



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.

  
C O M M O N S D E E D

**Attribution-NonCommercial-NoDerivs 2.5**

**You are free:**

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

**Under the following conditions:**

 **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/>

# **Congestion and Safety: A Spatial Analysis of London**

**Robert B. Noland**

**Mohammed A. Quddus**

Centre for Transport Studies

Dept. of Civil and Environmental Engineering

Imperial College London

London SW7 2AZ

Ph: 44(0)20-7594-6036

Fx: 44(0)20-7594-6102

email: [r.noland@ic.ac.uk](mailto:r.noland@ic.ac.uk)

# **Congestion and Safety: A Spatial Analysis of London**

## **Abstract**

A disaggregate spatial analysis, using enumeration district data for London was conducted with the aim of examining how congestion may affect traffic safety. It has been hypothesized that while congested traffic conditions may increase the number of vehicle crashes and interactions, their severity is normally lower than crashes under uncongested free flowing conditions. This is primarily due to the slower speeds of vehicles when congestion is present. Our analysis uses negative binomial count models to examine whether factors affecting casualties (fatalities, serious injuries and slight injuries) differed during congested time periods as opposed to uncongested time periods. We also controlled for congestion spatially using a number of proxy variables and estimated pedestrian casualty models since a large proportion of London casualties are pedestrians. Results are not conclusive. Our results suggest that road infrastructure effects may interact with congestion levels such that in London any spatial differences are largely mitigated. Some small differences are seen between the models for congested versus uncongested time periods, but no conclusive trends can be found. Our results lead us to suspect that congestion as a mitigator of crash severity is less likely to occur in urban conditions, but may still be a factor on higher speed roads and motorways.

## **Introduction**

Congestion reduction is often stated to be one of the primary goals of transport policy. This is seen as desirable due to the economic costs associated with traffic congestion. Travel delay is seen as both reducing economic productivity and reducing the quality of life of those stuck in traffic. All else equal, congestion reduction is a desirable policy.

However, all else is not normally equal. In particular, it has been speculated that there may be an inverse relation between congestion reduction and improved safety (Shefer & Rietveld, 1997). Free flowing traffic will normally travel at speeds that make the likelihood of fatalities higher when accidents occur. On the other hand, congested traffic can slow traffic such that fatalities are unlikely to occur in the event of an accident. Increased traffic levels and congested traffic will increase the number of interactions between vehicles and potential collisions. Therefore, it is plausible that while more congested traffic may lead to more accidents, the outcomes of those accidents are less severe.

This poses a potential policy dilemma for decision makers. External costs associated with congestion may be off-set by external benefits associated with fewer traffic fatalities due to congestion.

Recent empirical evidence suggests that these effects do occur. Zhou & Sisiopiku (1997) examined a sixteen mile segment of an Interstate freeway in Michigan. They correlated accidents with the volume to capacity ratio of the freeway finding that this tended to follow a U-shaped curve. That is, more accidents occurred when vehicle flow was relatively low and when it was relatively high. However, more severe injury and fatal accidents tended to decrease as the volume to capacity ratio increased, strongly suggesting that more congested links lead to lower accident fatalities. Another study of traffic flow on interurban French motorways found that crash severity levels are greater during night-time hours (Martin, 2002). A study of highways in Connecticut also found that larger volume to

capacity ratios are associated with a reduced number of crashes (Ivan et al., 2000). This latter result is not necessarily consistent with the theory formulated by Shefer & Rietveld as this result included all levels of crash severity.

These studies were not focused on urbanized road conditions. Dickerson et al. (2000) used data from London (as we do) to examine the relationship between accident externalities and traffic flow. While not disaggregating by severity levels, they conclude that at high traffic flows there is a substantial negative externality while at low flow levels there is not. This result suggests that the theory elaborated by Shefer & Rietveld (1997) may not hold but one can't say for certain without disaggregating accidents by severity level.

The study reported here attempts to do this, again using data from London. Our data set consists of highly disaggregate spatial units, enumeration districts, for which we have over 15,000 units. This data is overlaid with accident data from the STATS19 database by severity level of accident and time of day of the accident. This gives us data on fatal, serious injury, and slight injury accidents which we disaggregate by day and nighttime periods. We also disaggregate pedestrian casualties since about 19% of all the casualties in our database are pedestrians. Our models include other spatial data that may be associated with accidents, including road network data and demographic data. We control for inner and outer London areas and various other proxies for congestion. Therefore we estimate models that attempt to control for congestion both temporally and spatially.

Results are inconclusive and we cannot confirm the hypothesis of Shefer & Rietveld (1997). We find that our proxy variables for relative congestion do not seem to be giving the anticipated result. In general we find no differences between Inner and Outer London. Minor differences are found between models for congested versus uncongested time periods but no conclusive trends can be found. Overall, our results tend to suggest that these effects are minor or non-existent in London. We speculate that this may be because speeds are

generally low both in congested and uncongested areas and that those areas that are congested already have infrastructure in place that mitigates the safety effects associated with high speed traffic.

The data sources used for this analysis are described in some detail in the next section. We then present the statistical method used followed by a presentation and interpretation of the results. Conclusions are then presented.

## **Data**

Our analysis focuses on the Greater London metropolitan area. The total area of the metropolitan area is about 1580 sq km of which Inner London is about 320 sq km (20.2%) and Outer London is 1260 sq km (79.7%). The boundaries are shown in Figure 1. The spatial data unit used in our study is the Enumeration District (ED). These consist of an average of 200 households each and for the Greater London area there are 15,366 units. These are shown graphically in Figure 2, which also shows the distribution of traffic casualties within each ED. The geographical size of the ED's varies since they are delineated based on the number of households in each. The average size of each ED is about 10 hectares, with the smallest being 0.065 hectares and the largest being about 950 hectares.

The digital ED boundary data and road network data were obtained from UK Ordnance Survey (OS) data via EDINA services (UKBORDERS and Digimap respectively). Data on traffic casualties was extracted from the STATS19 database and is geo-coded and assigned to the ED's using a geographic information system (GIS). Demographic data on population and employment was obtained from the Office of National Statistics. Car ownership data was based on 1991 UK census data. Information on hospital locations was supplied from NavTech by Saturn Technology UK. All these databases were integrated using a GIS that allowed aggregation to the ED level.

