Safety in numbers: Do road-traffic injuries scale non-linearly with travel?

github.com/robj411/safety_in_numbers_power_law

October 17, 2021

Contents

Abstract

We consider the mathematical model used to describe the phenomenon of "safety in numbers" in relation to road-traffic injuries. We focus on a model commonly used to describe number of injuries as a function of total use of two transport modes – motor cars and cyclists – in large contiguous geographic areas, such as cities and counties. While models often include many covariates, the basic functional form does not account for the size of the area under study (or, equivalently, the density of road use per unit time and/or space). We present some consequences of this modelling choice in terms of interpretation of meaning and application in predictive modelling, and explore it further through simulations and analysis. We develop the simulation and analytic results into a new model that we demonstrate using data from England. We conclude that omission of the size of areas under study leads to confounding which results in a bias in the model, rendering it inadequate to make predictions of change in injury number following a change in transport mode use. Equivalently, the finding that the 'square root is a central estimate of "safety-in-numbers"' could be an artefact of a model fit to data that span scales, and interpreted as applying to numbers that change against a background of a fixed size.

1 Introduction

The term "safety in numbers" reflects the observation that a change in the number of road users of a specific type (commonly cyclists) is not met by a proportional change in the number of injuries caused to or by the road-user group. Equivalently, from an individual-level perspective, the more cyclists there are, the lower the risk is per cyclist. We describe this as "linearity": is the expected (total) number of injuries a linear function of the number of cyclists? If it is "sublinear", we might say that there is safety in numbers. Our challenge is to formulate this question in a mathematical model and take account of all relevant covariates.

The number of injuries over a certain period and space, I, is generally formulated as follows, in terms of a base rate, α, the number of road users of one type (say, motorists), M, the number of road users of the second type (say, cyclists), C, and "safety in numbers" exponents for each, β_1 and β_2 (Elvik and Bjørnskau, 2017):

$$I \sim \mathcal{F}(\lambda), \tag{1}$$

$$\lambda = \alpha M^{\beta_1} C^{\beta_2}, \tag{2}$$

where $\mathcal{F}$ is a function specifying a distribution, such as Poisson or negative binomial, and the expected number of injuries, $E(I)$, is the parameter λ. Then, an exponent less than 1 implies sublinear scaling. For completeness, one might include other covariates, e.g.

$$\lambda = \alpha M^{\beta_1} C^{\beta_2} \exp\left(\sum_{i=3}^{P} \beta_i X_i\right), \tag{3}$$

but these are not central to the present discussion, which focuses around the variables in Equation 2.

Beyond descriptions of observed data, such analyses are proposed to inform public-health forecasting (Schepers and Heinen, 2013) and to make predictions in novel scenarios (de Sá et al., 2017; Jaller et al., 2020). This includes assessment of likely health benefits following policy change, as well as forecasting healthcare needs given the expected change to transport-related behaviour, e.g. the increase in motor vehicle ownership expected in cities in fast-growing economies in the coming years, especially in light of Sustainable Development Goal 3.6 to reduce road-traffic casualties to half of their 2011 value.[1] We have previously used this formulation to make predictions of numbers of road injuries for specific cities (Accra, Sao Paulo), as well as in the developing generic software ITHIM-R. (ITHIM: integrated transport and health impact model.) To make predictions in novel scenarios requires an assumption of causality, and the validity of a causal relationship relies on having accounted for all confounders.

[1] `http://iris.wpro.who.int/handle/10665.1/12878`

In this work, we question the validity of the assumptions implicit in the model as well as its applications. Our intention here is to examine (1) the interpretation of Equation 2, (2) its application to health-impact modelling, and (3) how the relationship between safety and number of road users might be measured. Note that there have been two types of study regarding safety in numbers: studies comparing multiple roads or road junctions in the same city, and studies comparing different cities (or geographical regions containing whole road networks). In this work we focus on the latter. However, we must keep the former in mind, because its results have been used in applications making predictions for the latter type. Ultimately, we would like to have a single comprehensive framework that joins observations at the two scales together.

Using a theoretical approach, we derive the result that coefficients $\beta_1 + \beta_2 = 1$ in Equation 2 learnt from multiple settings, which differ in scale but have the same road-user density, correspond to linear scaling across time and space. Put another way, these coefficients permit "tiling" of a small space to create a larger space whose properties are the same in terms of risk per unit space. We confirm this result through simulation, which shows also that, again under Equation 2, the coefficients $\beta_1 + \beta_2 = 2$ will tend to describe the variations seen in injury count data from multiple areas of the same size but different road-user density. These provide null hypotheses that injuries are linear in road use for studies assessing the impact of road-user number on collision risk. It is useful to understand "tiling" in contrast with mode shifts: tiling preserves properties across space and time, allowing us to make predictions for longer or shorter time periods, or larger or smaller areas, where we expect road use to follow scale. We contrast this with predictions where the time and space stay the same and the extent of road use changes.

We present an alternative model which includes the components of Equation 2 and additionally it explicitly includes size as a covariate. (Note that size might be parametrised in terms of area, population, road length, or any other relevant metric.) We use exponents δ_1 and δ_2 to parametrise this model, to distinguish it from existing models. The cases discussed above (scaling size, constant density ($\beta_1 + \beta_2 = 1$) and constant size, scaling density ($\beta_1 + \beta_2 = 2$)) are special cases of this model. We apply this model to data for England. Our results are consistent with the preceding theoretical and simulation analyses. With this model, we find variously that the data support or do not support the null hypothesis of linearity of injuries with respect to road-user density – that is, for doubled travel, we expect fewer than twice the number of injuries – depending on their particular conformation.

The main purposes of this work are (a) to identify an acceptable model for prediction of injury numbers under a shift in road use, (b) to open new avenues for research into road-traffic injury dynamics, and (c) to encourage mode-shift modellers to consider exponents closer to 1, recognising that the "square root" is the central estimate of a system where the size of the space changes and the density stays the same. We highlight some important features of the existing model and propose an amended model for prediction. We identify areas we believe are most in need of attention, which are: accounting for size/density; the definition of "density"; the link between small-scale and city-level studies; how we interpret non-linearity where a predictor is a group of modes; and how we identify "non-linearity".

2 Background

2.1 Equation 2 in the literature

"Safety in numbers" exponents are estimated in analyses of road-traffic injuries through fitting a regression model such as Equation 2 or Equation 3 to data. At minimum, these data consist of counts of road injuries, and of two road user types, often cars and pedestrians or cyclists. The exponents β_1 and β_2 are estimated, and reported with confidence intervals and p-values corresponding to the probability of observing the data under the null hypothesis $\beta_1 = 0$, $\beta_2 = 0$, which is the default in most statistical software.

Various different units of measurement have been used in studies reporting safety-in-numbers effects. For example, Miranda-Moreno et al. (2011); Nordback et al. (2014); Geyer et al. (2006); Gårder et al. (1998); Schepers et al. (2011) and Leden (2002) count vehicles (usually in terms of average daily number of vehicles per unit space, though sometimes the time unit is annual or hourly). In contrast, Prato et al. (2016) and Schepers and Heinen (2013) work in terms of km travelled by each road-user group. Injuries are counted as the total in one or more years. The areas considered in the studies range in size from intersections (Nordback et al., 2014) to municipalities (Schepers and Heinen, 2013) and "local authorities" (Aldred et al., 2017).

Time and space are not typically explicitly included in published models. Instead, it is implicit that the numbers of injuries are linear in time and space. This might be crucial to the results found in safety studies (see Section 2.3).

In terms of prediction, exponents can be used as "laws", as in de Sá et al. (2017) and Jaller et al. (2020), where learnt exponents are applied alone, outside the context of the training data and the full model. This differs from what we usually understand of making predictions from models, in which all parameters from a model are used to compute its prediction(s). For ITHIM-R/Tigthat, as in de Sá et al. (2017), we take our current observations (obs) and make a prediction (pred) as follows:

$$I_{\mathrm{pred}} = I_{\mathrm{obs}} \frac{\alpha M_{\mathrm{pred}}^{\beta_1} C_{\mathrm{pred}}^{\beta_2}}{\alpha M_{\mathrm{obs}}^{\beta_1} C_{\mathrm{obs}}^{\beta_2}} = I_{\mathrm{obs}} \left(\frac{M_{\mathrm{pred}}}{M_{\mathrm{obs}}}\right)^{\beta_1} \left(\frac{C_{\mathrm{pred}}}{C_{\mathrm{obs}}}\right)^{\beta_2}. \tag{4}$$

Note that the dependence on the base rate α is lost and injuries are predicted as a function of the old observed injury counts and travel, the new/predicted travel, and the parameters β. Similarly, were we to use a model of the form of Equation 3, the exponential terms would cancel out.

