

This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



CC creative commons
COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

BY: **Attribution.** You must attribute the work in the manner specified by the author or licensor.

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

A Spatially Disaggregate Analysis of Road Casualties in England

Robert B. Noland
Mohammed A. Quddus

Centre for Transport Studies
Dept. of Civil & Environmental Engineering
Imperial College London
London, SW7 2AZ
Ph: 44(0)20-7594-6036
Fx: 44(0)20-7594-6102
Email: r.noland@imperial.ac.uk
Email: m.quddus@imperial.ac.uk

Revised: May 28, 2003

A Spatially Disaggregate Analysis of Road Casualties in England

Abstract

Spatially disaggregate ward level data for England is used in an analysis of various area-wide factors on road casualties. Data on 8414 wards was input into a geographic information system that contained data on land use types, road characteristics and road casualties. Demographic data on area-wide deprivation (the index of multiple deprivation) for each ward was also included. Negative binomial count data models were used to analyze the associations between these factors with traffic fatalities, serious injuries and slight injuries. Results suggest that urbanized areas are associated with fewer casualties (especially fatalities) while areas of higher employment density are associated with more casualties. More deprived areas tend to have higher levels of casualties, though not of motorized casualties (except slight injuries). The effect of road characteristics are less significant but there are some positive associations with the density of "A" and "B" level roads.

Key words: transport safety, negative binomial models, spatial analyses, geographic information systems, deprivation, land use

Introduction

The United Kingdom (UK) has one of the lowest incidents of traffic-related fatalities in the world. Reductions in both total fatalities and injuries have been significant over the last 30 years, despite large increases in vehicle travel. These reductions are frequently related to improved vehicle design, seat-belt usage, detailed safety audits after crashes, and various engineering safety measures.

One factor not normally considered are the land use characteristics of an area, for example, the extent of urban development and relative population density. The relative socio-economic condition of a neighborhood is also not normally evaluated. Lower income neighborhoods may be subjected to larger amounts of traffic than wealthier neighborhoods but there may also be some additional reasons why relatively more casualties occur in these neighborhoods.

Controlling for land use and economic condition is best evaluated through a spatial disaggregation of the data. This also allows one to control for population characteristics (such as relative age cohorts) and some of the road infrastructure features, such as the number of roundabouts and junctions.

This study uses a spatial analyses approach to evaluate these issues. This is made possible by geographic UK ordnance survey data that features road lengths and other infrastructure features. This is overlaid with land use data on percent urban and rural populations, population data, and data on economic deprivation. The level of aggregation is over 8414 wards (similar to precincts) for England (i.e., we do not analyze data from Scotland, Wales, or Northern Ireland). Data on traffic casualties is geo-coded and can be assigned to the geographic units within our geographic information system (GIS). Negative binomial (NB)

models are estimated with the dependent variable being total fatalities, serious injuries and slight injuries and for these same categories but with only motorized casualties (i.e., excluding pedestrian and bicyclist casualties). This is done because there may be different relationships between motorized casualties as opposed to total casualties. However, the latter is of primary interest since total casualties is the key outcome variable of policy interest. We find some minor differences in effects for total casualties versus those for motorized casualties only.

Casualty data was derived from the STATS19 (1999) UK national road accident data and aggregated to the ward level. This allows statistical analyses of associations between various categories of road casualties and the independent variables within the data set.

Our results show that both land use and area deprivation are associated with traffic casualties. Wholly urban wards experience a lower level of fatalities compared to wholly rural wards whereas wards with higher employment density and lower population density are associated with increased fatalities. Area deprivation is positively associated with total traffic casualties, but not motorized casualties; the magnitude of the effect is higher for slight injuries than other casualties in both cases. Number of nodes and roundabouts and length of motorway per square meter of ward do not exhibit any effects on traffic casualties, albeit length of A-roads, B-roads and minor roads show some effects on serious injuries.

The paper is organized as follows. First we provide a brief literature review of other work that has evaluated both land use and economic deprivation and their relationship to traffic casualties. This is followed by a detailed description of the data sources and the geographic information system that has been constructed. Statistical methods are then described followed by a presentation of results. This is followed by conclusions and suggestions for further research in this area.

Literature Review

Much of the research on traffic casualties has evaluated the influence of human factors, road geometry, traffic characteristics, environmental factors and vehicle characteristics on casualties. There are very few studies that evaluate spatial factors such as land use and the level of economic deprivation within a given spatial unit.

Kim and Yamashita (2002) examine the relationship between land use and accidents. Their results show that in areas of commercial land use there appear to be both a higher frequency of crashes overall as well as on a per acre basis. Their study linked ten-years of crash data with one year of land use data, thereby perhaps missing the effect of changes in land use over time.

Senior et al. (2000) investigate the relationship between premature mortality and material deprivation, and the differences in this relationship between urban and rural areas. Their results show that high mortality can be found in any area of Wales although it is more pervasive in the industrial valleys and the low-income areas of cities, towns and villages. More rural areas have lower average premature death at the deprived end of the socio-economic spectrum. This shows the highly differentiated nature of rural areas which themselves harbor pockets of severe mortality and deprivation. . Inequalities in premature mortality are widest in the cities, narrowest in the more rural areas and of intermediate comparable value in the rest of Wales. This work suggests, that many factors associated with deprivation may affect mortality rates, and we would expect similar variations with how deprivation may affect traffic-related mortality.

Ossenbruggen et al. (2001) examine roadway safety in rural and small urbanized areas. The hypothesis tested in their study is that land use activity, roadside design and other site

characteristics are important factors in predicting roadway risk, which is the probability of a crash or injury. Logistic regression analysis is used to identify statistically significant factors that are associated with crash and injury risk. Comparative risk assessment is used to analyze 'typical' village, shopping and residential sites and to help explain why one site is more hazardous than another one. Owing to their relatively large traffic exposure, typical shopping sites proved to be most hazardous. When the three sites studied are adjusted to have the same traffic exposure, village sites proved to be least hazardous and residential and shopping sites proved to be most hazardous. Village sites, which are relatively more pedestrian friendly and have the lowest vehicle operating speeds, may help explain why they are least hazardous.

Chichester et al (1998) analyze the associations between road traffic accidents and socio-economic deprivation on Scotland's west coast. This study analyzed hospital accident and emergency room admissions and investigated associations between road traffic accidents and socio-economic status. Findings strongly suggest ($p=0.0046$) a positive trend between road traffic accidents and deprivation. Significance held for gender, victim role, purpose of journey and age, except for drivers 60 or over. This study demonstrated a link between individual deprivation and road traffic accidents. In our study we control for the social deprivation of geographic areas and the road traffic accidents that occur in those areas.

