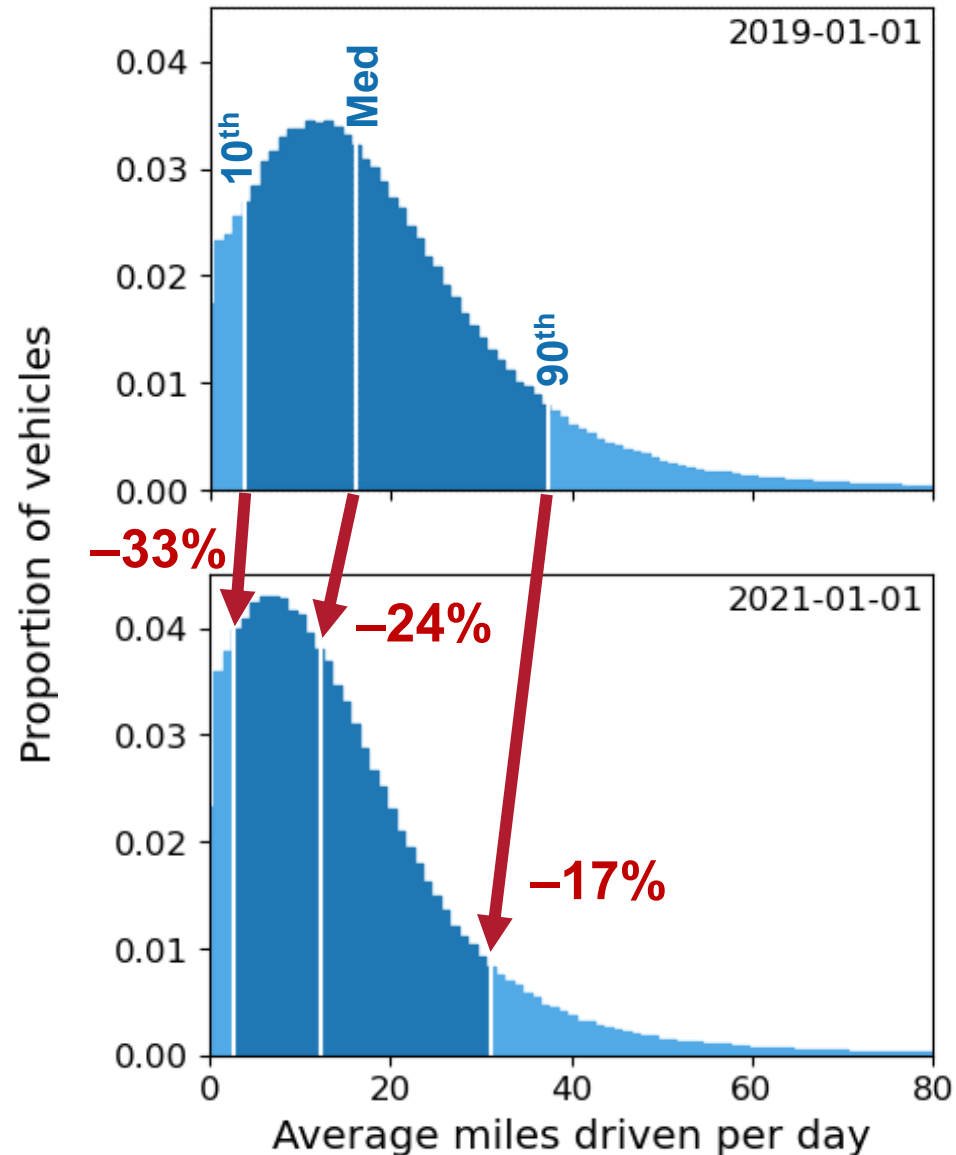


But individual vehicles are driven less



- ▶ Average per-vehicle mileage has fallen year on year
- ▶ Steady decline is particularly marked in vehicles **< 3 years old**
- ▶ Estimation method for vehicles **< 3 years old** causes smoothing and do not see sharp decrease during COVID that seen in vehicles **3+ years old**