In this way, the method relates to power-scaling laws in the biological sciences, dating back to metabolic rate as a function of size (Kleiber, 1947), and extending to social and economic sciences (Bettencourt et al., 2007). Laws hypothesised resulting from models such as these have been widely discussed in terms of both methods and implications (Stumpf and Porter, 2012; Clauset et al., 2009), which have some relevance for our example.

2.2 Illustration with data for England

To give some context, we present data pertaining to road-traffic injuries recorded in England in the years 2005 to 2015 for 148 areas, including counties and boroughs. From among these data we isolate all injuries to cyclists that occurred in events involving at least one car. Alongside these data, we use Road Traffic Statistics estimates of distance travelled by cars and bikes in these areas[2], as well as census estimates for population numbers in the year 2011.

We begin with some descriptive figures demonstrating what the data look like, and what a safety-in-numbers analysis represents. For simplicity we illustrate the principle using one mode alone: just cyclists, without consideration of other modes with which there were collisions. We show how injuries to cyclists (involving cars) scales with the total distance travelled by cyclists (Figure 1). The theory behind safety in numbers is that if the number of injuries is sublinear in

[2]We use data covering all road types provided by RTS minus that on motorways, calculated from `https://roadtraffic.dft.gov.uk/downloads`.

the independent variable (here, bike distance), then there is safety in numbers. As illustrated in the scatter plots, the question becomes "can we reject the null hypothesis that line of best fit has a gradient of one?" In this Figure, we show how the gradient changes as we include only more severe cases. We will not return to the issue of severity but it is interesting to note that the conclusion we draw about safety in numbers might depend on what we classify as an "injury". We repeat the analysis with London boroughs only in Figure 10.

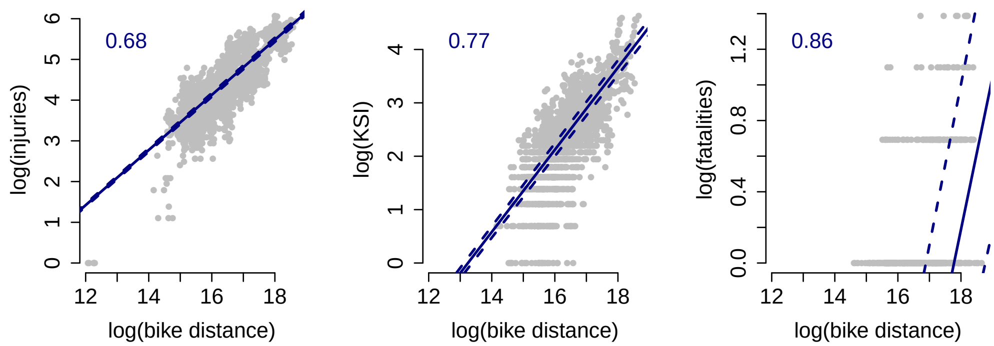

Figure 1: Fit of equation $I = \alpha C^{\beta}$ to data for local areas in England. I is (left) total incidents involving a car in which a cyclist was injured; (middle) the subset of these that were KSI; and (right) the subset of these that were fatal. C counts the total cycle distance travelled in the area. In blue is the line of best fit, with the gradient of the line (the value for β) in the top-left corner.

We look also at many possible relationships between single predictors (independent variables) and injury counts (Figure 11). These values are summarised in Table 1. We notice that the values of the coefficients are similar for each level of severity for all predictors, and that the coefficients are largest when the less serious casualties are excluded (KSI: killed or seriously injured). Finally, we have included the two-variable model of Equation 2, and reported the sum of the coefficients along with the singular coefficients for the single predictors. We note that this row aligns well with the other rows.

Table 1: Coefficients

Predictor	All injuries	KSI	Fatalities
Total travel (β)	0.77	0.90	1.11
Population (β)	0.78	0.89	0.98
Bike (β)	0.68	0.77	0.86
Car (β)	0.75	0.88	1.11
Bike and Car ($\beta_1 + \beta_2$)	0.67	0.79	0.97

This represents the key insight of this work: in general, it is posited that $\beta_i < 1$ for an equation of the form of Equation 2 implies safety in numbers for mode i (Elvik and Bjørnskau, 2017). Given the form of Equation 2, we should expect to find $\beta_i < 1$ for all i for any study counting motorists, cyclists and injuries across scales, because the mode counts C and M are functions of scale (see e.g. Figure 12). Importantly, this finding would be completely consistent with linearity in injuries with respect to cyclists and with respect to motorists.

The act of decomposing the population N into two component parts (M and C) splits the exponent associated with N (the coefficient β) between M and C (the coefficients β_1 and β_2). This presents an explanation for the strong agreement of the sum of coefficients with the singular coefficients in Table 1.

2.3 Linearity in Equation 2

Equation 2 is parametrised in terms of amounts of road users, and does not include terms to set these amounts into context, such as time or space. The equation does not distinguish between differences in number due to extended measurement, i.e. recording injuries over a larger area or time period, and an increase in number due to there being more road users in the same area and time period, i.e. a greater density of road users. The consequence is that each parametrised model is particular to its own setting, which vary in scale from hours to years, and junctions to counties. Despite these ranges in scale, remarkably consistent values are reported for effects of "safety in numbers".

The absence of accounting for scale might be what gives rise to the ubiquitous observation of "safety in numbers". To illustrate with a simple example, consider a single observation of injury number $I = \alpha M^{\beta_1} C^{\beta_2}$, and make a second observation identical to the first, which, added to the first, gives $2I = \alpha(2M)^{\beta_1}(2C)^{\beta_2}$. We can solve these equations together to learn about the β values where we have assumed linearity:

$$\frac{2I}{I} = \frac{\alpha(2M)^{\beta_1}(2C)^{\beta_2}}{\alpha M^{\beta_1} C^{\beta_2}}; \tag{5}$$

$$2 = \frac{2^{\beta_1} M^{\beta_1} 2^{\beta_2} C^{\beta_2}}{M^{\beta_1} C^{\beta_2}}; \tag{6}$$

$$= \frac{2^{\beta_1} 2^{\beta_2}}{1}; \tag{7}$$

$$= 2^{\beta_1 + \beta_2} \tag{8}$$

So $\beta_1 + \beta_2 = 1$ when injuries are linear in observation sizes. Note that this result is independent of the supposed or true mechanism, and independent of the relationship between the covariates M and C. Therefore, as $\beta_1 + \beta_2 = 1$ is consistent with a simple linear relationship between injuries the sizes of observations.

Many studies have estimated similar values for β_1 and β_2, and this is taken as evidence for safety in numbers. Note, however, that studies examined areas with different scales, and the effect does not diminish as size increases (Elvik and Bjørnskau, 2017). Therefore, the effect identified is not a function of proximity. We should expect a phenomenon such as safety in numbers to exhibit a proximal effect: protection should be conferred most to those closest in space and time. Instead, the trend of coefficients towards $\beta_1 + \beta_2 = 1$ across all scales is consistent with linearity in injuries with respect to the mode types, and datasets that vary by scale.

2.4 The scope of this work

We focus our attention on city-level models (i.e. large contiguous geographic areas such as cities, counties and boroughs), rather than small-scale models. We would ultimately like to join up the work presented here with studies of small-scale areas, as the results in small-scale studies are already used in city-level models (Jaller et al., 2020). We emphasise that scale must be explicitly considered and accounted for in modelling, both in inferring effects from data, and in applying learnt coefficients to new settings. We propose a general framework that includes scale and that can be used for both of these applications.

3 Simulation and motivation for an alternative model

Here we look explicitly at city-level models, that is, studies where the unit of observation is the city (rather than small localities such as junctions) in order to understand how safety scales as a function of road use. As at other scales, coefficients $\beta_1, \beta_2 < 1$ have conventionally been understood to correspond to safety in numbers and, again, coefficients learnt in such studies have been suggested as parameters to be used in predictions of injury numbers in mode-shift scenarios i.e. causal effects of mode usage.