Abdalla al (1997) examined the relationship between area social characteristics and road accident casualties. Initially, consideration was given to the relationship between casualty frequencies and the distance of the accidents from the zones of residence. As might be anticipated, the casualty frequencies were higher near the zones of residence, possibly due to higher exposure. Investigation of the relationships between casualty rates and social deprivation indicators for the casualties' zone of residence found that the casualty rates amongst residents

from areas classified as relatively deprived were significantly higher than those from relatively affluent areas. They also found that the areas of deprivation tended to have higher accident rates, suggesting some factors associated with the area itself was related to the likelihood of an accident. They did not control for other effects, such as the level of infrastructure or the land use of the zone.

Beattie et al (2001) found that admission to accident and emergency departments is not only related to severity of injury but also to deprivation category. The paper did not just evaluate traffic-related casualties, but all injury accidents. This suggests some underlying relationship between increased risk of all accidents and economic deprivation.

Roberts et al (1992) examined the geographical distribution of traffic-related child injury morbidity in Auckland, New Zealand, between 1982 and 1987. Their results suggest that total injury morbidity and pedestrian injury morbidity rates are strongly correlated with census area unit unemployment rates, which were used as a measure of socio-economic deprivation. Their conclusion was that injury prevention programs should be targeted at socio-economically disadvantaged communities.

An analysis was made of the death rates from accidents in children aged 0-14 by health districts in England and Wales during the five years periods between 1974-79 and 1980-1984 by Avery et al (1990). The results show that death rates were higher in urban areas than in rural areas. There was a very strong correlation with social deprivation.

Another study by Jessop (1996) could not confirm the hypothesis that the relationship between deprivation and accident related mortality is weaker among residents of non-metropolitan areas of England and Wales than among residents of metropolitan areas.

These studies suggest, first, that those living in more deprived areas are more likely to be involved in traffic crashes that result in casualties and also that those who live in more deprived areas suffer increased risk of mortality from all accidents. Evidence on the relationship to specific land use factors is less clear, but generally suggests that commercial areas have higher casualties or traffic accidents. For the most part, many of these studies do not control for other factors that influence traffic-related casualties, such as population age cohorts and the type of road infrastructure. In addition, most have not used more advanced statistical techniques that allow for the properties of count data to be correctly modeled. This is a concern, in particular given the large number of zero counts in the data. For this reason we use negative binomial models with the correct distributional properties (Jovanis & Chang, 1986; Miaou & Lum, 1993).

In the next section we describe the data used in this study in some detail. This is followed by the statistical methods used and the results of the models estimated.

Description of Spatial Data

One of the key elements of this study was the integration of multiple databases into a comprehensive geographic information system. This was made possible by the geo-coding of accident locations within the 1999 STATS19 UK national road accident database. The recorded location of an accident can be identified by its x and y coordinates which correspond with the British National Grid coordinate system. STATS19 disaggregates the outcome of each accident by fatalities, serious injuries and slight injuries, as well as many other variables that identify the individual, the vehicles involved, and other factors associated with the accident.

The total accident counts, disaggregated by fatalities, serious injuries, and slight injuries were aggregated to 8414 census wards for England (this data was not available for Scotland and Wales). The average number of households in each ward is about 2000, which represents an

aggregation of smaller scale census enumeration districts. Geographic data on ward boundaries was obtained from the EDINA UK BORDERS data set. This enabled us to aggregate the counts at the ward level within a geographic information system (both ArcView and MapInfo were used in our analyses). Figure 1 displays the boundary data for the wards used in our analyses. The average size of each ward is about 1490 hectares, with the smallest being 0.09 hectares and the largest being about 44800 hectares. This unit of analyses allows a very detailed spatial analyses of the factors associated with road accidents.

To obtain data on features of the road network, data was obtained from Meridian2™ data, which is derived from the latest available versions of Ordnance Survey (OS) data. This data uses the National Grid coordinate system for describing locations within Great Britain and is therefore easily entered into our GIS. The resolution of the data is 1 meter. It is a geometrically structured vector database customized from a variety of OS datasets that define the real world geographic entities (objects) as point and line features. It has a scale of 1:50000 containing a comprehensive road network, main passenger railways, county, district and unitary authority boundaries, developed land use areas, woodlands and water features. Different infrastructure features were extracted from the data including motorway, A road, B road, minor road and nodes (i.e., intersections). Figure 2 displays a sample of this data for central London. From this we calculate the road lengths for each category of road, the number of nodes and the number of roundabouts within each census ward.

The EDINA UKBORDERS data specifies categories of land use for each ward. The associated land use data of wards has six categories, which are “wholly urban” (100%), “predominantly urban” (+75%), “mixed urban” (50-75%), “mixed rural” (25-50%), “predominantly rural” (1-25%) and “wholly rural” (0%). These are defined as categorical

variables in our dataset. Unfortunately, more specific data on type of land use, such as residential, commercial, or industrial uses was not available. Therefore, we use employment density and population density as proxies for the relative intensity of commercial/industrial and residential land uses.

Total residential population and population age cohorts under 16, 16-65 and over 65 were also available at the ward level. Percent population in each of the age cohorts is often associated with road accidents, especially for high risk younger population groups. Unfortunately, we did not have a finer disaggregation of age cohorts at the ward level.

We also use a measure of economic deprivation (poverty) as do Graham et al. (2003). The UK specifies an index of deprivation (ID2000) for each ward in the UK (DETR, 2000). This is a combined index based on various socioeconomic factors associated with each ward. Therefore it is a spatial measure based upon where the road accident occurs and may have no relationship to the socioeconomic conditions of the victim of the accident.

ID2000 is made up of six domain indices at the ward level. These are income, employment, health deprivation and disability, education skills and training, housing and geographical access to services. The income score is based upon the percent of the population living on means-tested welfare benefits. The employment score is based upon the percent of the population who want to work but are unable to do so due to unemployment, sickness and disability. The health score is based upon the fraction of people whose quality of life is impaired by either poor health or disability. The education, skill and training score measures education deprivation. This is based upon the lack of educational qualifications amongst adults and children of different ages in a local area. The housing score identifies the number of people living in unsatisfactory housing, and, in extreme cases, homelessness. This score is derived from:

(1) homeless households in temporary accommodation, (2) household overcrowding, and (3) poor quality private sector housing. The score for geographical access to services is comprised of (1) access to a Post Office, (2) access to food shops, (3) access to a general practitioner, and (4) access to primary schools for all 5-8 years olds. The focus of the first three indicators of the access score is on lower income households as they are more likely to not have access to private vehicles or may not be able to afford public transport. Therefore, this score basically measures the difficulty involved with gaining access to basic activities. A detailed description of the data behind the ID2000 is available in DETR (2000).