The STATS19 UK national road accident database has information on the recorded location of an accident, identified by its  $x$  and  $y$  coordinates corresponding to 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. Of these other factors we also disaggregate by time of day of the accident and weekdays and weekends. This allows us to have accident data for a congested time period (7:00am – 8:30pm on weekdays) and uncongested time period (8:30pm-7:00am on weekdays). We disregarded weekend data for this analysis since weekend patterns of congestion may be quite different. Casualty data was aggregated for three years of data, 1999-2001.

The total number of casualties appears to vary significantly between the congested and uncongested time periods. Over 82% of total casualties occur during congested periods while only about 18% occur during the uncongested times of the day (excluding weekends). Table 1 shows that during the congested time period the percent of accidents that are fatalities is lower, at 69% relative to the percent of serious and slight injuries (79% and 83%, respectively). Inner London, which is relatively more congested also has a lower percent of fatalities (43%) relative to Outer London (57%), while the percent of serious and slight injuries is larger in Inner London. Similar trends also hold for pedestrian casualties, with the percent fatalities being lower during congested time periods and in Inner London, relative to injuries. This would tend to suggest that congestion would tend to mitigate the more severe accidents, as suggested by Shefer and Rietveld (1997).

To obtain data on features of the road network, data was obtained from EDINA Digimap Meridian2<sup>TM</sup> 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 coordinate

resolution of the data is one 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. Different infrastructure features were extracted from the data including motorway, A road, B road, minor road, roundabouts and nodes. Three leg and four leg junctions were derived from the node data using a MapBasic program. These are identified within the database distinct from just nodes as shown in Figure 3. In our analysis we normalize these by land area so that we can represent the intensity of these road network features within a given ED.

Total residential population was based on 1998 estimates from the Office of National Statistics (ONS). Total employment data at the ED level were obtained from ONS, Labour Force Survey, 2000. Data on household car ownership at the ED level were obtained from the 1991 UK census. Unfortunately, we do not have recent year car ownership data at the ED level.

Information on public transport accessibility was obtained from Transport for London (TfL). This included the location of bus stops, underground stations and railway stations. These were aggregated to the ED level using our GIS. Bus stop data was normalized by area while for underground and railway stations we include a dummy variable if at least one is located in the ED. Spatial location of national railway stations and underground stations within Greater London are shown in Figure 4.

Recent work has shown the importance of medical technology in reducing fatalities in accidents (Noland & Quddus, in press; Noland, in press). To control for this we included data on the location of hospitals. This data was obtained from NavTech-European data supplied by Saturn Technology Ltd UK. There are a total of 52 hospitals within Greater London. Distance of the nearest hospital was calculated from the centroid of each ED to the



geo-coded location of the hospital. The minimum distance of the nearest hospital from the ED centroid was 35m and the maximum distance was 7542m.

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 ED. Instead we use a method devised by Graham and Glaister (in press) to represent the level of travel activity based upon the levels of employment in each ED relative to all other EDs. This proxy variable for the volume of traffic flowing through each ED is derived as,

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

in which  $d_{ij}$  is the centroid distance from ED  $i$  to ED  $j$ . Origin and destination traffic within an ED is proportional to the level of total employment,  $E_i$  for ED  $i$ . Therefore, following Graham and Glaister (in press), we define proximate employment,  $PE_i$  for each ED to represent the volume of traffic passing through the ED. Clearly this will be heavily weighted based upon the employment of neighboring EDs with more distant EDs receiving less weight. Intuitively this is essentially a gravity measure of the distribution and interaction between activities, commonly used to determine how traffic is distributed within a region.

Summary statistics for the variables used in the analyses is presented in Table 2 and Table 3. Note that the total casualty figures are only for the weekdays; summary statistics for the congested versus the uncongested period are also shown.

### **Statistical Methodology**

The Poisson regression model is a natural first choice for modeling discrete, random, nonnegative and sporadic events such as traffic accidents. The Poisson distribution has only one adjustable parameter, namely the mean  $\mu$ , which must be positive. A log-linear

relationship between the expected number of casualties and explanatory variables is a more commonly used formulation, i.e.,

$$\mu_i = E(n_i | \mathbf{X}_i) = \exp(\boldsymbol{\beta}\mathbf{X}_i), \quad n_i = 0, 1, 2, \dots \quad (2)$$

In this case  $n_i$  is the number of casualties occurring on each ED over a given time period,  $\mathbf{X}_i$  is a vector of explanatory variables indicating infrastructure, demographic and other characteristics of the ED  $i$  and  $\boldsymbol{\beta}$  is a vector of estimable coefficients. The probability density function can be expressed as

$$\Pr(n_i | \mathbf{X}_i) = \frac{\exp(-\mu_i)\mu_i^{n_i}}{n_i!} \quad (3)$$

However, the Poisson regression model has some potential problems. The rate at which events occur during a period of observations may not be a constant, that is, the variance of  $n_i$  may not be equal to the mean. Various deviations from the basic Poisson process can result in overdispersion in the data (Washington et al., 2003). The solution is to apply the Negative Binomial (NB) model. The NB model is derived from the Poisson model by incorporating a stochastic component in the relationship between  $\mu$  and  $\mathbf{X}$  where the mean  $\mu$  is replaced with the random variable  $\tilde{\mu}$  i.e.,

$$\tilde{\mu}_i = \exp(\boldsymbol{\beta}\mathbf{X}_i + \varepsilon_i) \quad (4)$$

Here  $\varepsilon$  is a non-negative random term, representing the unobserved variation across observations. One can think of  $\varepsilon$  either as the combined effects of unobserved variables that have been omitted from the model or as another source of pure randomness (Washington et al. 2003). The probability density function for the NB distribution can be expressed as

$$\Pr(n_i | \mu_i) = \frac{\Gamma(n_i + (1/k))}{n_i! \Gamma(1/k)} \left( \frac{1/k}{(1/k) + \mu_i} \right)^{1/k} \left( \frac{\mu_i}{(1/k) + \mu_i} \right)^{n_i} \quad (5)$$

Where  $k(\geq 0)$  is often referred to as the overdispersion parameter. In the analysis that follows we estimate NB models since most often accident data will display overdispersion and tests of our models indicate that this is the case for our data.