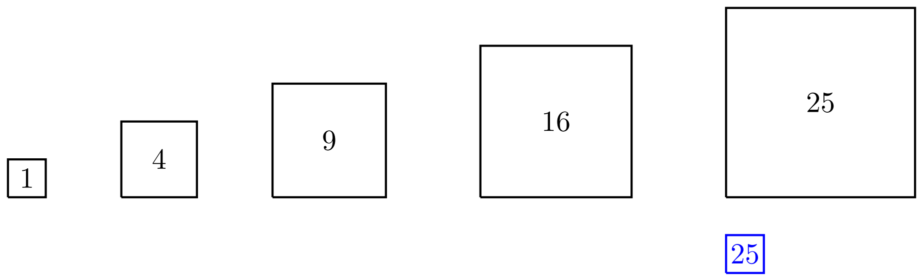

Figure 2: Schematic representation of an inter-city study, demonstrating scaling across cities. If we fit the model in Equation 2 using these cities, and then make a prediction for “City 1” with 25 times the travel, then the best estimate we make will correspond to the observation for “City 25” (black). However, when modelling mode-shift scenarios, what we actually want to predict is represented in blue: the total travel of “City 25”, taking place in a city of the size of “City 1”.

Figure 2 demonstrates the fundamental problem in using cities that vary across scales to inform models intended for use in predicting for a city that will change in number but not scale, i.e. it will change in density. In this section we develop the ideas summarised in Figure 2 through a simulation study based on a simple individual-level physical mechanism for generating collisions between different parties, and injuries.

3.1 Simulation model

We create a simulator with three variables: the number of cyclists, the number of motorists, and the dimension of the space, which we equate to the size of the city. Each cyclist and each motorist is a five-pixel by five-pixel square. At each time step each body moves one body-width. We simulate 500 time steps. Each body undertakes a biased random walk: that is, with probability 5/6 it continues in the same direction, and with probability 1/6 it chooses a direction randomly from among the directions available to it (including the same direction). Moves that would take a body out of the frame wrap around to the other side of the frame, emulating balanced migration into and out of the area. The frame size is some multiple of the step size that we vary as an input parameter.

We count the number of collisions between cyclists and motorists, which is defined as any overlap in pixels between a cyclist and a motorist. Upon collision the cyclist disappears and reappears randomly in the next time step. The cyclist disappears immediately, so a collision involving two motorists and one cyclist will be counted as one event. There is no “safety in numbers” in this model: a cyclist’s risk of collision is independent of the number of other cyclists in the simulation.

3.2 Simulation study

In order to emulate inter-city regression studies, we simulate 50 times frames of different sizes and constant density. We define our vector of sizes to be the integers $x = \{5, ..., 14\}$. The numbers of cyclists and motorists are Poisson-distributed random variables with mean x^2, and the dimensions of the frames are $20{\times}5x$ pixels by $20{\times}5x$ pixels. There are 500 simulations in total: ten sizes with fifty repetitions of each. The results of the simulations are summarised in Figure 3. We fit a Poisson regression model to the resulting number of collisions to learn the parameters β_1 and β_2, finding $\beta_1 \approx \beta_2 \approx 0.5$ as expected, since the density remains constant, but the area sizes vary. We then use (a) the model to make new predictions and (b) the simulator to test them.

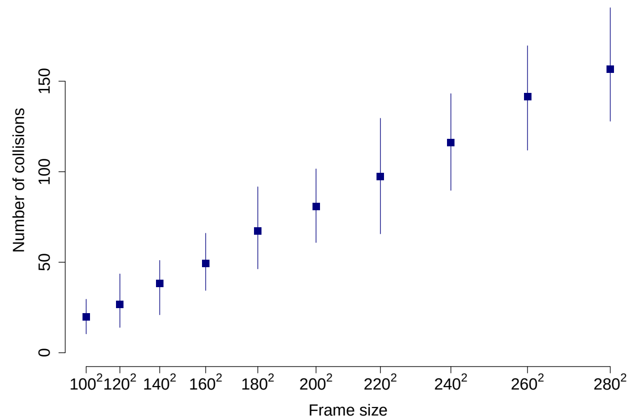

Figure 3: Simulation study. The number of collisions is a function of frame size, number of cyclists and number of motorists. Each bar shows 90% of the range of 50 simulations. In every simulation the density of cyclists and the density of motorists is the same.

3.3 Comparison of model predictions to simulations

First, we consider each frame size with 100 motorists and 100 cyclists. This was one of our simulation inputs for generating the simulated data: frame size 200^2. Because our model is a function of road-user numbers alone, we make the same prediction for every frame size: 80 collisions. In Figure 4 we plot this against the simulated number over 50 repetitions for each frame size, demonstrating the extent of the bias resulting from having left frame size out of the regression model. Note that only for the frame size corresponding to mode numbers of 100 does the 90% confidence range overlap the predicted value.

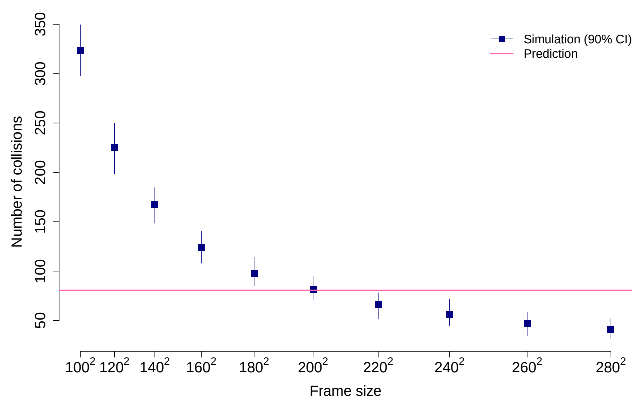

Figure 4: Simulated vs. predicted results for 100 motorists and 100 cyclists with varying frame sizes. The predicted value is 80 for all frame sizes.

Second, using the same set up, we consider that the number of motorists is constant at 100 in the corresponding frame size of 200^2, and we vary the number of cyclists. Again, we predict the number of collisions using our model, which this time varies with cyclist number. Note in Figure 5 that the prediction aligns with the simulation when the cyclist value corresponds to the frame size and density of the predictive model. When this number is exceeded, the estimate is too low, and at a lower density, the prediction is too high. We should keep this picture in mind when we consider hypothetical mode-shift scenarios such as cyclist numbers changing but the other mode, such as truck, not changing.

Third, we consider that the number of cyclists is constant at 150 in a frame size of 200^2, and

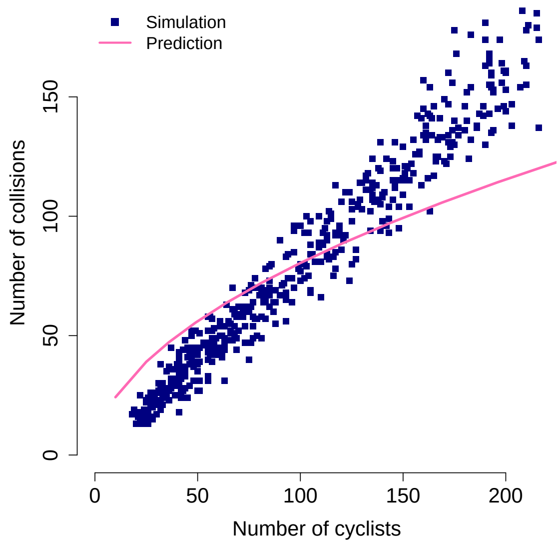

Figure 5: We fix motorists at 100 and frame size at 200^2, and vary the number of cyclists as Poisson random variables with means from 25 to 196. We simulate and predict using the model the number of collisions. This represents the type of error we might make if we predict for one mode that varies when another stays constant.

we vary the number of motorists. Again, we predict the number of collisions using our model, which this time varies with motorist number. Note in Figure 6 that the prediction aligns with the simulation when the motorist value corresponds to the difference between the motorist value corresponding to the frame size and density of the predictive model, and the change in cyclist number from its corresponding value. (Precisely: the errors cancel out when the number of new motorists = the number of old motorists $\times$ the number of old cyclists / the number of new cyclists $= 100 \times 100/150 = 66.7$.) When this number is exceeded, the estimate is too low, and at a lower density, the prediction is too high. We should keep this picture in mind when we consider hypothetical scenarios where, for example, we have six modes each with 100 road users, and we predict for a scenario in which cyclists increase to 150 and all other modes decrease to 90.

3.4 Simulated safety in numbers

We repeat the simulation study, this time imposing a safety-in-numbers effect. We specify that the probability of collision given an overlap in pixels is represented by some number $p = p(C)$, which is a function of the number of cyclists C. We recreate Figure 3 with $p = C^{-0.25}$. In terms of "safety in numbers", this corresponds to a raw exponent of 0.75. The probability p doesn't depend on M, so its exponent is 1.