The index of multiple deprivation is based upon a weighted average of these individual scores. The defined weights are: income (25%); employment (25%); health deprivation and disability (15%); education, skills and training (15%); housing (10%); and, geographic access to services (10%). The IMD2000 score is the combined sum of the weighted, exponentially transformed domain rank of the domain score. The larger the IMD2000 score, the more deprived the ward.

One of the key determinants of the likelihood of a road accident is the relative exposure to traffic. Unfortunately, we do not have data on the volume of vehicle travel within each ward. Instead we use a method devised by Graham and Glaister (in press) to represent the level of travel activity based upon the levels of population and employment in each ward relative to all other wards. Therefore, we derive a proxy variable for the volume of traffic flowing through each ward. This is derived as,

$$PE_i = \sum_j \frac{E_j}{d_{ij}} \quad \text{and} \quad PP_i = \sum_j \frac{P_j}{d_{ij}} \quad \text{where } i \neq j \quad (1)$$

in which d_{ij} is the centroid distance from ward i to ward j . Originating and destination traffic within a ward is proportional to the level of total employment, E_i , and resident population, P_i , for ward i . Therefore, following Graham and Glaister (in press), we define proximate employment, PE_i , and proximate population, PP_i , of ward i as a proxy for the volume of traffic passing through the ward. Clearly this will be heavily weighted based upon population or employment of neighboring wards with more distant wards receiving less weight (Graham and Glaister, in press). This makes sense in that we would expect wards which are more proximate to generate more traffic, as would those with more population or employment.

Summary statistics for the data used in the analyses is presented in Table 1.

Statistical Methodology

One of the features of road accident data is that it is count data and therefore is not normally distributed. Therefore, either a Poisson or negative binomial model has the correct distributional properties for model estimation. Poisson distributions assume the variance is equal to the mean, a condition that is frequently violated and is known as overdispersion. Negative binomial models provide a generalization of the Poisson model that can account for this (Miaou, 1994; Shankar et al., 1995; Vogt and Bared, 1998). Although the source of overdispersion in count data cannot be distinguished, its presence can be adjusted by introducing a stochastic component in the log-linear relationship between the expected numbers of accident in an observation unit i , μ_i and the covariates \mathbf{X}

$$\ln \tilde{\mu}_i = \mathbf{X}_i \boldsymbol{\beta} + \varepsilon_i \quad (2)$$

where $\boldsymbol{\beta}$ is a vector of estimable coefficients representing the effects of the covariates. ε is a random error that is assumed to be uncorrelated with \mathbf{X} . One can think of ε either as the combined effects of unobserved variables that have been omitted from the model (Gourieroux et

al. 1984) or as a source of pure randomness (Hausman et al. 1984). In the Poisson regression model, variation in μ is introduced through observed heterogeneity. Different values of X results in different values of μ . In the negative binomial model, variation $\tilde{\mu}$ is due both to variation in X among individuals but also due to unobserved heterogeneity introduced by ε .

The probability density function for the NB distribution can be expressed as

$$\Pr(n_i | \mu_i, k) = \frac{\Gamma(n_i + 1/k)}{\Gamma(1/k)\Gamma(n_i + 1)} \left(\frac{k\mu_i}{1 + k\mu_i} \right)^{n_i} \left(\frac{1}{1 + k\mu_i} \right)^{1/k} \quad (3)$$

in which $k(\geq 0)$ is often referred to as the overdispersion parameter. If k goes to zero then the negative binomial model reduces to the Poisson regression model. In this way, the Poisson regression model is nested within the NB and statistical tests for $k=0$ can be used to evaluate the significant presence or amount of overdispersion in the data. We find that the NB model appears to be justified in the analyses that follows.

Results of Model Estimates

Our objective was to explore statistical associations between road casualties and the spatial variables associated with a given area. Therefore, we are able to test hypotheses concerning the association of road casualties with various land use variables, road characteristics, and area-wide demographics including the level of social deprivation. Tables 2 and 3 contain results for negative binomial model estimates for fatalities, serious injuries, and slight injuries for both all accidents and just those affecting motorists (i.e., excluding pedestrians and bicyclists). In addition, we examine the substantive effect of the independent variables by examining the percent changes in count frequencies associated with the dummy variables and by calculating elasticities for mean values of the continuous variables. These are shown in Tables 4 and 5.

Table 4 also shows the actual count for each of the casualty categories and the predicted count. We can see that the fatality and serious injury models perform quite well, while the slight injury models over predict the number of actual incidents. This may be a result of the better accuracy of accident data associated with more severe injuries and fatalities, and the results of the analyses should be considered with this in mind.

The variable in the models can be grouped into four distinct categories. These are the land use variables, described by both population and employment density and the fraction of urbanized land; road characteristics, which are described by the length of various road types and the number of junctions and roundabouts in a given area; demographic characteristics, which include age cohort divisions, the index of multiple deprivation, and the percent of population that is economically active; and, finally, proxies for the traffic flow through a given area, the proximate employment variable described earlier plus total employment and residential population. Each category of variables is discussed in the following sections and summary statistics are presented in Table 1.

Land use variables

Employment and population density were tested as proxy variables for land use in each model but were found to be statistically insignificant. Better results were obtained by including dummy variables based upon the distribution of observed employment density and population density (by square km). Results are interpreted in comparison to the reference case that represented the middle range of densities. The likelihood of a casualty appears to be lower when employment density is less than 1500 / km². This effect is strongest for slight injuries. When employment density is higher than 3000 / km² casualties are more likely. Both these effects are

only prevalent with regard to total casualties and do not appear to be statistically significant when considering only motorized casualties (Table 3), except for slight injuries.

Table 4 shows the expected change in count frequency for the employment measures. This can be interpreted as changing all the wards to the specified employment density range. For example, this means that if all the wards have an employment density of 3000 / km² or greater, then total fatalities would increase by 37.94%. This should be considered in the context that only 6.5% of the wards in the sample have this level of employment density (see Table 1 and repeated in first column of Table 4).

Results for population density have a more pronounced and stronger effect, with large levels of statistical significance. In general, those areas with lower population density experience relatively more casualties (of all types) while those areas with higher population densities have lower casualties. Actual estimated increases in frequencies, as shown in Table 4 are also quite large. While population density and employment density are not highly correlated in our dataset (due to the small areas analyzed), this suggests perhaps that traffic within commercial as opposed to residential areas may increase the risk of casualties, while those areas with high residential population density have relatively fewer casualties.