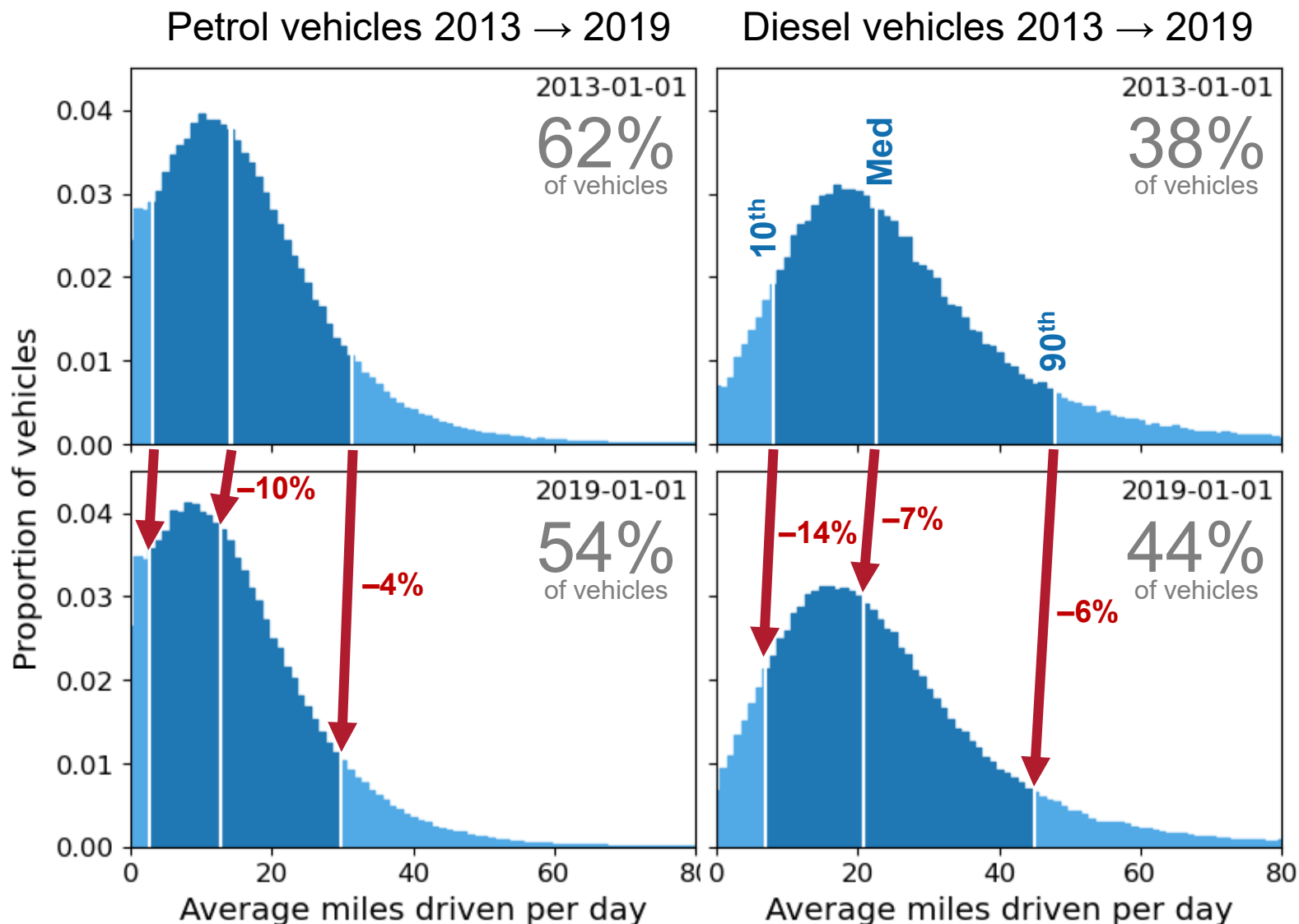
Distribution of *mileage rates* shifted during COVID



- ▶ Distribution of per-vehicle mileage rates shifted during COVID
- ▶ Mean mileage rate for 3+ y.o. vehicles: 19.3 → 15.4 (–20%) miles per day
- ▶ Reductions in the mean were:
 - ▶ **Less-marked** in urban SW London (–16%)
 - ▶ **More-marked** in rural NW Wales (–25%)
- ▶ Reductions were not uniform across the distribution — high mileage tails persisted