We see the effect of cyclist number on collision number in Figure 7. Note the non-linear gradient, compared to Figure 3. For these data, we infer for a model of the form of Equation 2 exponents $\beta_1 + \beta_2 \approx 0.75$, which is equal to the sum of the raw exponents input into the simulation (defined above) minus 1: $\beta_1 + \beta_2 \approx 0.75 = 1 + 0.75 - 1$.

Next, we simulate the number of collisions for this model with doubled numbers of cyclists and motorists and varying frame sizes. The results are shown in Figure 8 along with the prediction from the exponents inferred before ($\beta_1 + \beta_2 \approx 0.75$) from Figure 7.

Concatenating simulated data in Figures 7 and 8, we fit a model with exponents $\beta_1 + \beta_2 \approx 1$, recovering the tiling parameters discussed previously, showing that the model is fitting across scales – hence we refer to them as scaling exponents. When we control for frame size (i.e. we specify a different model (Equation 10), which eliminates confounding), we isolate the density effect: we will refer to these exponents (those belonging to the density model, which we formally define later in Section 4) as δ_1 and δ_2, analogous to β_1 and β_2. We find that $\delta_1 + \delta_2 \approx 1.78$, close to the original

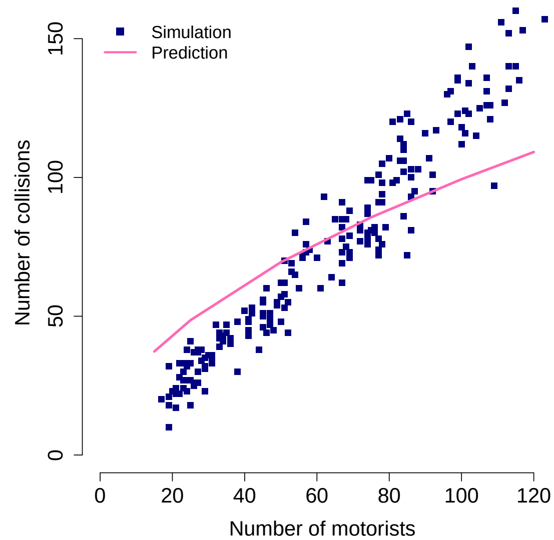

Figure 6: We fix cyclists at 150 and frame size at 200^2, and vary the number of motorists as Poisson random variables with means from 25 to 100. We simulate and predict using the model the number of collisions. This represents the type of error we might make if we predict for one mode that varies as a result of reallocation of other mode types. Note the deviation between model and prediction, which is minimised close to the point of constant density, i.e. that the increase in cyclists is close to the decrease in motorists.

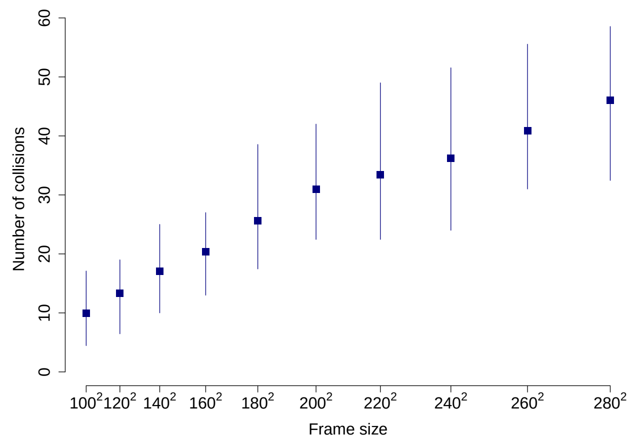

Figure 7: Simulation study. The number of collisions is a function of frame size, number of cyclists and number of motorists. Each bar shows 90% of the range of 50 simulations. In every simulation the density of cyclists and the density of motorists is the same. Collisions are a function of the number of cyclists and occur with probability $p = C^{-0.25}$.

exponents.

We repeat the whole process for $p = C^{-0.5}$, i.e. exponents 0.5 and 1. Analogously for Figure 7 we find $\beta_1 + \beta_2 \approx 0.5$, the sum of the original exponents minus one. Again we simulate data for doubled cyclists and motorists which, concatenated to the original data, yield scaling exponents $\beta_1 + \beta_2 \approx 1$ and density exponents $\delta_1 + \delta_2 \approx 1.5$ when size is accounted for.

We are trying to build up a picture of how Equation 2 behaves across scales and densities. It seems from Figures 4, 5, 6 and 8 that the predictive equation we are aiming for requires density exponents; that studies across scales return scaling exponents; and that, for two modes, the sum of the density exponents is one more than the sum of the scaling exponents: $\delta_1 + \delta_2 \approx \beta_1 + \beta_2 + 1$. We

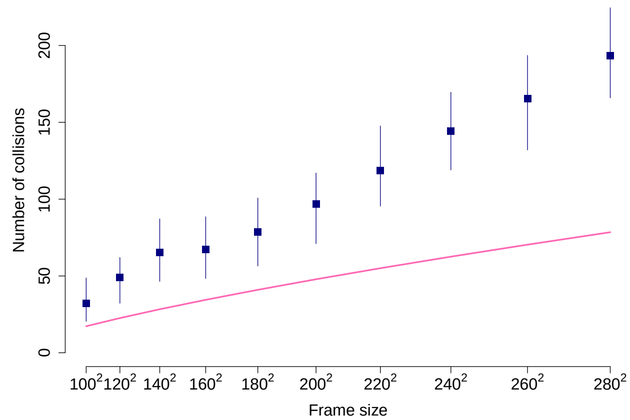

Figure 8: As in Figure 7, with double the number of motorists and cyclists. The prediction generated from data in Figure 7 is shown in pink.

will explore in the following Section if this observation generalises, and what its mechanistic basis might be.

3.5 Discussion

We have demonstrated through simulation the construction of a dataset in which there is no safety in numbers; how application of regression of the form of Equation 2 would lead us to conclude that there is safety in numbers; and that predictions resting on that model's assumptions are systematically biased. That is, we have shown how such studies might be confounded by scale. This phenomenon could potentially explain the "square root as the central estimate of 'safety-in-numbers"' (Jaller et al., 2020).

There are some key differences between our simulation study and real data: we have no spatial or temporal factors governing the rate of collision, and no relationship between e.g. density and speed. One respect in which our simulation differs from the England study of Section 5 is that we consider here constant density of road users across scales, whereas, on average, the areas in England decline in density as size increases. In contrast, in general, it is posited that, globally, as city size increases, so does density (see schematic of the model space, Figure 9). However, we could tailor our simulation to match in some way a phenomenon we are interested to capture.[3]

The point here is not how much our simulation set up resembles what one imagines happens in cities. The point is that we can contrive data using a transparent physical mechanism from which we can learn coefficients that fit the general trend $\beta_1 \approx \beta_2 \approx 0.5$ using Equation 2, and we can test this model against data simulated from the same original source. This means we can test the range of applicability of the model. We simulated data that exhibits no safety in numbers, and we inferred a safety-in-numbers effect using Equation 2. To correct for this bias, we used a size-adjusted model that accounts explicity for scale, and which we will expand upon in the next Section.

3.6 Conclusion

The conclusion of this test is that a predictive model that does not take into account factors of scale fails to predict outside its training space. We can identify the direction of the bias: uncaptured increases in density lead to underprediction of collisions, and uncaptured decreases in density lead to overprediction. This calls into question the assumption that density-independent regression models can be used to predict the number of collisions, or injuries, that will occur in mode-shift scenarios.

[3]For example, we could introduce density-dependent speeds. Concretely, in a fixed space, doubling cyclists and doubling motorists increases density by a factor of four. If we then divide every mode speed by that value - four - we undo exactly the effect of the increased density.