## **Results**

The basic objective of our analyses was to test whether the hypothesis that congestion on urban streets and highways provides a safety benefit. For this reason our analyses strategy was to evaluate several models. First, we disaggregate the data by severity of the casualty that results, by fatality, serious injury, and slight injury. We also disaggregate the analyses by time of day, specifying a congested time period of 7:00am to 8:30pm on weekdays and an uncongested time period of 8:30pm to 7:00am, also just on weekdays. We control for spatial congestion by including a dummy variable to control for Inner London versus Outer London, with the former tending to be more congested. In addition, we analyze similar models for pedestrian casualties. This is done mainly because pedestrian casualties constitute about 19% of total casualties in London.

Table 4 displays results for a NB model using data from the congested time period (7:00am to 8:30pm, weekdays). Table 5 shows the same analyses using the uncongested time period (8:30pm to 7:00am, weekdays). Similar pedestrian casualty models are shown in Tables 6 and 7. All models contain numerous explanatory variables which are discussed in turn in the sections that follow.

### *Proxy variables to measure congestion*

An indicator variable was included for those ED's located in Inner London. These correspond to the most congested parts of the city. This variable is not statistically different from zero (at the 95% confidence level) for the three models of the congested time period (Table 4). In other words, we are picking up no residual difference between Inner and Outer London that could explain differences in all three types of casualties. We do pick up a slight

negative effect for slight injuries being less likely in Inner London compared to Outer London during the congested time period. One could speculate that congestion levels are similar throughout the region during these hours and thus one would not expect to see large differences, however, on balance congestion is probably more persistent within Inner London compared to Outer London.

In the models for uncongested time periods (Table 5), the Inner London dummy variable has a low level of significance (about the 75% level) in the fatality model. The coefficient value is relatively high at 0.3134 compared to other estimates for this variable. This seems to imply that during uncongested time periods there is a slight difference between Inner and Outer London with the former being more likely to have fatalities, all else equal. Whether this is due to less congestion during these time periods cannot be determined.

The models of pedestrian casualties (Tables 6 and 7) show no significant difference between Inner and Outer London. Inner London clearly has more pedestrian traffic than suburban areas, so perhaps this merely reflects increased exposure levels in Inner London versus increased risk in Outer London off-setting each other.

Another key variable which proxies for relative congestion levels is our measure of proximate employment. This has a positive but weak level of statistical significance in the fatality model for the congested time period but is positive and significant for serious and slight injuries (Table 4). In the models of uncongested time periods only the injury models are clearly positive and significant (Table 5). In both cases the coefficient value is larger for serious injuries than for slight injuries, opposite of what we would expect. However, we also see that during the uncongested time period these coefficient values are larger than during the congested time period. This means that proximate employment (a VMT proxy) seems to result in more injuries during uncongested time periods than during congested time periods. This would seem to suggest that if this variable represents congestion that congestion may

lead to fewer accidents with serious and slight injuries. The same general trend holds in our pedestrian models (Tables 6 and 7) with a much stronger effect for pedestrian fatalities during uncongested time periods (at a 90% level of significance).

A third variable that could proxy for localized congestion within the ED's is our measure of employment density. These areas would presumably be more congested during the daytime and uncongested at night. Slight injuries appear less likely during congested time periods but during uncongested time periods coefficient values are all negative (Tables 4 and 5). Results are also indeterminate in the pedestrian models with conflicting results for the effects during congested time periods, with slight injuries being positive and significant while serious injuries have a negative value (Table 6).

Our results appear to be relatively indeterminate as to whether congestion mitigates the severity of traffic crashes. A clear conclusion could have been made if the Inner London dummy variable showed a negative and significant effect for fatalities during congested periods with this effect disappearing during uncongested time periods. A similar effect for the proximate employment variable would have led to a clearer conclusion. Clearly we do not have details on actual vehicle speeds and flows within each ED and this complicates interpretation of these results. There seems to be some weak evidence that during uncongested time periods the probability of injuries is greater than during congested time periods, but no evidence that the probability of fatalities increases.

#### *Interactions with road infrastructure*

Congestion effects could perhaps be mitigated by differences in the road infrastructure within Inner London compared to Outer London. An analysis of variance (ANOVA) confirms that all the independent variables in our analysis, with the exception of railway stations, have a different average value when Outer and Inner London ED's are compared (see Table 8). Inner London has a higher density of Underground stations and bus

stops, relative to Outer London. Inner London also has a higher density of road length (with the exception of Motorways) and a higher density of both three-leg and four-leg junctions, compared to Outer London. The density of roundabouts is higher in Outer London. Both employment density, population density and proximate employment are also higher in Inner London. The distance to the nearest hospital is higher in Outer London. More households in Inner London do not own cars compared to Outer London.

The effect of the infrastructure variables on casualties is of most interest as differences in these may be endogenous. That is, if these factors have an effect on safety, it is possible that changes have been made over time such that they mitigate adverse safety impacts associated with the traffic volumes in these areas. In general, we would expect that denser primary (A and B) road networks would lead to more vehicle interactions and more accidents. For the congested time period (Table 4) we see that increased A road density is positive and statistically significant. The coefficient for slight injuries is highest, but not much higher than that for fatalities. For the uncongested models (Table 5) the coefficient value is highest for the fatality model. These results suggest that while A road density is associated with higher fatalities and injuries, during uncongested time periods the effect seems to be associated more with fatalities than with injuries. Interestingly, in the pedestrian casualty models (Tables 6 and 7) while these effects are still positive (and in most cases statistically significant), parameter values are generally lower.

Results are somewhat different for B road density. For congested time periods, B road density is associated with more slight injuries. There is no significant effect for either uncongested time periods or in the pedestrian models.

Minor road density is generally associated with fewer casualties in all cases, except there is no association with fatalities during uncongested periods (for both total fatalities and pedestrian fatalities). The value of the coefficient is largest for the association with fatalities,

that is, there is a stronger effect on reducing fatalities in congested areas, relative to both injury categories, when minor road density is higher. This does not hold during uncongested time periods where the association is stronger for the injury models. Clearly, given that minor roads will tend to have lower speeds, those ED's with denser minor road networks will have fewer casualties of all types probably due to the lower average speeds on minor roads. Our results suggest that as the density of the road classes moves from minor to B road to A roads, we see an increasing association with all types of casualties.