Dummy variables for the level of urbanization within each ward were also included in the model. In general, we find that the larger the level of urbanization the fewer casualties. Those areas that are predominantly urban (75% or more urban development) have significantly fewer casualties of all types. Interestingly, results in Table 4 imply that if all areas were similar to the category of 75% or more urban development, total casualties would increase since fewer areas would then be 100% urban (which represents the vast majority, 59.3%, of the wards in the data). Mixed urban and the rural categories result in relatively more casualties compared to more

urbanized areas, especially for slight injuries. Clearly, urbanized wards dominate our data set, primarily since these tend to be smaller in area than more rural wards, since ward size is set by the number of households within it (approximately 2000 per ward). We also tested models that explicitly controlled for area, but found no major difference in results.

Our data do not completely explain why these land use effects may be occurring. We suspect that the ability of vehicles to travel faster in rural areas may play a role. These effects do seem to reinforce the conclusion that land use conditions may affect the relative speed of vehicles. More congested urban areas would be less likely to result in fatalities (all else equal) than uncongested rural areas, though the likelihood of more minor (slight injury crashes) is relatively more likely in congested urban areas, though still less likely than in rural areas. The results for population density would also tend to confirm the finding that more congested areas result in relatively fewer casualties, although this is less likely to occur in areas with higher employment density, which may be equally congested. Some studies that have found more congested roads to be safer (Zhou & Sisiopiku, 1997) and suggests that policies of reducing congestion in urbanized areas may have unanticipated safety consequences (Shefer & Rietveld, 1997).

Another interpretation may be that road designs and layouts lead to slower traffic within urbanized areas with consequent safety benefits independent of the actual land use. Our analyses also controls for and tests the associations with various roadway characteristics but not in the detail needed to address this question. These effects are discussed in the next section.

Road characteristics

In general, we would expect higher functional roadway categories (such as motorways) to be associated with fewer casualties. We also expect roundabouts to be associated with fewer

casualties and increased number of traffic nodes (essentially junctions) to be associated with more casualties. We normalize the variables by area as this gives a better indication of the intensity of road development in each ward.

We find that roadway characteristics seem to have little effect on total fatalities (Table 2). None of our coefficient estimates are statistically significant. Road lengths are, however, associated with increases in serious injuries. “B” roads tend to have the highest association with serious injuries while “A” roads have a positive but slightly lower association. Motorways, surprisingly, have a positive association with increased serious injuries, although only at the 90% level of significance. Minor road length does not affect serious injuries and even seems to be associated with decreases in slight injuries, especially with motorized injuries (Table 3). Increased slight injuries are also associated with increased “B” road length. Surprisingly, increased number of roundabouts also tends to increase slight injuries, although only at the 90% confidence level. These results generally hold for both total casualties and motorized casualties.

“A” and “B” roads would naturally tend to have lower speeds than motorways, thus it is not surprising that they are not associated with increased fatalities. Motorways, which are controlled access highways, will tend to be relatively safer and thus it is also not surprising that they are not associated with increased fatalities, despite the higher speeds on these roads. The larger association of “A” and “B” roads with serious and slight injuries may suggest some problems with these roads, for example, perhaps design speeds are frequently exceeded or interactions between vehicles tend to be greater. The intensity of junctions and roundabouts, surprisingly, has no major association with any of our casualty measures, except slight injuries. In all cases, the elasticity values associated with these variables are negligible, implying that small differences in road characteristics between the various wards, has little effect on casualties.

Demographic characteristics

The key area-specific demographic measure that we analyze is the index of multiple deprivation (IMD). Our results show that areas that have higher IMD scores (i.e., are more deprived) are associated with increased serious and slight injuries (Table 2). We also find a positive association with increased fatalities but only at the 90% level of significance. This is consistent with the results of Abdalla et al (1997), Chichester et al. (1998) and Beattie et al. (2001) who also found that there is a positive correlation between area deprivation and traffic casualties. However, our results are less convincing when we look at only motorized casualties. In this case, we see that only slight injuries have a positive association with the index of multiple deprivation. Graham et al. (2003) find a positive association with child and adult pedestrian casualties, which is consistent with our result and suggests that this is the effect driving the results of the models in Table 2.

Elasticities of casualties with respect to the deprivation score range from 0.153 for total slight injuries to -0.073 for motorized fatalities. Interestingly, the total fatality elasticity is 0.077 slightly larger the absolute value of the negative elasticity for total motorized fatalities. This may imply that any reduction in motorized fatalities from increased deprivation may be off-set by near equivalent increases in non-motorized fatalities.

It is unclear how area deprivation affects casualties. Graham et al. (2003) by analyzing pedestrian casualties argue that most of the pedestrians are casualties in wards in which they most likely live. This is probably most true of children. For our results, the strongest association for motorized casualties is for slight injuries and it is unlikely that many of the casualties are residents of the ward in which the accident occurred. Therefore, it poses an interesting question as to what factors associated with deprived areas lead to this association with slight injuries for

accidents with no pedestrians or bicyclists? One interpretation may be that relatively lower income deprived areas are located in areas with less safe traffic characteristics.

One source could perhaps be unmeasured differences in road characteristics between more deprived and less deprived areas. An analyses of the road characteristics for which we have data (and which are controlled for in the analyses) shows that there is no inherent pattern between road characteristics and deprivation. This is shown in Table 6. We see, for example, that the most deprived and the least deprived areas tend to have the least amount of road kilometers per square mile. There is also no pattern to the number of roundabouts and node junctions in each area. While there may be some unmeasured road factor not accounted for, this data suggests that road characteristics are similar for the most and least deprived areas.

Unmeasured road factors may display a similar pattern, although we cannot say for sure.

Other demographic factors that were controlled for include age cohorts of the ward population and the percent of the population that is economically active. The latter is associated with more slight injuries (Tables 2 and 3). We find some effects from the age cohorts at the 95% confidence level for total fatalities, but little difference between coefficient values. However, the elasticity values associated with these variables are quite large in all cases. Ideally we would like to have examined smaller cohort ranges, especially encompassing young drivers (aged 18-25), but this data was not available.

Proxy variables for traffic flow

Finally, we control for exposure to traffic with three different variables, all of which are highly significant. The measure of proximate employment (described above), total employment and total resident population are all proxies for the level of traffic in a given area. Proximate population was also tested and had the same effect but was not included with the final models

due to a high level of collinearity with the proximate employment variable. Elasticity values are relatively high for the proximate employment variable, suggesting that changes in traffic flow (as measured by this variable) can have a substantive effect on increasing casualties, relative to some of the other parameters.

Conclusions

Our analyses has utilized highly disaggregate spatial data for an analyses of area-specific variables associated with road traffic casualties. Our findings suggest that urbanized more densely populated areas will tend to have fewer traffic casualties while areas with higher employment density have more traffic casualties. The first effect may be due to reduced speeds and higher levels of congestion or possibly the lower design speeds of roads in urbanized areas. The latter effect may be due simply to increased levels of street activity in areas with more employment. Results suggest that increasing speeds in urbanized areas by reducing congestion may have adverse safety consequences.