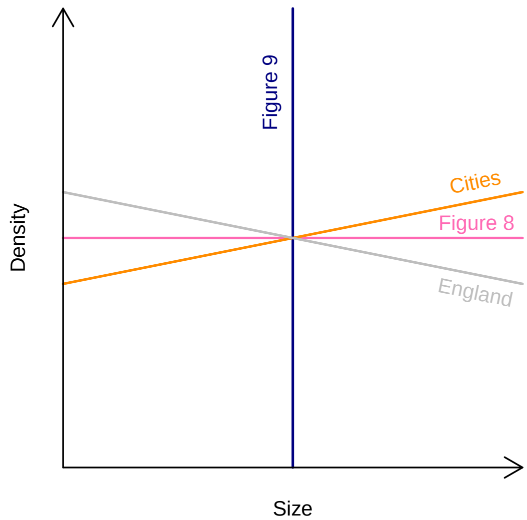

Figure 9: A depiction of the model space covered by the simulation model. We mark out in pink and navy blue a line of constant density and a line of constant size through the model space, depicted in Figures 3 and 4, respectively. The study of areas in England might be represented in this model as occupying the space of increasing size and decreasing density (grey), and expect that a study of many cities will occupy the space of increasing size and increasing density (orange). (NB: these are schematic.)

A heuristic for correction is offered (formalised in Section 4): for our simulation, scaling exponents of 0.5 corresponded to density exponents of 1, with two modes simulated. A starting point for application of inter-city studies is then the translation of a pair of scaling exponents β_1, β_2 to a pair of density exponents $\beta_1' + \beta_2' = \beta_1 + \beta_2 + 1$. This would correct the errors seen in Figures 4, 5 and 6. Then our null hypotheses, assuming road-user risk per individual road user is independent of all other road users of that type, are, for models of the form of Equation 2, that $\beta_1 + \beta_2 = 1$ for constant density and varying scale, and $\beta_1 + \beta_2 = 2$ for constant scale and varying density.

4 An alternative model

Recall that our null hypotheses are, for models of the form of Equation 2, that

(1) $\beta_1 + \beta_2 = 1$ for models fitted to datasets in which there is constant density and varying scale, and

(2) $\beta_1 + \beta_2 = 2$ for models fitted to datasets in which there is constant scale and varying density.

Testing these hypotheses would allow us to learn about the effect of road-user number on injury risks. However, as it's unlikely that we will find datasets that have either constant size or constant density, we create a model that brings both together, so that scale/density is accounted for in the model. The model we write to capture both those special cases is parametrised by exponents δ_1 and δ_2 and also by size; i.e. (1) and (2) are "special cases" and complementary presentations of the overarching model. Here, we present the model used in place of Equation 2 to correctly capture the dynamics in the simulations for general use.

4.1 The model

We assume that there are constant numbers of road users per unit in order to simplify the model and its derivation, i.e., we consider that there are n units, with $c = C/n$ cyclists per unit and $m = M/n$ motorists per unit. A more precise model would accommodate heterogeneity in road-use density. Then our model is written:

$$I_n \sim \mathcal{F}(\gamma_n) \tag{9}$$
$$\gamma_n = \frac{\alpha}{n}(nm)^{\delta_1}(nc)^{\delta_2} = \frac{\alpha}{n}M^{\delta_1}C^{\delta_2} \tag{10}$$

where n is the number of spatial units and γ_n is the expected number of injuries in those n units. Then, in a scenario, we can distinguish between a doubling of density,

$$\gamma_n^{(\text{double density})} = \frac{\alpha}{n}(2M)^{\delta_1}(2C)^{\delta_2} \tag{11}$$
$$= 2^{\delta_1+\delta_2}\frac{\alpha}{n}M^{\delta_1}C^{\delta_2} \tag{12}$$
$$= 2^{\delta_1+\delta_2}\gamma_n, \tag{13}$$

and a doubling of size, i.e. we double the number of areas n, which has the consequence of doubling M and C:

$$\gamma_n^{(\text{double size})} = \frac{\alpha}{2n}(2M)^{\delta_1}(2C)^{\delta_2} \tag{14}$$
$$= 2^{\delta_1+\delta_2-1}\frac{\alpha}{n}M^{\delta_1}C^{\delta_2} \tag{15}$$
$$= 2^{\delta_1+\delta_2-1}\gamma_n. \tag{16}$$

When density doubles, because M and C double and the size remains the same, we expect injury number to multiply by $2^{\delta_1+\delta_2}$. When density stays the same, so that when M and C double, so too does the size, we expect the injury number to multiply by $2^{\delta_1+\delta_2-1}$.

We choose this particular formulation as the simplest means of accounting for scale. It is directly comparable with previous work, such as Aldred et al. (2017) and, additionally, we can derive the relationship with β_1 and β_2 in Equation 2. Assuming observations for a size of 1 and a size of n, from Equation 10 we have

$$\frac{\gamma_n}{\gamma_1} = \frac{\frac{\alpha}{n}(nm)^{\delta_1}(nc)^{\delta_2}}{\alpha m^{\delta_1}c^{\delta_2}} \tag{17}$$
$$= n^{\delta_1+\delta_2-1}. \tag{18}$$

We can derive the equivalent relation using Equation 2:

$$\frac{\gamma_n}{\gamma_1} = \frac{\alpha(nm)^{\beta_1}(nc)^{\beta_2}}{\alpha m^{\beta_1} c^{\beta_2}} \tag{19}$$

$$= n^{\beta_1+\beta_2}. \tag{20}$$

Thus we see that for these two models to be consistent, we require $\beta_1 + \beta_2 + 1 = \delta_1 + \delta_2$.

4.2 Discussion

We propose the model of Equation 10 as a model that is consistent with our observations and our insights. There will be other models that also fulfill those criteria. There will be other models that contain ours within them as a subset or a special case. An example would be an extension that takes account also of speed.

In terms of application, we have identified a relationship between $\beta_1 + \beta_2$ and $\delta_1 + \delta_2$, but we don't know how to transform β_1 and β_2 separately. We could define $\delta_i = \beta_i + 0.5$, adding an equal amount to each variable. This assumes that a cross-scale application of the model in Equation 2 is accurately capturing the contributions of the individual modes to risk. However, inference of β_1 and β_2 is likely confounded by space and/or time, so we might instead choose $\delta_1 = \delta_2 = (\beta_1 + \beta_2 + 1)/2$, an equitable solution, which assumes that inference using Equation 2 has identified the correct sum $(\beta_1 + \beta_2)$ but not their individual contributions. These options should be tested, particularly for cases where one mode vastly outnumbers the other. For models expressing uncertainty, distributions can be assigned to these parameters and a sensitivity analysis conducted to ascertain the impact of the values on the outcome.

5 A size-adjusted study of the England data

We apply the previously presented model (Equation 10) to the England data, using the distance travelled as covariates and the total road length for each area as offsets, assuming a Poisson distribution for the counts. We use the 148 areas of England and consider only A, B and minor roads. We consider urban and rural areas both separately and together, defining urban areas as those with at least 98% of their population registered as living in a city, town or minor conurbation in the 2011 census (2079 data points; 2889 for rural). We use the software Stan in R with default (uniform) priors to test the hypothesis $\delta_1 + \delta_2 = 2$ by evaluating the posterior probability that $\delta_1 + \delta_2 < 2$ and inspecting its distribution, e.g. whether it is concentrated around a value that is practically different from 2.

5.1 Results

In Table 2 we present the 95% credible intervals for the sum $\delta_1 + \delta_2$. Our data and model give good evidence against the hypothesis of linearity ($\delta_1 + \delta_2 = 2$) for all injuries and KSI in urban areas, but not in rural areas. The evidence against this hypothesis for fatalities in all areas is weak. The ranges are much larger for fatalities, presumably because the signal is weaker as there are fewer events. It might be that with more data we would be able to make stronger statements with more confidence about the data in relation to the null hypothesis.

We note that the sums are about 1 more than the sums $\beta_1 + \beta_2$ that we learn with the same data and a simpler model that considers numbers alone and not road length (Table 1 for "All areas": 0.67 for all injuries, 0.79 for KSI, and 0.97 for fatalities). (In fact, the difference is slightly greater than 1, as the areas in England have a slightly negative correlation between size and cumulative cycling, and hence greater adjustment is required, depicted as the gap between the grey and navy lines in the schematic in Figure 9, which exceeds 90°.)

Table 2: 95% credible intervals for $\delta_1 + \delta_2$ and probability $\delta_1 + \delta_2 < 2$.

	All injuries		KSI		Fatalities	
	95% CI	Probability	95% CI	Probability	95% CI	Probability
All areas	1.79–1.80	1.00	1.90–1.93	1.00	2.00–2.18	0.02
Urban areas	1.62–1.65	1.00	1.74–1.83	1.00	1.42–2.08	0.93
Rural areas	1.92–1.93	1.00	1.96–2.01	0.92	1.90–2.14	0.37

In Tables 3 and 4 we present the 95% credible intervals for the coefficients for cycle and car travel, and in Table 5 the intervals for the intercepts. Note the negative correlations between coefficients: for "All areas" in Tables 3 and 4, as casualty severity increases, more of the coefficients' sum is attributed to car at the expense of cycle. Similarly, across all urbanicity levels, there is a negative correlation between the sums in Table 2 and Table 5, suggesting a tradeoff between base rate and safety effect.