Motorway density is found to be associated with more slight injuries during congested time periods. During uncongested time periods increased motorway density is also associated with more serious injuries. To some extent this supports the idea that less congestion increases the severity of accidents and we would expect to find this effect stronger in ED's with motorways. This seems to be the case, although motorway density does not appear to be associated with more fatalities. Pedestrian casualties are not associated with motorway density, not surprisingly, since pedestrians are prohibited from using motorways.

Junctions are normally considered relatively hazardous as this is where more vehicles interact and accidents are more likely. We find no statistically significant effect with respect to fatalities in any of the models. There are strong associations between junction density and injuries. Four-leg junction density appears to be positively associated with slight injuries during congested time periods, while three-leg junctions are positively associated with both serious and slight injuries. During uncongested time periods we find that four-leg junction density is positively associated with both serious and slight injuries while three-leg junction density is not. In the case of pedestrian casualties (Tables 6 and 7), we find that both three and four-leg junctions are associated with both serious and slight injuries during congested time periods. Four-leg junctions show no association with pedestrian injuries during

uncongested time periods while three-leg junctions do (opposite the effect found for all injuries).

These results would tend to support the assertion that increased potential vehicle interactions increases accidents. In any case, this result would seem to support the hypothesis that increased vehicle interactions leads to more accidents, but that these accidents are not necessarily fatal accidents.

We find the surprising result that roundabout density does not seem to be associated with fewer casualties but has a positive association with increased slight injuries. This occurs during both congested and uncongested time periods. This effect is not present for pedestrian injuries but during uncongested time periods there is a small (at the 80% confidence level) positive association with increased pedestrian fatalities. This latter result could be an indication of the risks associated with uncongested traffic.

The surprisingly positive association of roundabout density with slight injuries could actually be an indicator of the relative safety of roundabouts. Roundabout locations may be endogenous, in that they were constructed at junctions with high accident probabilities. There is no way to know if the positive association with slight injuries may be due to the inherent danger of the conditions at these junctions that have been mitigated by roundabout construction. In other words, if these were normal junctions, then perhaps they would be associated with serious injuries and fatalities.

The location of public transport stations would tend to be associated with high levels of pedestrian activity and relatively more congested traffic. This would be particularly true in the case of London Underground stations. We include a dummy variable for whether an ED has an Underground station and find a significant and positive association with all casualties during congested time periods (Table 4). This association also occurs during uncongested time periods and coefficient values are actually slightly higher, especially for fatalities (Table



5). Therefore, this result would tend to support our conclusion that congestion may mitigate the severity of accidents, although the differences are quite minor.

One possibility is that these effects are due to pedestrian traffic accessing the Underground station. Our pedestrian models show similar associations during congested time periods but no association with fatalities during uncongested time periods (Tables 6 and 7). Of course, pedestrian activity around the station would be very low when the stations are closed during nighttime and early morning hours.

The location of railway stations shows an association with serious and slight injuries for both congested and uncongested time periods and for both all injuries and just pedestrian injuries.

The density of bus stops was also included in our models. These were strongly associated with all classes of casualties, except pedestrian fatalities during uncongested time periods. These effects may be due to the presence of buses causing vehicle accidents but the results for the pedestrian models suggest that greater pedestrian activity around bus stops may lead to the association with more pedestrian casualties.

#### *Demographic variables*

We include various demographic variables in our model to control for other effects associated with casualties. We find that higher residential population density is strongly associated with reductions in casualties. This is consistent with the findings of Noland & Quddus (2003) in a spatial analyses of all of England. This effect appears to be strongest (i.e. highest negative coefficient value) for the association with fatalities (both in total and for pedestrians). This effect may be due to the lower speed limits in areas with higher population density.

We also find an association with the percent of households with no cars. That is, the fewer cars, the more casualties in all our models. Our car ownership data dates from 1991

while our casualty data is from 1999-2001, which may introduce some uncertainty in this result as car ownership in London continues to increase. However, we would expect those areas with lower car ownership in 1991 to still have lower car ownership in 1999-2001. This result possibly suggests some association of casualties with more deprived areas, as suggested by Graham & Glaister (in press) for pedestrian casualties.

Finally, we include a proxy variable to control for access to medical care. This is the distance to a hospital for each ED. We would expect greater access to a hospital to decrease the likelihood of a fatality but have less affect on injuries. We find no consistent pattern in our results. One possibility is that we do not have data on which hospitals have accident and emergency services and thus we may not be capturing the effect of fast access to medical care. In addition, ambulances may be dispatched from other locations, so actual time from the occurrence of an accident until the patient is transferred to a hospital may not be correlated with these distances.

## **Conclusions**

Our analysis attempted to determine if the hypothesis that traffic congestion may result in some safety benefits could be confirmed. For the data analyzed here, for the greater London metropolitan area, we find little evidence to support this hypothesis. Spatial differences between Inner and Outer London appear to be minor. Differences between our models for congested time periods and uncongested time periods, in general, do not lead us to any firm conclusions.

Some of our results suggest that there may be some small negative safety effects associated with reduced congestion. We found a small positive association between roundabout density and pedestrian fatalities. Roundabouts tend to be difficult for pedestrians to cross and during nighttime areas with higher roundabout density appear to be more hazardous. Exposure also appears to be a key confounding factor, especially for pedestrian

activity. We find that public transport stations tend to be associated with more casualties, but less so during uncongested time periods when pedestrian activity would be much less (i.e., when stations are closed).

The density of various road classes also shows some intriguing results. Areas with higher minor road density are associated with lower casualties, while those with higher A road density tend to be associated with more casualties. This may be due to higher traffic levels on A roads but may also be due to lower speeds on minor roads.

Our results for motorway density indicates some effects that might confirm the hypothesis of congestion being beneficial for safety. Results showed increased serious injuries during uncongested time periods, but no increased association with fatalities. This evidence is clearly weak, but does suggest that while these effects may not be occurring on urban roads, they may still be present on motorways or higher speed roads. Clearly, more research is needed into these effects to fully understand the interactions between congestion and road safety.

## References

- Dickerson, Andrew, Peirson, John, and Vickerman, Roger, 2000, Road Accidents and Traffic Flows: An Econometric Investigation, *Economica*, 67: 101-21
- 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*.
- Ivan, John N., Wang, Chunyan, Bernardo, Nelson R., 2000, Explaining two-lane highway crash rates using land use and hourly exposure, *Accident Analysis and Prevention*, 32: 787-795.
- Noland, Robert B., in press, "Medical Treatment and Traffic Fatality Reductions in Industrialized Countries", *Accident Analysis and Prevention*.
- Noland, Robert B., and Mohammed A. Quddus, in press, "Improvements in Medical Care and Technology and Reductions in Traffic-related Fatalities in Great Britain", *Accident Analysis and Prevention*.