The finding that areas with higher levels of social deprivation have relatively higher casualty rates is also consistent with other findings (Abdalla et al., 1997), although this effect is less strong when only motorized casualties are considered. It is unclear what residual factor associated with more deprived areas may be leading to increased casualty rates from traffic flowing through those areas. We have controlled for many of the road characteristics that would be likely to vary between areas. It might be that lower income groups tend to locate in areas with more unsafe traffic, due to lower housing costs being associated with less safety, but we cannot determine this from our data. Additional work should be conducted to identify the characteristics of these areas (or their populations) that may adversely affect road safety.

In general, we did not find that roadway characteristics (as we defined them) mattered in terms of affecting fatalities. Increased road length was associated with increased serious injuries, especially for “B” roads, which were also associated with increased slight injuries. We did not find that more roundabouts in an area made a difference, nor do more junctions.

Further analyses of our data can exploit additional spatial relationships. This includes potential spatial autocorrelation that can help identify similar unidentified factors between areas that affect casualties. Time series data would also allow a more rigorous analyses of changes over time and would enable us to avoid problems of heterogeneity in the data. In addition, disaggregation of the STATS19 data offers the potential for analyzing additional effects on more specific types of traffic-related casualties.

Acknowledgements

This work was funded by the Engineering and Physical Sciences Research Council. We would like to thank Daniel Graham for helpful comments on the analyses and data acquisition. STATS 19 data was supplied by the ESRC UK data archives at the University of Essex.

References

- Abdalla, Ibrahim M, Robert Raeside, Derek Barker, and David R.D. McGuicagan, 1997. An investigation into the relationships between area social characteristics and road accident casualties. *Accident Analysis and Prevention*, 29(5): 583-593.
- Avery, J G., Vaudin, J. N., Fletcher, J.L., Watson, J M. 1990. Geographical and social variations in mortality due to childhood accidents in England and Wales 1975-1984. *Public Health*. 104 (3).
- Beattie, T.F., Gorman, D. R, Walker, J. J. 2001. The association between deprivation levels, attendance rate and triage category of children attending a children's accident and emergency department. *Emergency Medicine Journal*, 18(2).
- Chichester, B. M., Gregan, J. A., Anderson, D. P., Kerr, J. M. 1998. Associations between road traffic accidents and socio-economic deprivation on Scotland's west coast. *Scottish Medical Journal*, 43(5).
- Department of Environment, Transport and the Regions (DETR), 2000, *Indices of Deprivation 2000*.
- Gourieroux, C., Monfort, A. and Trognon, A. Pseudo maximum likelihood methods: applications to Poisson models. *Econometrica*, 52, pp. 701-720. 1984.
- Graham D.J. and Glaister S. (in press) 'Spatial variation in road pedestrian casualties: the role of urban scale, density and land use mix' *Urban Studies*.
- Graham DJ, Glaister S and Anderson R (2003) 'Child pedestrian casualties in England: the effect of area deprivation.' Submitted to the *Journal of Regional Science*.
- Hausman, J.C., Hall, B.H. and Griliches, Z. Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, 52(4), pp. 909-938. 1984.
- Jessop, E G. 1996. Deprivation and mortality in non-metropolitan areas of England and Wales. *Journal of Epidemiology and Community Health*, Vol 50 (5).
- Jovanis, P., Chang, H., 1986, Modeling the relationship of accidents to miles travelled, *Transportation Research Record* 1068, 42-51.
- Kim, K. and Yamashita, E. 2002. Motor vehicle crashes and land use: an empirical analysis from Hawaii. Working paper. Department of Urban and Regional Planning. University of Hawaii.
- Miaou, Shaw-Pin and Harry Lum, 1993, Modeling Vehicle Accidents and Highway Geometric Design Relationships, *Accident Analysis and Prevention*, 25(6): 689-709.
- Miaou, S.-P. The Relationship Between Truck Accidents and Geometric Design of Road section: Poisson versus Negative Binomial Regression. *Accident Analysis and Prevention*, 26(4), pp. 471-482. 1994.

Ossenbruggen, P.J., Pendharkar, J., Ivan, J. 2001. Roadway safety in rural and small urbanized areas. *Accident Analysis and Prevention* 33 (4).

Roberts, I; Marshall, R; Norton, R; Borman, B. 1992. An area analysis of child injury morbidity in Auckland. *Journal of Paediatrics and Child Health*, Vol 28(6).

Senior, M., Williams, H., Higgs, Gary. 2000. Urban-rural mortality differentials: controlling for a material deprivation. *Social Science and Medicine*. Vol 51.p. 289-305.

Shankar, V.N., Mannering, F. and Barfield, W. Effect of Roadway Geometric and Environmental Factors on Rural Freeway Accident Frequencies. *Accident Analysis and Prevention*, 27(3), pp. 371-389. 1995.

Shefer, Daniel and Piet Rietveld, 1997, Congestion and safety on highways: towards an analytical model, *Urban Studies*, 34: 679-692.

Vogt, A. and Bared, J. Accident models for two-lane rural segments and intersections. *Transportation Research Record*, 1635, pp. 18-29. 1998.

Zhou, Min and Virginia P. Sisiopiku, 1997, On the relationship between volume to capacity ratios and accident rates, Paper presented at the 76th Annual Meeting of the Transportation Research Board, Washington, DC.

Figure 1: Census wards boundaries for England (all 8414 wards)