Table 3: 95% credible intervals for δ_2 (cycle travel) and probability $\delta_2 < 1$.

	All injuries		KSI		Fatalities	
	95% CI	Probability	95% CI	Probability	95% CI	Probability
All areas	0.59–0.61	1.00	0.52–0.57	1.00	-0.01–0.30	1.00
Urban areas	0.68–0.71	1.00	0.64–0.72	1.00	-0.00–0.49	1.00
Rural areas	0.19–0.22	1.00	0.18–0.26	1.00	-0.16–0.28	1.00

Finally, we present the coefficients we estimate when we combine all datasets, including the new levels as factor predictors in a multivariable regression model (Table 6). Using factors for casualty severity and urbanisation, we end up with coefficients that resemble those fit to the dataset with all injuries and pooled urbanisation.

Table 4: 95% credible intervals for δ_1 (car travel) and probability $\delta_1 < 1$.

	All injuries		KSI		Fatalities	
	95% CI	Probability	95% CI	Probability	95% CI	Probability
All areas	1.18–1.20	0.00	1.35–1.39	0.00	1.79–2.10	0.00
Urban areas	0.93–0.95	1.00	1.06–1.14	0.00	1.20–1.80	0.00
Rural areas	1.70–1.73	0.00	1.72–1.81	0.00	1.72–2.23	0.00

Table 5: 95% credible intervals for α (the intercept).

	All injuries	KSI	Fatalities
All areas	-22.97–-22.76	-27.87–-27.27	-38.24–-34.59
Urban areas	-19.65–-19.06	-25.29–-23.55	-35.57–-23.69
Rural areas	-27.52–-27.16	-30.91–-29.94	-37.93–-32.52

5.2 Discussion

We present the test of $\delta_1 + \delta_2 = 2$ as the test for "size-adjusted safety in numbers", which can be done in the Bayesian framework with Stan. Using glm in R, we can test $\delta_1 = 1$ and $\delta_2 = 1$ (i.e., test per capita as in Shalizi (2011)) using the offset function, but we cannot test their sum.

Both in the glm framework and with Stan, it is unclear to what extent we can interpret the values for δ_1 and δ_2. Were we to use data in which there is no correlation at all between cycling volume and car volume, they will be uniquely identifiable parameters. On the other hand, if in our data cycling and car volumes were perfectly correlated, we could not separate out separate values at all. In reality, cycling and car volume are somewhat correlated, and therefore the values for δ_1 and δ_2 are somewhat entangled. Therefore, the values presented in the tables are unlikely to be representative of causal effects. It is possible that there is more safety in numbers for cyclists in rural areas, and a corresponding danger in numbers from cars in rural areas. It is also possible, given our model and data, that there is sharing of the coefficients when the model is fit, so that both values should be closer to 0.95.

There appear to be correlations between coefficients across the models. What does it mean that, for "All areas" in Tables 3 and 4, as casualty severity increases, more of the coefficients' sum is attributed to car at the expense of cycle? It seems likely that there is a pattern underlying the data, which is consistent across severity levels, and that the model is unable to describe. It seems less likely that casualty severity impacts on the dynamics of the non-linear relationship between distance travelled and road-traffic collision rates. There is also a negative correlation between base rate and safety effect in Tables 5 and 2: do places that are more safe have less of a safety effect? Or is there some transference between the parameters?

This analysis is comparable to the population-adjusted KSI model of Aldred et al. (2017). In that study, there are three covariates: cycle commuters, motor vehicle volume, and population. Here, we have two covariates, cycle distance and motor distance, and one offset, the total length of A, B and minor roads.

The number of cycling commuters is similar to estimated cycle distance; motor vehicle volume and motor distance are measuring the same thing; and population and road length are correlated, and are both a proxy for an area's size. The key difference between the two models is the treatment of the city proxy as a covariate (population) and as an offset (road length), which is equivalent to a covariate with a fixed coefficient, which we set to -1.

The sums of the means for the covariate coefficients of the population-adjusted models of Aldred et al. (2017) are 0.98, 1.06, 1.02, and 0.99. In our framework, these values would correspond to $\delta_1 + \delta_2 - 1$. We cannot test a hypothesis in this regression framework but it's worth noting the proximity to our corresponding estimates, and the proximity to $\delta_1 + \delta_2 - 1 = 1$, which implies linearity of injury with respect to scale.

Table 6: 95% credible intervals the δ coefficients using a factor covariate for severity and probabilities of the values being less than 1 and their sum being less than 2.

	δ_2 (cycle travel)		δ_1 (car travel)		$\delta_1 + \delta_2$	
	95% CI	$P(\delta_2 < 1)$	95% CI	$P(\delta_1 < 1)$	95% CI	$P(\delta_1 + \delta_2 < 2)$
All areas	0.59–0.61	1.00	1.18–1.20	0.00	1.79–1.80	1.00
Urban areas	0.68–0.71	1.00	0.93–0.95	1.00	1.62–1.65	1.00
Rural areas	0.19–0.23	1.00	1.70–1.73	0.00	1.92–1.93	1.00
Area as factor	0.58–0.60	1.00	1.23–1.25	0.00	1.83–1.84	1.00

That there is a systematic difference between "urban" and "rural" areas in terms of the coefficients fit by our model is indicative that the model does not capture the whole effect of travel density on injury rates. This, and the spurious correlations in coefficients, show there are features of the data not explained by the model of Equation 10.[4]

Mis-specification of any of (a) the relationship in the model between size and rate, (b) the probabilistic description of the error term, (c) the component contributions of the two mode distances, and (d) quantification of size through road length might contribute to the failure to explain the data, and all would benefit from further consideration and testing. All of these things can be improved upon, incrementally, as we have taken an incremental step from Equation 2 to a size-based model in Equation 10. However, taken together, they highlight that it has not been shown that a model with the fundamental form of Equation 2 might be capable of answering the question of safety linearity.

[4]The difference does, however, open the possibility that there is "nonlinearity in the nonlinearity" – i.e., that the effect is greater (and the exponents smaller) at higher densities. This is an avenue that could be explored as a link between city-scale and small-scale studies, which typically consider only high-density areas of a city.

6 Conclusions

For studies that scale across sizes, $\beta_1+\beta_2=1$ for Equation 2 represents linearity in numbers. That the same (or similar) coefficients are observed across scales is perhaps the biggest indication that whatever is being captured is not an effect of cyclists conferring protection to other nearby cyclists.

The missing component in Equation 2 is size. Omission of size results in misleading interpretations and predictive models unsuitable for fixed settings. Therefore, we recommend departing from this model and developing new expressions, which include size explicitly. We propose, in the first instance, a very simple adjustment to Equation 2 in Equation 10, in which we include size as defined by total road length.

We recognise the challenge of specifying testable hypotheses in this setting. We are trying to formulate the hypothesis that the risk to an individual road user colliding with another road user is independent of the number of other road users of their type, and linear in the number of road users of the other type, within a particular space. Then, for each cyclist, the risk would be αm. For c cyclists, the expected number of injuries would be αmc. For n spatial units, the expected number of injuries would be $n\alpha mc$.

Ideally we would test per capita rates for δ_1 and δ_2 in Equation 10, but, as the model stands, they cannot be confidently identified, so we test instead the null hypothesis $\delta_1+\delta_2=2$ in Stan. With this model, hypothesis, and the data for England, we found some subsets of the dataset supported a size-adjusted safety-in-numbers effect, and some did not.

The model we present is best described as "size-adjusted safety in numbers"; we have not developed or tested a model of "safety in density", which might be interesting and relevant to explore. It would allow for heterogeneity in density over space and time and exploration of their impacts on results. That will be particularly important for small-scale studies, e.g. of junctions, and could be addressed in the first instance through simulations such as those presented in Section 3. In addition to heterogeneity in density over time and space, an improvement to the model in Equation 10 would be inclusion of mode speeds.