Noland, Robert B., and Mohammed A. Quddus, 2003, A Spatially Disaggregate Analysis of Road Casualties in England, paper presented at the 82<sup>nd</sup> Annual Meeting of the Transportation Research Board and submitted to *Accident Analysis and Prevention*.

Shefer, Daniel and Rietveld, Piet, 1997, Congestion and Safety on Highways: Towards an Analytical Model, *Urban Studies*, 34(4): 679-692.

Washington, Simon P., Karlaftis, Matthew G., and Mannering, Fred L., 2003, *Statistical and Econometric Methods for Transportation Data Analysis*, Chapman & Hall / CRC: Boca Raton, FL.

Zhou, Min and Sisiopiku, Virginia, 1997, On the relationship between volume to capacity ratios and accident rates, *Transportation Research Record*, 1581: 47-52.

**Table 1: Breakdown of crash severity by congested time periods and congested locations**

<b>Variables</b>	<b>Congested time period</b>	<b>Uncongested time period</b>	<b>Inner London</b>	<b>Outer London</b>
<b>Total casualties</b>	82.2%	17.8%	46.9%	53.1%
<b>Fatalities</b>	69.3%	30.7%	43.0%	57.0%
<b>Serious injuries</b>	79.0%	21.0%	47.7%	52.3%
<b>Slight injuries</b>	82.7%	17.3%	46.9%	53.1%
<b>Pedestrian fatalities</b>	80.5%	19.5%	46.9%	53.1%
<b>Pedestrian serious injuries</b>	83.6%	16.4%	55.2%	44.8%
<b>Pedestrian slight injuries</b>	88.9%	11.1%	56.7%	43.3%

**Table 2: Summary statistics for enumeration district data, casualty data**

Summary statistics (N=15366)	Spatial unit: EDs			
	Mean	Std. Dev	Min	Max
Variable name				
Total fatalities	0.040	0.224	0	6
Total serious injuries	0.832	1.689	0	33
Total slight injuries	5.761	10.470	0	245
Congested fatalities	0.028	0.175	0	3
Congested serious injuries	0.658	1.378	0	29
Congested slight injuries	4.764	8.592	0	204
Uncongested fatalities	0.012	0.130	0	5
Uncongested serious injuries	0.174	0.557	0	11
Uncongested slight injuries	0.997	2.480	0	67
Pedestrian fatalities	0.019979	0.148069	0	3
Pedestrian serious injuries	0.246128	0.70289	0	20
Pedestrian slight injuries	1.010933	2.439881	0	67
Pedestrian congested fatalities	0.016075	0.13183	0	3
Pedestrian congested serious injuries	0.205714	0.601264	0	16
Pedestrian congested slight injuries	0.898347	2.133629	0	60
Pedestrian uncongested fatalities	0.003905	0.063403	0	2
Pedestrian uncongested serious injuries	0.040414	0.227596	0	4
Pedestrian uncongested slight injuries	0.112586	0.505565	0	23

**Table 3: Summary statistics for enumeration district data, independent variables**

Summary statistics (N=15366)	Spatial unit: EDs			
	Mean	Std. Dev	Min	Max
Variable name				
Proximate employment /1000	0.389	0.175	0.136	1.306
Employment per sq km of area /10000	0.439	1.261	0	31.000
Motorway length(km) per sq km of area	0.007	0.190	0	11.504
A road length (km) per sq km of area	0.798	2.066	0	60.013
B road length (km) per sq km of area	0.325	1.373	0	49.344
Minor road length (km) per sq km of area	3.416	4.397	0	97.754
Three-leg junctions per sq km of area/100	0.555	0.625	0	30.769
Four-leg junctions per sq km of area/100	0.095	0.236	0	6.667
Number of roundabouts per sq km of area	0.451	3.247	0	105.042
Underground stations	0.019	0.146	0	4.000
Railway stations	0.021	0.147	0	3.000
Bus stops per sq km of area	16.907	25.435	0	451.467
Resident population per sq km of area /1000	10.718	8.793	0	162.882
Distance to nearest hospital (m)/10000	0.233	0.121	0.003	0.754
Percent of households with no car	40.990	18.668	0	97.140

**Table 4: NB models for congested time period (7:00am to 8:30pm, weekdays)**

	Fatalities		Serious Injuries		Slight Injuries	
	Coef.	t-stat	Coef.	t-Stat	Coef.	t-stat
<b>Proxy variables for congestion</b>						
Indicator variable for Inner London	-0.0478	-0.30	0.0176	0.61	-0.0419	-1.37
Proximate employment /1000	0.7768	1.54	1.3681	2.63	1.0602	9.17
Employment per sq km of area /10000	-0.0298	-0.70	-0.0917	0.18	-0.0278	-2.31
<b>Infrastructure variables</b>						
Motorway length(km) per sq km of area	0.1403	1.00	0.3031	0.46	0.4494	5.97
A road length (km) per sq km of area	0.0958	4.59	0.0746	2.90	0.0960	16.19
B road length (km) per sq km of area	0.0343	0.78	0.0215	-0.87	0.0287	3.53
Minor road length (km) per sq km of area	-0.0652	-2.76	-0.0427	-7.24	-0.0436	-13.13
Three-leg junctions per sq km of area	0.0346	0.37	0.1152	4.15	0.0908	4.69
Four-leg junctions per sq km of area	-0.2774	-1.09	0.3051	0.54	0.3968	8.09
Number of roundabouts per sq km of area	0.0094	0.72	0.0080	0.69	0.0119	3.49
Underground stations (dummy variable)	0.4422	1.88	0.4436	2.45	0.5831	8.70
Railway stations (dummy variable)	-0.3022	-1.05	0.3348	3.25	0.3291	5.24
Bus stops per sq km of area	0.0103	5.79	0.0113	16.21	0.0127	28.28
<b>Demographic variables</b>						
Resident population per sq km of area /1000	-0.1394	-10.56	-0.1121	-19.35	-0.0938	-50.24
Distance to nearest hospital (m)/10000	-0.3617	-0.81	-0.0719	-2.57	-0.1144	-1.35
Percent of households with no car	0.0142	4.25	0.0115	14.41	0.0120	16.91
Constant	-3.3839	-15.42	-0.7725	-23.35	1.0984	24.76
Overdispersion parameter	1.5442	3.05	0.9763	26.78	1.0378	62.33
Number of observations	15336		15336		15336	
Log likelihood function at convergence	-1805.21		-15224.6		-36900.83	
Log likelihood ratio	0.078		0.093		0.075	