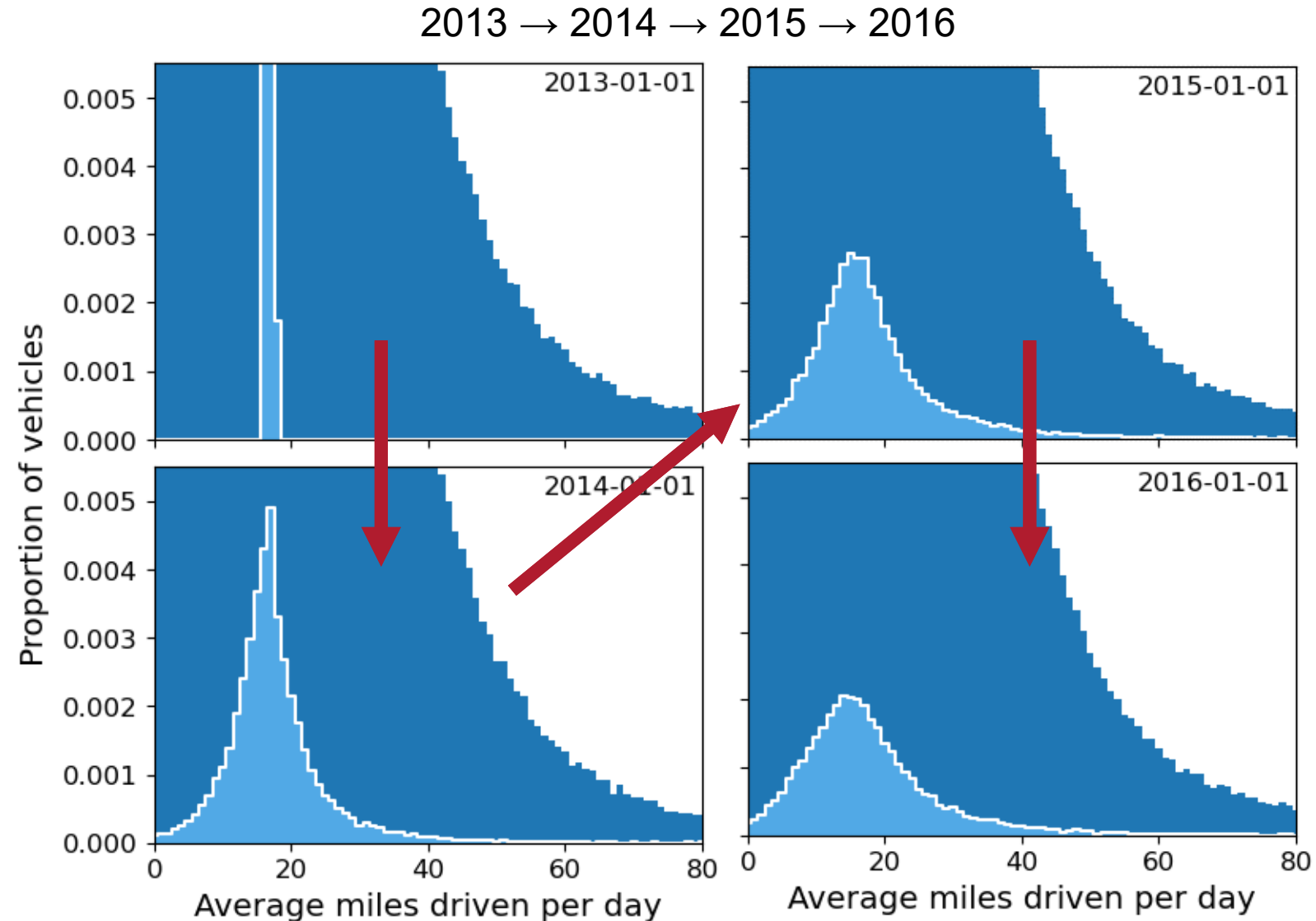
Interesting findings are not limited to COVID

- ▶ From 2013-2019, mean mileage rate reduced by:
 - ▶ 7% for petrol vehicles
 - ▶ 7% for diesel vehicles
 - ▶ **But** only 3% overall
- ▶ Share of petrol/diesel vehicles also changed:
 - ▶ 62% : 38% for 2013
 - ▶ 54% : 44% for 2019



Following individual vehicles from year to year

- ▶ Individual vehicle behaviour can also be tracked
- ▶ 5% slice of vehicles around median (2013)
- ▶ Distribution of mileage rates in following years is a convolution with a kernel function
- ▶ Interquartile range:
 - 0.8 (2013)
 - 6.3 (2014)
 - 8.9 (2015)
 - 10.6 (2016)



Conclusions

- ▶ In GB, vehicle safety test (MOT) data provides detailed descriptions of mileage across the full vehicle population
- ▶ Analyses show that:
 - ▶ It might replace official road traffic estimates (based on sectional counts)
 - ▶ Distributional differences in driving reductions across the COVID pandemic
- ▶ Current project is exploring further strata, plus a range of data linkage opportunities with other datasets



**Connecting Administrative
vehicle data for Research
on Sustainable transport**

<https://cars-project.github.io>

James Thomas j.thomas@bristol.ac.uk

Also: R. Eddie Wilson, Jillian Anable,
Ian Phillips, Malcolm Morgan,
Sally Cairns

DOI: [10.5281/zenodo.8051273](https://doi.org/10.5281/zenodo.8051273)