In terms of hypothesis testing, we aim to engage a wider audience and share data in order to find alternative ways to understand and formulate the problem. The forms we have considered in this work bear much similarity to the Cobb-Douglas production function, which relates capital and labour supply to production using output elasticities (Cobb and Douglas, 1928). We recognise also a correspondence between our model and city-level metrics that have been claimed to exhibit power-law scaling properties (Leitão et al., 2016). Input from these models might greatly benefit progress in this topic. In addition, there might be parallels with the practice of discretisation of space (and time) in a "contact matrix" to describe interactions across partitions in infectious-disease modelling (Birrell et al., 2011). We identify some areas for further testing or development of the model described by Equation 10, such as specification of the error term, use of density proper rather than size adjustment, the expression of the mode distances, how to quantify size, and the problem of mode (dis)aggregation. We would welcome development of other models, as well as methods for assessment and hypothesis testing.

Complementing data-driven analyses, simulations can be used to develop relational models. These can test implications of comprehensive mechanistic models of injuries as a function of space and its occupancy. Simulation and theory provide us with a number of null hypotheses, which (a) give us an objective to test and reject, and (b) provide a justified basis for prediction. In this work, the theory and simulation led to the following hypotheses, which connect Equation 2 to Equation 10:

(1) City-level size-scaling exponents $\beta_1+\beta_2=1$ correspond to linearity (Section 2.3).

(2) City-level size-scaling exponents $\beta_1+\beta_2=1$ (Equation 2) correspond to city-level density-scaling exponents $\delta_1+\delta_2=2$ (Equation 10), where mode speeds are assumed independent of density.

(3) To predict the consequence of a change in density using city-level size-scaling exponent β, we can use density exponents $\delta_1+\delta_2=1+\beta_1+\beta_2$, where there are two modes involved (Section 3.4).

These hypotheses were consistent with the results of the England data, in that the coefficients δ_1

and δ_2 fit with the model of Equation 10 had sum approximately one greater than that of β_1 and β_2 fit with the model of Equation 2. We therefore propose this framework as a first amendment for making predictions of road-injury burden in mode-shift scenarios. With reference to those initial objectives, we hope to have offered a new perspective, and to have generated more questions and pointed to new lines of inquiry.

A Supplementary figures

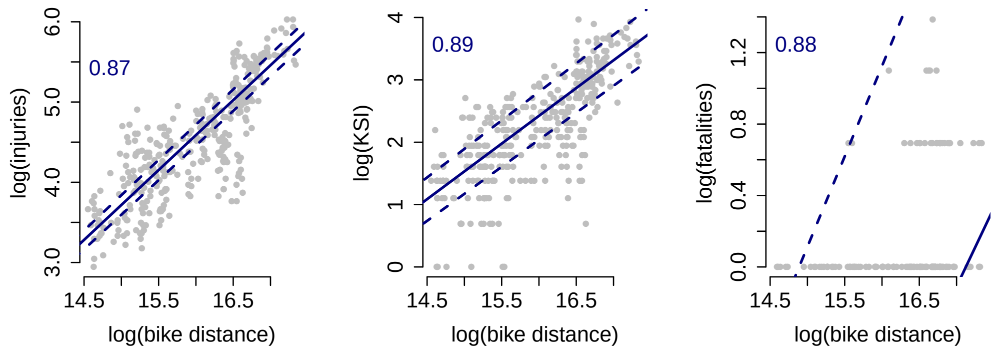

Figure 10: As in Figure 1, with only the London boroughs.

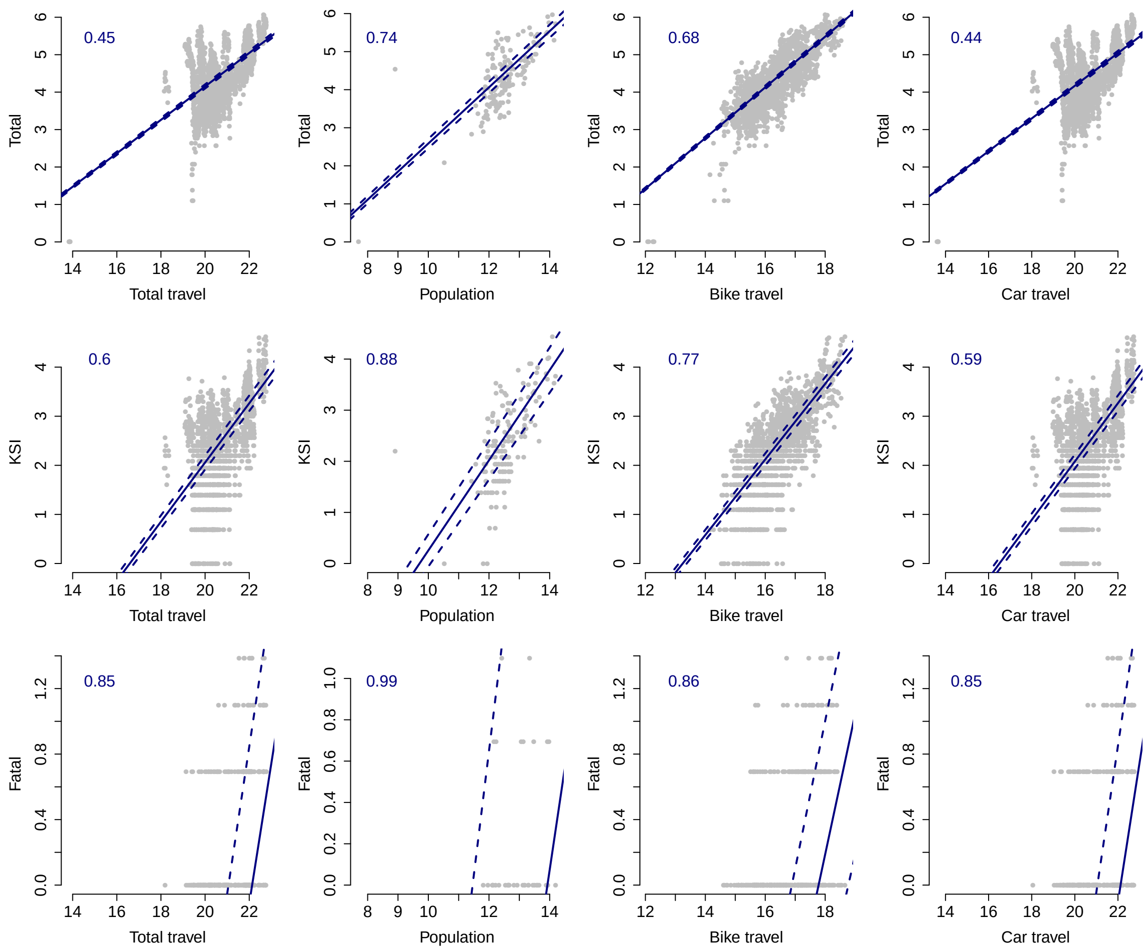

Figure 11: As in Figure 1, showing different relationships and the β coefficients they generate.

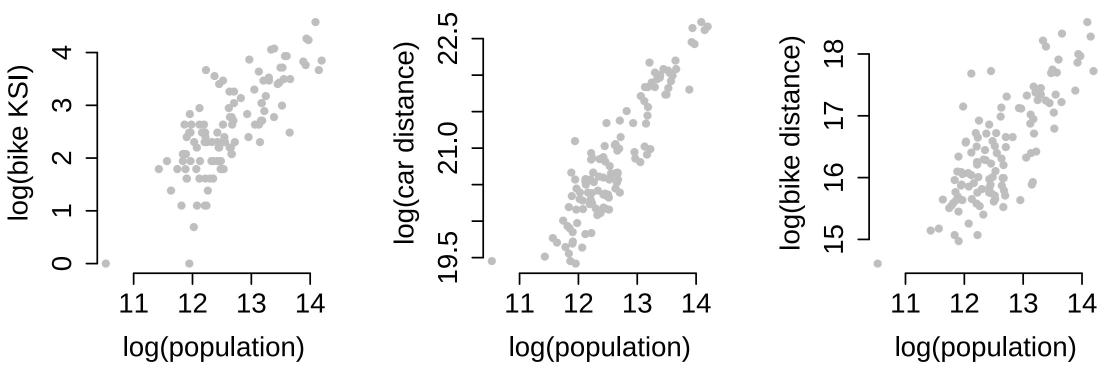

Figure 12: Linear relationships between population (N) and cyclist KSI, between population and car distance, and between population and cyclist distance in English counties.

B Worked example: implementation in ITHIM-R

We consider the setting of Accra, for which we have a list of recorded fatalities over multiple years. Each record contains the following information: the year, the mode of the casualty, the mode of the other party, the age of the casualty, and the gender of the casualty. In addition, we have a travel survey, from which we learn total travel by each mode (and by demographic group, which we omit for now, for simplicity).