**Table 5: NB models for uncongested time period (8:30pm to 7:00am, weekdays)**

	Fatalities		Serious Injuries		Slight Injuries	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
<b>Proxy variables for congestion</b>						
Indicator variable for Inner London	0.3134	1.23	0.0874	1.17	-0.0113	-0.19
Proximate employment /1000	0.8486	0.92	1.6090	6.02	1.3043	6.01
Employment per sq km of area /10000	-0.1391	-1.46	-0.1319	-4.96	-0.0644	-2.82
<b>Infrastructure variables</b>						
Motorway length(km) per sq km of area	0.0584	0.18	0.2184	2.20	0.3355	2.66
A road length (km) per sq km of area	0.1111	2.66	0.0896	7.40	0.0953	8.57
B road length (km) per sq km of area	0.0566	0.71	0.0224	1.06	0.0218	1.37
Minor road length (km) per sq km of area	-0.0041	-0.11	-0.0599	-5.70	-0.0338	-5.41
Three-leg junctions per sq km of area	0.0857	0.58	0.0632	1.42	-0.0260	-0.72
Four-leg junctions per sq km of area	-0.1908	-0.38	0.2973	2.34	0.5164	5.52
Number of roundabouts per sq km of area	-0.0171	-0.52	0.0082	1.15	0.0124	1.92
Underground stations (dummy variable)	0.6815	1.76	0.6025	4.85	0.6185	4.97
Railway stations (dummy variable)	-0.1109	-0.26	0.3272	2.61	0.4350	3.73
Bus stops per sq km of area	0.0096	2.70	0.0094	9.55	0.0132	15.84
<b>Demographic variables</b>						
Resident population per sq km of area /1000	-0.2211	-8.65	-0.1172	-20.23	-0.0944	-26.32
Distance to nearest hospital (m)/10000	0.9533	1.33	-0.1619	-0.76	0.0797	0.50
Percent of households with no car	0.0163	2.89	0.0117	6.93	0.0128	9.45
Constant	-4.3965	-11.46	-2.0671	-18.97	-0.6189	-7.39
Overdispersion parameter	10.4010	4.03	1.9468	14.81	3.4481	40.81
Number of observations	15336		15336		15336	
Log likelihood function at convergence	-900.18		-6745.58		-17726.29	
Log likelihood ratio	0.084		0.081		0.050	

**Table 6: NB models for pedestrian casualty congested time period (7:00am to 8:30pm, weekdays)**

	Pedestrian Fatalities		Pedestrian Serious Injuries		Pedestrian Slight Injuries	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
<b>Proxy variables for congestion</b>						
Indicator variable for Inner London	-0.0681	-0.34	0.0363	0.58	0.0059	0.15
Proximate employment /1000	0.4535	0.68	1.0764	5.10	0.7925	5.57
Employment per sq km of area /10000	-0.0389	-0.63	-0.0354	-1.78	0.0373	2.72
<b>Infrastructure variables</b>						
Motorway length(km) per sq km of area	-61.3129	-0.16	0.0413	0.43	-0.0913	-1.10
A road length (km) per sq km of area	0.0697	2.57	0.0398	3.88	0.0471	6.78
B road length (km) per sq km of area	0.0556	1.18	-0.0023	-0.13	0.0122	1.15
Minor road length (km) per sq km of area	-0.0563	-1.95	-0.0415	-4.98	-0.0379	-7.93
Three-leg junctions per sq km of area	0.0464	0.39	0.1818	4.71	0.1937	7.65
Four-leg junctions per sq km of area	-0.0023	-0.01	0.1621	1.64	0.3426	5.42
Number of roundabouts per sq km of area	-0.0127	-0.57	0.0030	0.53	0.0018	0.45
Underground stations (dummy variable)	0.4812	1.62	0.1995	1.82	0.5722	7.48
Railway stations (dummy variable)	-0.1536	-0.43	0.3093	2.87	0.3594	4.79
Bus stops per sq km of area	0.0122	5.80	0.0125	16.14	0.0147	27.26
<b>Demographic variables</b>						
Resident population per sq km of area /1000	-0.1201	-7.55	-0.0987	-21.47	-0.0824	-31.65
Distance to nearest hospital (m)/10000	-0.1543	-0.26	-0.4901	-2.55	-0.1290	-1.10
Percent of households with no car	0.0212	4.80	0.0199	13.90	0.0211	22.75
Constant	-4.3345	-14.84	-2.2362	-24.04	-1.0923	-18.25
Overdispersion parameter	2.0039	2.11	0.9833	12.77	1.0085	31.88
Number of observations	15336		15336		15336	
Log likelihood function at convergence	-1177.33		-7539.43		-17491.20	
Log likelihood ratio	0.070		0.098		0.105	