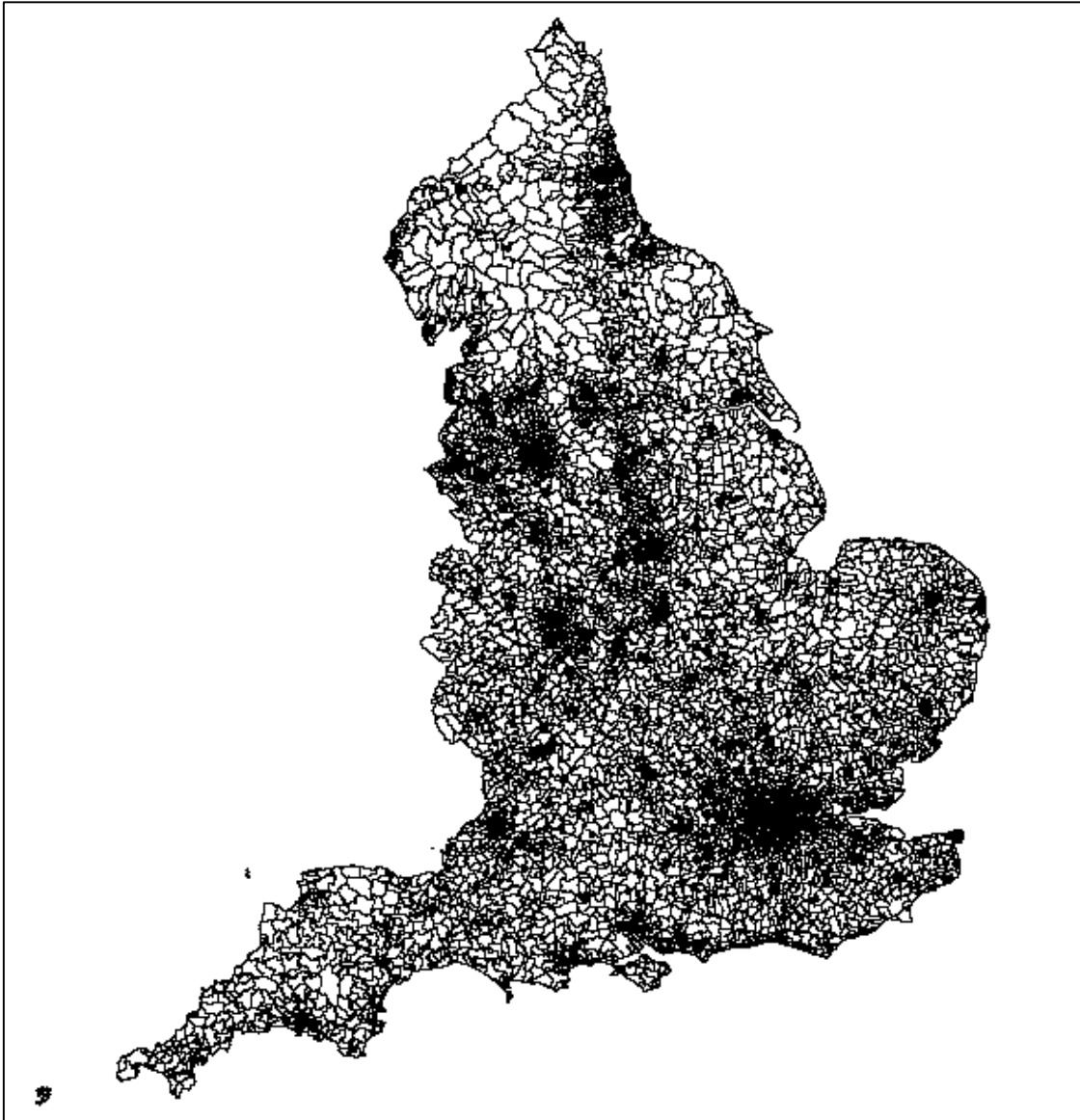


Figure 2: A portion of road network from inner London

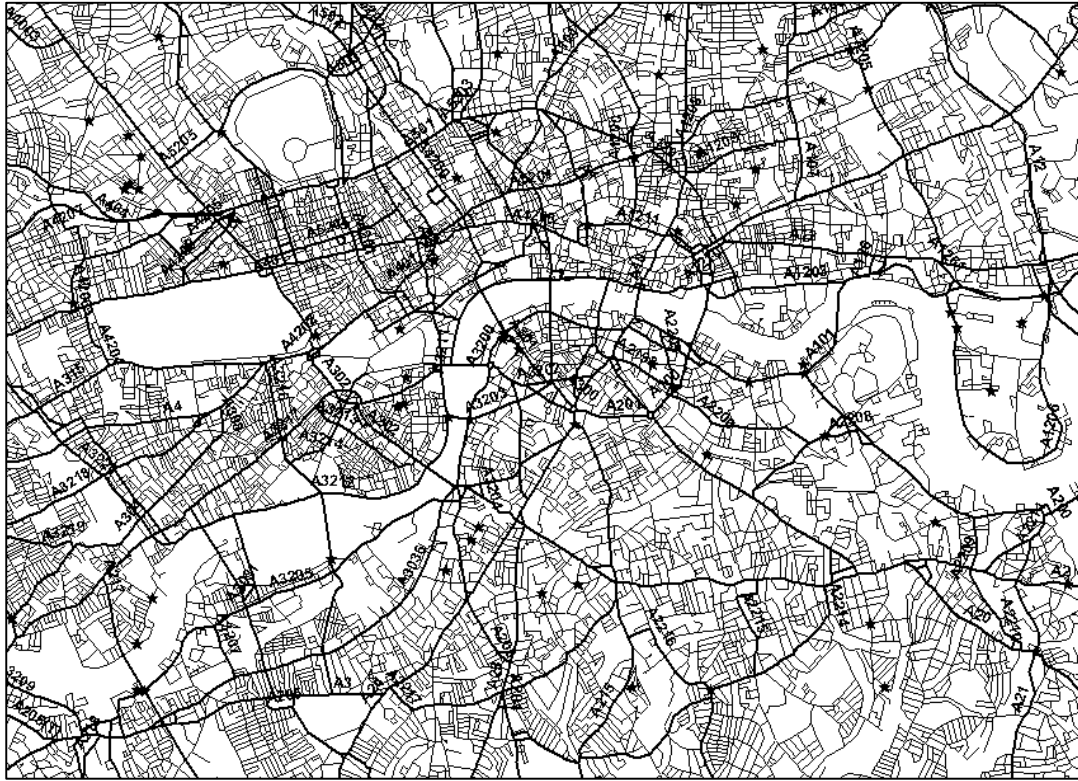


Table 1: Summary statistics of variables included in the models

Variables	N	Mean	Std. Dev.	Min	Max
Dependent variables					
Total fatalities	8200	0.342	0.72	0	10
Total serious injuries	8200	3.979	4.29	0	78
Total slight injuries	8200	29.358	32.62	0	573
Land use (indicator variables)					
Employment density $\leq 1500 / \text{km}^2$	8200	0.846	0.36	0	1
Employment density $\leq 3000 / \text{km}^2$ and $> 1500 / \text{km}^2$	8200	0.089	0.29	0	1
Employment density $> 3000 / \text{km}^2$	8200	0.065	0.25	0	1
Population density $\leq 3000 / \text{km}^2$	8200	0.712	0.45	0	1
Population density $\leq 6000 / \text{km}^2$ and $> 3000 / \text{km}^2$	8200	0.208	0.41	0	1
Population density $> 6000 / \text{km}^2$	8200	0.08	0.27	0	1
Wholly urban	8198	0.593	0.49	0	1
Predominantly Urban (75%+)	8198	0.113	0.32	0	1
Mixed Urban (50 - 75%)	8200	0.061	0.24	0	1
Mixed Rural (25 - 50%)	8198	0.194	0.4	0	1
Predominantly Rural (1-25%)	8198	0.025	0.16	0	1
Wholly rural	8198	0.012	0.11	0	1
Road characteristics					
Motorway length (m) per m^2 of area	8200	2.01×10^{-4}	7.92×10^{-3}	0	0.529
A road length (m) per m^2 of area	8200	3.015×10^{-3}	8.94×10^{-2}	0	7.159
B road length (m) per m^2 of area	8200	1.385×10^{-3}	2.81×10^{-2}	0	1.41
Minor road length (m) per m^2 of area	8200	16.427×10^{-3}	0.312	0	16.909
Number of roundabouts per m^2 of area*10000	8200	1.104×10^{-2}	0.223	0	12.16175
Number of junctions per m^2 of area	8200	7.089×10^{-5}	9.97×10^{-4}	0	1.35×10^{-2}
Demographic characteristics					
Index of multiple deprivation score	8197	21.684	15.41	1.16	83.77
Percentage of residential population under 16	8197	19.776	3.64	1	41
Percent of residential population aged 16 to 59	8197	58.433	5.22	30	97
Percent of residential population over 60	8197	21.795	6.01	2	62
Percent of population that are economically active	8197	47.07	5.24	13	73
Proxy for traffic flow					
Proximate.emp	8200	196.484	99.04	42.47	887.57
Total employment	8197	5850.555	4223	300	137191
Total resident population	8197	2285.868	4748.33	14	35100

Table 2: Negative binomial (NB) models for total fatalities, serious injuries, and slight injuries

Dependent Variable	Total Fatalities		Serious Injuries		Slight Injuries	
	Model A		Model B		Model C	
Variables	Coef.	β/S.E.	Coef.	β/S.E.	Coef.	β/S.E.
Land use						
Employment density ≤ 1500 / km ²	-0.151	-1.62	-0.093	-2.59	-0.156	-5.60
Employment density ≤ 3000 / km ² and > 1500 / km ² (Reference variable)						
Employment density > 3000 / km ²	0.209	1.63	0.136	2.75	0.067	1.73
Population density ≤ 3000 / km ²	0.404	5.38	0.165	5.87	0.205	9.71
Population density ≤ 6000 / km ² and > 3000 / km ² (Reference variable)						
Population density > 6000 / km ²	-0.564	-4.43	-0.210	-4.76	-0.232	-6.90
Wholly urban (100%)	-1.177	-15.86	-0.752	-24.51	-21.350	-10.35
Predominantly Urban (75%+)	-0.679	-7.98	-0.456	-12.44	-11.480	-4.35
Mixed Urban (50 - 75%)	-0.043	-0.49	-0.118	-2.81	-2.240	3.08
Mixed Rural (25 - 50%)	0.175	1.47	0.092	1.59	1.800	5.83
Predominantly Rural (1-25%)	-0.179	-0.96	0.109	1.34	1.200	6.54
Wholly rural (Reference variable)						
Road characteristics						
Motorway length (m) per m ² of area	-3.34	-0.65	1.9	1.61	1.34	1.3
A road length (m) per m ² of area	0.355	0.43	0.515	2.33	0.272	1.5
B road length (m) per m ² of area	-0.232	-0.16	1.05	2.67	0.834	2.38
Minor road length (m) per m ² of area	-0.432	-0.85	-0.129	-1.25	-0.103	-1.49
Number of roundabouts per m ² of area*10000	0.135	0.42	0.089	0.79	0.144	1.7