B.1 Constructing the model

We fit the observed data (the number of injuries, I) to an equation of the form

$$I \sim \text{Poisson}(\lambda), \tag{21}$$

$$\lambda = \alpha M^{\beta_1} C^{\beta_2} \exp\left(\sum_{i=3}^{P} X_i \beta_i\right) \tag{22}$$

with α a fixed intercept, C and M the distances travelled by cyclists and cars, respectively, based on the travel survey, and X the model matrix built from all the covariates (here, we consider only the two modes; gender and age of the casualty are omitted for simplicity). We do not use the "year" covariate but instead suppose that we have multiple observations for a single "year" (i.e. we reuse the distance data). Finally, the coefficients to fit using `glm` are α and β_i for $i \geq 3$, and we supply β_1 and β_2 as fixed parameters so that $M^{\beta_1} C^{\beta_2}$ is our offset.

Note that there are many combinations of modes, so this model is linked via the model matrix X to the number of pedestrian casualties in collisions with buses, etc. The contingency table of injury counts between all mode pairings forms the "who hit whom" matrix for the city.

B.2 Making predictions

We use the same model equation to make predictions in hypothesised scenarios. The prediction equation requires us to specify the distances travelled in the scenario: call them $\hat{M}$ and $\hat{C}$. Then we predict the expected number of injuries in the scenario, $\hat{I}$, as:

$$\hat{I} = \alpha \hat{M}^{\beta_1} \hat{C}^{\beta_2} \exp\left(\sum_{i=3}^{P} X_i \beta_i\right). \tag{23}$$

To aid interpretation, we can consider the ratio of the expected injuries in the scenario to the expected injuries in the baseline:

$$\frac{\hat{I}}{\mathbb{E}(I)} = \frac{\hat{M}^{\beta_1} \hat{C}^{\beta_2}}{M^{\beta_1} C^{\beta_2}} \tag{24}$$

$$= \left(\frac{\hat{M}}{M}\right)^{\beta_1} \left(\frac{\hat{C}}{C}\right)^{\beta_2}. \tag{25}$$

Then we can immediately read out, for example, that if M does not change ($\hat{M} = M$) then the fold change in injuries is equal to the fold change in cycling raised to the power β_2: if $\beta_2 = 1$, then if cycling increases 25 times, so does the injury count. If $\beta_2 = 0.5$, then if cycling increases 25 times, the injury count increases five times.

B.3 β_1 and β_2 parameters

The question we need to answer is, given that we are using this model, what values should we choose for β_1 and β_2? This choice will impact on the other parameters to fit (α and β_i for $i \geq 3$) and, crucially, on the number of injuries we predict in scenarios.

Recall that there are multiple casualty modes and multiple "other party" modes, including NOV (no other vehicle). Another question we need to answer is how these values should differ for different modes, in particular (a) where a mode's distance is not changing at all (or even very little) in scenarios, (b) where a mode is a combination of multiple modes, and (c) where there is no other mode.

References

Aldred, R., Goel, R., Woodcock, J. and Goodman, A. (2017), 'Contextualising Safety in Numbers: a longitudinal investigation into change in cycling safety in Britain, 1991–2001 and 2001–2011', *Injury Prevention* pp. injuryprev–2017–042498.
URL: *http://injuryprevention.bmj.com/lookup/doi/10.1136/injuryprev-2017-042498*

Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C. and West, G. B. (2007), 'Growth, innovation, scaling, and the pace of life in cities', *PNAS* **104**(17), 7301–7306.

Birrell, P. J., Ketsetzis, G., Gay, N. J., Cooper, B. S., Presanis, A. M., Harris, R. J., Charlett, A., Zhang, X.-S., White, P. J., Pebody, R. G. and De Angelis, D. (2011), 'Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London', *Proceedings of the National Academy of Sciences* **108**(45), 18238–18243.
URL: *http://www.pnas.org/cgi/doi/10.1073/pnas.1103002108*

Clauset, A., Shalizi, C. R. and Newman, M. E. (2009), 'Power-law distributions in empirical data', *SIAM Review* **51**(4), 661–703.

Cobb, C. W. and Douglas, P. H. (1928), 'Theory of production', *American Economic Association* **2**, 139–165.

de Sá, T. H., Tainio, M., Goodman, A., Edwards, P., Haines, A., Gouveia, N., Monteiro, C. and Woodcock, J. (2017), 'Health impact modelling of different travel patterns on physical activity, air pollution and road injuries for São Paulo, Brazil', *Environment International* **108**(June), 22–31.
URL: *http://dx.doi.org/10.1016/j.envint.2017.07.009*

Elvik, R. and Bjørnskau, T. (2017), 'Safety-in-numbers : A systematic review and meta-analysis of evidence', *Safety Science* **92**(0349), 274–282.
URL: *http://dx.doi.org/10.1016/j.ssci.2015.07.017*

Geyer, J., Raford, N., Pham, T. and Ragland, D. (2006), 'Safety in numbers: data from Oakland, California', *Transportation Research Record: Journal of the Transportation Research Board* **1982**, 150–154.
URL: *http://dx.doi.org/10.3141/1982-20*

Gårder, P., Leden, L. and Pulkkinen, U. (1998), 'Measuring the safety effect of raised bicycle crossings using a new research methodology', *Transportation Research Record* **1636**(98), 64–70.
URL: *http://nacto.org/wp-content/uploads/2010/08/Measuring-the-Safety-Effect-of-Raised-Bicycle-Crossings-Using-a-New-Research-Methodology.pdf*

Jaller, M., Pourrahmani, E., Rodier, C., Maizlish, N., Zhang, M., University of California, D., Center for Transportation and Community Health, E., and Technology, O. o. t. A. S. f. R. (2020), 'Active Transportation and Community Health Impacts of Automated Vehicle Scenarios: An Integration of the San Francisco Bay Area Activity Based Travel Demand Model and the Integrated Transport and Health Impacts Model (ITHIM)', p. 49p.

Kleiber, M. (1947), 'Body size and metabolic rate', *Physiological Reviews* **27**(4).

Leden, L. (2002), 'Pedestrian risk decrease with pedestrian flow. A case study based on data from signalized intersections in Hamilton, Ontario', *Accident Analysis and Prevention* **34**, 457–464.

Leitão, J. C., Miotto, J. M., Gerlach, M. and Altmann, E. G. (2016), 'Is this scaling nonlinear?', *Royal Society open science* **3**, 150649.
URL: *http://dx.doi.org/10.1098/rsos.150649*

Miranda-Moreno, L., Strauss, J. and Morency, P. (2011), 'Disaggregate exposure measures and injury frequency models of cyclist safety at signalized intersections', *Transportation Research Record: Journal of the Transportation Research Board* **2236**, 74–82.
URL: *http://trrjournalonline.trb.org/doi/10.3141/2236-09*

Nordback, K., Marshall, W. E. and Janson, B. N. (2014), 'Bicyclist safety performance functions for a U.S. city', *Accident Analysis and Prevention* **65**, 114–122.
URL: *http://dx.doi.org/10.1016/j.aap.2013.12.016*

Prato, C. G., Kaplan, S., Rasmussen, T. K., Hels, T., Giacomo, C., Kaplan, S. and Rasmussen, T. K. (2016), 'Infrastructure and spatial effects on the frequency of cyclist-motorist collisions in the Copenhagen Region', *Journal of Transportation Safety & Security* **8**(4), 346–360.
URL: *http://dx.doi.org/10.1080/19439962.2015.1055414*

Schepers, J. P. and Heinen, E. (2013), 'How does a modal shift from short car trips to cycling affect road safety?', *Accident Analysis and Prevention* **50**, 1118–1127.
URL: *http://dx.doi.org/10.1016/j.aap.2012.09.004*

Schepers, J. P., Kroeze, P. A., Sweers, W. and Wüst, J. C. (2011), 'Road factors and bicycle – motor vehicle crashes at unsignalized priority intersections', *Accident Analysis and Prevention* **43**(3), 853–861.
URL: *http://dx.doi.org/10.1016/j.aap.2010.11.005*

Shalizi, C. R. (2011), 'Scaling and hierarchy in urban economies', *arXiv* p. 1102.4101v2.

Stumpf, M. P. H. and Porter, M. (2012), 'Critical truths about power laws', *Science (New York, N.Y.)* **335**(6069), 665–6.