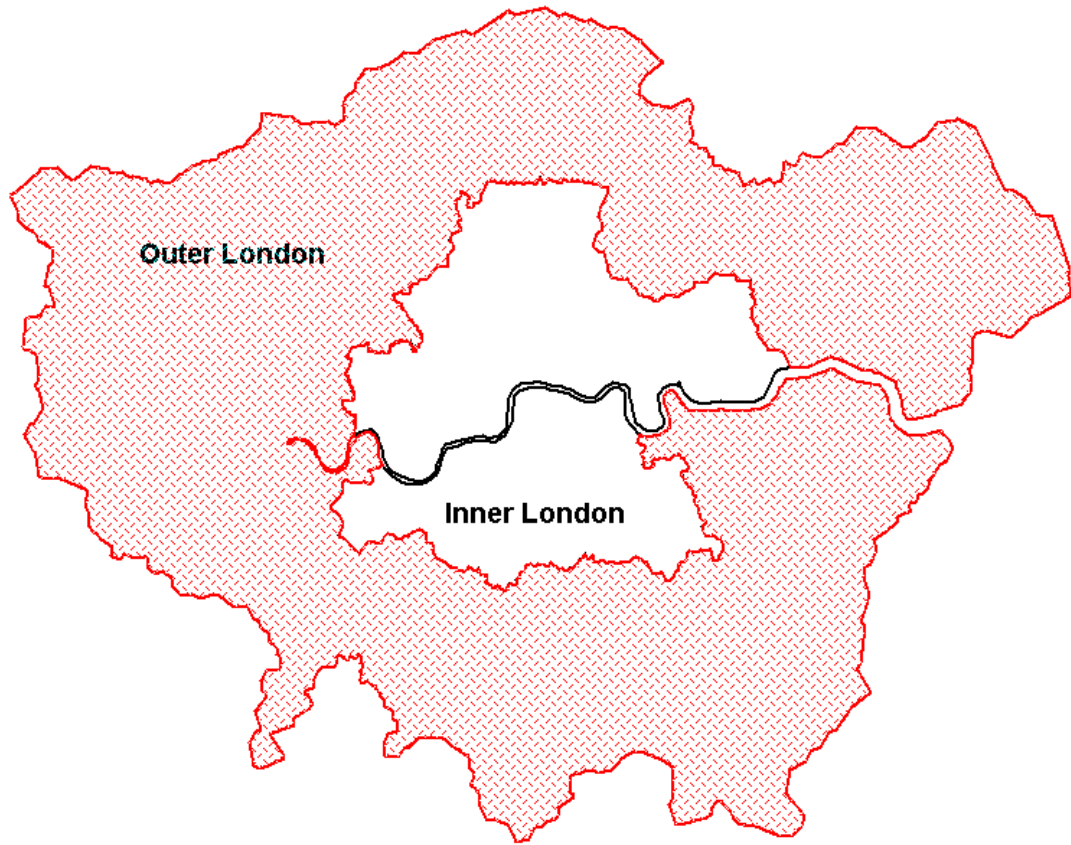
**Table 7: NB models for uncongested time period (8:30pm to 7:00am, weekdays)**

	Pedestrian Fatalities		Pedestrian Serious Injuries		Pedestrian Slight Injuries	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
<b>Proxy variables for congestion</b>						
Indicator variable for Inner London	-0.0204	-0.05	-0.0342	-0.26	0.0717	0.81
Proximate employment /1000	1.9018	1.65	1.7839	4.44	1.2151	4.18
Employment per sq km of area /10000	-0.0580	-0.54	-0.0390	-1.10	0.0122	0.46
<b>Infrastructure variables</b>						
Motorway length(km) per sq km of area	-0.2740	-0.20	-0.1548	-0.43	-0.2009	-0.73
A road length (km) per sq km of area	0.0727	1.56	0.0493	2.67	0.0809	6.49
B road length (km) per sq km of area	-0.1572	-0.86	0.0181	0.56	0.0066	0.29
Minor road length (km) per sq km of area	0.0112	0.25	-0.0458	-2.70	-0.0168	-1.66
Three-leg junctions per sq km of area	0.0742	0.61	0.1393	2.30	0.1448	2.75
Four-leg junctions per sq km of area	-0.0296	-0.06	0.2191	1.23	0.1587	1.19
Number of roundabouts per sq km of area	0.0272	1.40	0.0031	0.29	-0.0038	-0.44
Underground stations (dummy variable)	0.5785	1.07	0.7650	4.31	0.8913	6.55
Railway stations (dummy variable)	-0.5325	-0.67	0.7928	4.26	0.6812	4.78
Bus stops per sq km of area	0.0050	1.12	0.0108	7.80	0.0116	11.00
<b>Demographic variables</b>						
Resident population per sq km of area /1000	-0.1390	-4.55	-0.0823	-9.36	-0.0808	-14.05
Distance to nearest hospital (m)/10000	1.3686	1.15	-0.3575	-0.84	-0.2352	-0.83
Percent of households with no car	0.0282	3.33	0.0217	7.36	0.0238	11.58
Constant	-6.8929	-11.78	-4.4285	-22.56	-3.5147	-26.32
Overdispersion parameter	1.5048	0.56	1.7459	4.95	2.1629	11.73
Number of observations	15336		15336		15336	
Log likelihood function at convergence	-362.97		-2308.81		-4784.40	
Log likelihood ratio	0.074		0.106		0.105	

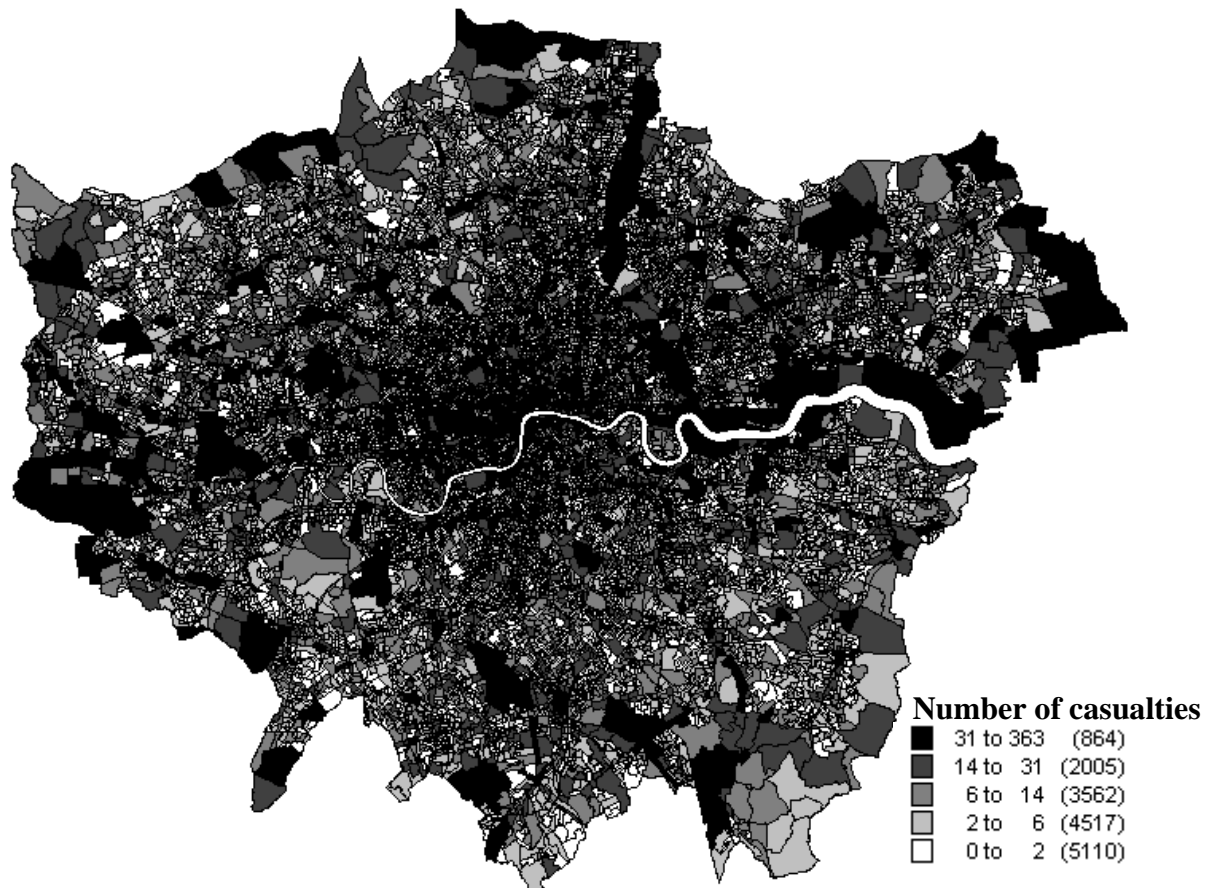
**Table 8: Analysis of variance of difference between independent variables for Inner London and Outer London**