Number of junctions per m ² of area	90.9	0.98	-20.1	-0.6	-23.4	-1.02
Demographic characteristics						
Index of multiple deprivation score	3.55E-03	1.57	5.36E-03	6.04	7.04E-03	10.47
Percentage of residential population under 16	7.39E-02	1.61	1.70E-02	0.90	-6.20E-04	-0.04
Percent of residential population aged 16 to 59	7.96E-02	1.73	1.80E-02	0.95	1.29E-02	0.91
Percent of residential population over 60	7.83E-02	1.71	2.48E-02	1.32	7.26E-03	0.51
Percent of population that are economically active	4.81E-03	0.53	9.76E-03	2.75	1.24E-03	0.46
Proxy for traffic flow						
Proximate employment	1.96E-03	6.17	2.59E-03	21.08	2.15E-03	22.14
Total employment	8.38E-05	13.04	8.83E-05	30.19	1.14E-04	45.67
Total resident population	1.49E-05	3.36	2.61E-05	8.99	3.99E-05	13.48
Constant						
Constant	-9.61	-2.10	-1.87	-1.00	1.11	0.79
k (constant) for NB model	0.918		0.384		0.338	
Number of observations	8197		8197		8197	
Log-likelihood function at zero	-6232.0		-20366.4		-35873.7	
Log-likelihood function at convergence	-5937.9		-18885.2		-32651.4	

Table 3: Negative binomial (NB) models for motorized fatalities, motorized serious injuries, and motorized slight injuries

Dependent Variable	Motorized Fatalities		Motorized Serious Injuries		Motorized Slight Injuries	
	Model D		Model E		Model F	
Variables	Coef.	β/S.E.	Coef.	β/S.E.	Coef.	β/S.E.
Land use						
Employment density ≤ 1500 / km ²	0.0298	0.23	-0.0214	-0.47	-0.1240	-4.01
Employment density ≤ 3000 / km ² and > 1500 / km ² (Reference variable)						
Employment density > 3000 / km ²	0.0621	0.32	0.0895	1.42	0.0319	0.74
Population density ≤ 3000 / km ²	0.5806	5.69	0.3396	9.58	0.2880	12.28
Population density ≤ 6000 / km ² and > 3000 / km ² (Reference variable)						
Population density > 6000 / km ²	-0.5289	-2.86	-0.2857	-5.09	-0.2845	-7.69
Wholly urban (100%)	-1.3713	-15.35	-1.0491	-28.66	-0.5519	-21.78
Predominantly Urban (75%+)	-0.7491	-7.46	-0.5889	-13.65	-0.2908	-9.51
Mixed Urban (50 - 75%)	-0.0532	-0.52	-0.1825	-3.75	0.0393	1.11
Mixed Rural (25 - 50%)	0.1283	0.93	0.0722	1.07	0.2288	4.59
Predominantly Rural (1-25%)	-0.1512	-0.72	0.1000	1.07	0.4039	5.87
Wholly rural (Reference variable)						
Road characteristics						
Motorway length (m) per m ² of area	3.66	0.57	2.85	1.87	1.78	1.5
A road length (m) per m ² of area	2.23	1.05	0.672	2.35	0.358	1.74
B road length (m) per m ² of area	-0.696	-0.13	1.69	3.32	1.22	3
Minor road length (m) per m ² of area	-1.28	-0.94	-0.139	-1.09	-0.145	-1.89
Number of roundabouts per m ² of area*10000	1.25	1.38	0.0856	0.61	0.239	2.28
Number of junctions per m ² of area	-170	-0.49	-28.2	-0.68	-40	-1.55
Demographic characteristics						
Index of multiple deprivation score	-3.38E-03	-1.16	6.62E-04	0.6	5.61E-03	7.47
Percentage of residential population under 16	0.1167	2.06	0.0212	0.93	-4.77E-03	-0.31
Percent of residential population aged 16 to 59	0.1277	2.25	0.0159	0.69	1.07E-02	0.68
Percent of residential population over 60	0.1163	2.06	0.0246	1.08	1.31E-03	0.08
Percent of population that are economically active	-1.43E-03	-0.13	0.0094	2.06	-2.67E-03	-0.84
Proxy for traffic flow						
Proximate employment	1.71E-03	3.94	3.35E-03	21.74	2.32E-03	21.54
Total employment	1.41E-05	2.44	2.27E-05	6.33	3.96E-05	12.03
Total resident population	7.49E-05	9.07	8.29E-05	22.96	1.15E-04	41.78
Constant						
Constant	-14.1350	-2.51	-2.1688	-0.95	1.4691	0.95
k (constant) for NB model	1.413162		0.570061		0.4085	
Number of observations	8197		8197		8197	
Log-likelihood function at zero	-6232		-20366.4		-35873.7	
Log-likelihood function at convergence	-4596.0847		-16822.173		-31639.976	