<b>Analysis of variance (ANOVA)</b>				
<b>Experimental unit: Enumeration Districts (EDs)</b>				
<b>Treatments: Inner London (6586) and Outer London (8780)</b>				
<b>Factors</b>	<b>F statistic</b>	<b>p-value (prob&gt;F)</b>	<b>Average: Inner London</b>	<b>Average: Outer London</b>
Area (sq km)	422.86	0.0000	0.0485	0.1429
Proximate employment /1000	14607.03	0.0000	0.5300	0.2829
Employment per sq km of area /10000	83.03	0.0000	0.0329	0.0186
Motorway length(km) per sq km of area	16.74	0.0000	0.000	0.013
A road length (km) per sq km of area	158.02	0.0000	1.039	0.617
B road length (km) per sq km of area	190.81	0.0000	0.500	0.193
Minor road length (km) per sq km of area	816.39	0.0000	4.556	2.560
Three-leg junctions per sq km of area/100	893.25	0.0000	0.724	0.428
Four-leg junctions per sq km of area/100	507.26	0.0000	0.144	0.059
Number of roundabouts per sq km of area	7.78	0.0053	0.366	0.514
Underground stations	30.95	0.0000	0.0264	0.0132
Railway stations	1.89	0.1691	0.0191	0.0224
Bus stops per sq km of area	254.80	0.0000	20.658	14.093
Resident population per sq km of area /1000	3316.83	0.0000	14.9961	7.5094
Distance to nearest hospital (m)/10000	2648.96	0.0000	0.1793	0.2734
Percent of households with no car	6112.00	0.0000	52.4888	32.3648

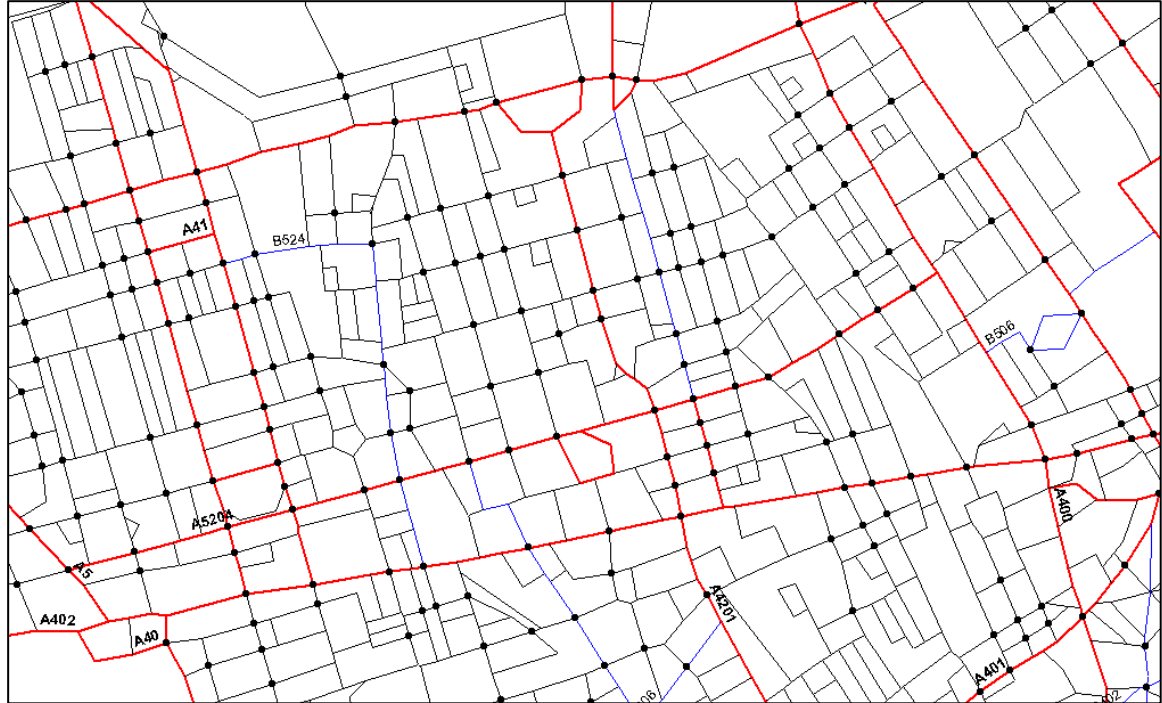
Note: p-value>0.1 implies that the 'mean' of factor (explanatory variables) at Inner London and Outer London are not statistically different at the 90% confidence level.



**Figure 1: Inner and Outer London as defined in our data**



**Figure 2: Spatial distribution of total traffic casualties for years 1999 to 2001 in Greater London by Enumeration District (ED)**



**Figure 3: MapBasic results for a part of central London identifying four-legged junctions**

**Figure 4: Underground and national railway stations within Greater London**