Table 4: Percent change in accident frequencies using predicted values

	Mean value in data	Total Fatalities: model A	Serious Injuries: Model B	Slight Injuries: Model C	Motorized Fatalities Model D	Motorized Serious Injuries: Model E	Motorized Slight Injuries: Model F
Actual count		2803	32625	240734	1930	22640	195200
Predicted count		2832	34812	344696	1937	23695	273632
Employment density $\leq 1500 / \text{km}^2$	84.6%	-3.73%	-4.65%	-9.25%	0.05%	-1.58%	-9.25%
Employment density $\leq 3000 / \text{km}^2$ and $> 1500 / \text{km}^2$ (Reference variable)	8.9%						
Employment density $> 3000 / \text{km}^2$	6.5%	37.94%	19.81%	13.49%	3.33%	9.97%	13.49%
Population density $\leq 3000 / \text{km}^2$	71.2%	17.84%	9.95%	12.99%	15.13%	15.91%	12.99%
Population density $\leq 6000 / \text{km}^2$ and $> 3000 / \text{km}^2$ (Reference variable)	20.8%						
Population density $> 6000 / \text{km}^2$	8.0%	-55.20%	-24.44%	-27.00%	-62.04%	-37.98%	-27.00%
Wholly urban (100%)	59.3%	-34.04%	-16.49%	-6.77%	-45.04%	-26.89%	-6.77%
Predominantly Urban (75%+)	11.3%	8.50%	12.24%	13.42%	2.39%	15.84%	13.42%
Mixed Urban (50 - 75%)	6.1%	104.89%	57.50%	47.93%	105.34%	73.93%	47.93%
Mixed Rural (25 - 50%)	2.5%	154.83%	94.32%	76.97%	146.19%	124.38%	76.97%
Predominantly Rural (1-25%)	1.2%	78.86%	97.50%	106.48%	86.17%	130.70%	106.48%
Wholly rural (Reference variable)	19.4%						

Table 5: Mean elasticity estimates for continuous variables

	Total Fatalities: Model A	Serious Injuries: Model B	Slight Injuries: Model C	Motorized Fatalities Model D	Motorized Serious Injuries: Model E	Motorized Slight Injuries: Model F
Road characteristics						
Motorway length (m) per m^2 of area	-0.001	0.000	0.000	0.001	0.001	0.000
A road length (m) per m^2 of area	0.001	0.002	0.001	0.007	0.002	0.001
B road length (m) per m^2 of area	-0.000	0.001	0.001	-0.001	0.002	0.002
Minor road length (m) per m^2 of area	-0.007	-0.002	-0.002	-0.021	-0.002	-0.002
Number of roundabouts per m^2 of area*1000	0.001	0.001	0.002	0.014	0.001	0.003
Number of junctions per m^2 of area	0.006	-0.001	-0.002	-0.012	-0.002	-0.003
Demographic characteristics						
Index of multiple deprivation score	0.077	0.116	0.153	-0.073	0.014	0.122
Percentage of residential population under 16	1.461	0.336	-0.012	2.308	0.419	-0.094
Percent of residential population aged 16 to 59	4.651	1.052	0.754	7.462	0.929	0.625
Percent of residential population over 60	1.707	0.541	0.158	2.535	0.536	0.029
Percent of population that are economically active	0.226	0.459	0.058	-0.067	0.442	-0.126
Proxy for traffic flow						
Proximate employment	0.385	0.509	0.422	0.336	0.658	0.456
Total employment	0.490	0.517	0.667	0.082	0.133	0.232
Total resident population	0.034	0.060	0.091	0.171	0.189	0.263

Table 6
Road characteristics versus deprivation scores

Quintile distribution of deprivation scores	Average Index of multiple deprivation score	Motorway length (m) per m ² of area	A road length (m) per m ² of area	B road length (m) per m ² of area	Minor road length (m) per m ² of area	Number of roundabouts per m ² of area*10000	Number of junctions per m ² of area
1: 20% most deprived	47.31	2.0 x 10 ⁻⁴	1.22 x 10 ⁻³	5.6 x 10 ⁻⁴	9.88 x 10 ⁻³	8.7 x 10 ⁻³	6.72 x 10 ⁻⁵
2	26.21	3.2 x 10 ⁻⁴	2.71 x 10 ⁻³	2.6 x 10 ⁻³	2.23 x 10 ⁻²	2.02 x 10 ⁻²	1.15 x 10 ⁻⁴
3	17.09	8.0 x 10 ⁻⁵	6.82 x 10 ⁻³	3.7 x 10 ⁻⁴	2.36 x 10 ⁻²	1.16 x 10 ⁻²	7.02 x 10 ⁻⁵
4	11.4	4.0 x 10 ⁻⁵	3.05 x 10 ⁻³	2.30 x 10 ⁻³	1.73 x 10 ⁻²	7.9 x 10 ⁻³	5.51 x 10 ⁻⁵
5: 20% least deprived	6.4	3.6 x 10 ⁻⁴	1.27 x 10 ⁻³	1.08 x 10 ⁻³	9.10 x 10 ⁻³	6.8 x 10 ⁻³	4.68 x 10 ⁻⁵