Multilevel modeling for the regional effect of enforcement on road accidents

George Yannis[1] | Eleonora Papadimitriou | Constantinos Antoniou
Assistant Professor | Research Assistant | Research Associate

National Technical University of Athens
Department of Transportation Planning and Engineering

Abstract

This paper investigates the effect of the intensification of Police enforcement on the number of road accidents at national and regional level in Greece, focusing on one of the most important road safety violations: drinking-and-driving. Multilevel Negative binomial models are developed to describe the effect of the intensification of alcohol enforcement on the reduction of road accidents in different regions of Greece. Moreover, two approaches are explored as far as regional clustering is concerned; the first one concerns an ad-hoc geographical clustering, and the second one is based on the results of mathematical cluster analysis through demographic, transport and road safety characteristics. Results indicate that there are significant spatial dependences among road

---

[1] Corresponding Author: Dr. George Yannis
National Technical University of Athens
School of Civil Engineering
Department of Transportation Planning and Engineering
5, Iroon Polytechniou str., GR-15773 Athens, Greece
Tel: +30 210 7721326
Fax: +30 210 7721454
E-mail: geyannis@central.ntua.gr

accidents and enforcement. Additionally, it is shown that these dependences are more efficiently interpreted when regions are determined on the basis of qualitative similarities than on the basis of geographical adjacency.

Key-words: road accidents; alcohol enforcement; negative binomial models; multilevel analysis; cluster analysis.

1. Introduction

Road accidents and related casualties presented an increasing trend during the past decade in Greece, mainly due to insufficient maintenance of the road network, inappropriate behaviour of the road users and lack of efficient and systematic enforcement. Since 1998, an important effort was devoted to the improvement of this situation in Greece, expressed in an intensification of enforcement. More specifically, in 1998 the Greek Traffic Police started the intensification of road safety enforcement, having set as general target the gradual increase of roadside controls for the two most important infringements: speeding and drinking-and-driving, aiming to improve driver behaviour.

This road safety enforcement intensification could be one of the two basic reasons (the other one may be congestion) that may explain the important decrease observed in the number of road accidents, persons killed and injured during the last five years in Greece. Previous research on enforcement assessment has indicated that only a significant increase in enforcement level may affect the number of accidents (Bjørnskau, Elvik, 2003). Additionally, very little validation of enforcement effect at national level has been available in international literature. In particular, most evaluation attempts concern a temporary increase in local resources or concentrated enforcement efforts in a selected area (ESCAPE, 2003).

Several longitudinal studies on road accidents trends include the effect of Police enforcement. A recent research (Weiki, Slattoper, 2006) examined several explanatory variables to describe the decrease of road accidents in Ohio, USA during the period 1973-2000 and found that the introduction of a drink-driving arrests law and its enforcement had a significant, though short term, positive effect on injury accidents. A relevant research in the Netherlands (Mathjissen, 2005) showed that each intensification of alcohol enforcement within the period 1970-2000 resulted in a short term drink-driving decrease.

In general, intensification of road safety enforcement is considered to affect both drivers' behaviour and the road safety outcome, in terms of road accidents and casualties, those two parameters being examined separately in most relevant studies. Glodenbeld and Van Schagen (2005) report a significant decrease in road accidents and a significant modification of drivers speeding behaviour, as a result of an intensification of speeding enforcement on the rural network of a large area in the Netherlands. Chen et al. (2002) also report a significant effect of speed cameras on both drivers speeding behaviour and accidents occurrence. De Waard and Rootjiers (1994) evaluated different types and intensities of speeding enforcement and found a significant, preventive rather than repressive, effect on drivers' behaviour.

Several studies also focus on the possible variations of the effect of the various types of enforcement over time or in space. In a research investigating the effect of speed radars (Chen et al. 2002) showed that a significant effect was obtained not only on and around the radars locations, but also along the entire enforcement corridor. Hauer and Ahlin (1982) report both a "time halo" effect of enforcement and a spatial dispersion (upstream and downstream the enforcement sites). Vaa (1997) shows a significant effect of road section speed enforcement on drivers' behaviour, as well as an important temporal variation, with a specific decrease of the effect during the morning peak hours. Another relevant research (Tay, 2005) evaluating drinking-and-driving enforcement and publicity campaigns revealed a significantly higher effect on "high alcohol hours".

A research on drinking-and-driving among college students in the US (Wechsler et al. 2003) showed that the occurrence of drinking-and-driving differs significantly according to the policy environment at local and regional levels and the enforcement of those policies. Moreover, the results of a drinking-and-driving survey in Belgium (Vanlaar, 2005) are analyzed according to the geographical hierarchical structure of the survey, and significant differences in drinking-and-driving among different survey sites are found. Finally, Hakkert et al. (2001) evaluated the regional effect of a general road safety enforcement project in Israel, by means of a grouping of geographic zones according to Police enforcement intensity, and found a significantly different regional effect.

As far as Greece is concerned, the measures were implemented at national level, and a systematic enforcement covering most types of violations was achieved, resulting to a significant reduction of road accidents and related casualties, as shown in Table 1, which presents the basic road safety related trends in Greece.

Table 1 to be inserted here

However, because of the lack of specific quantitative targets in the intensification of enforcement, the increase of Police roadside controls was not carried out in a uniform or fixed way across the administrative regions of Greece, unlike most relevant enforcement projects (e.g. Goldenbeld, van Schagen, 2005, Hakkert et al., 2001). In particular, both the amount and the intensification rate of the enforcement activity presented significant differences in different regions; consequently, the resulting decrease in the number of road accidents could not be assessed under a common framework for all regions. Moreover, important differences among individual regions, such as demographic characteristics, transport systems infrastructure and operation and road safety related attitudes and behaviours further complicate the interpretation of the effects of Police enforcement. Summarizing, the efficiency assessment of a spatially varying Police activity over regions with significantly different characteristics is an interesting yet complex issue.

2. Objectives and methodology

The objective of this research is the quantification of the national and regional effect of Police enforcement on road safety in Greece. In particular, it is assumed that the effect of enforcement on road safety may depend on (either or both):

- the administrative structure of the Police, which follows a strictly geographic regional hierarchy, resulting to different practices and amounts of enforcement,
- the spatial variations of the socioeconomic and traffic characteristics, which reflect another hierarchy, both spatial and qualitative, resulting to different effects of enforcement.

On that purpose, a hierarchical modeling approach shall be used, in order to capture the spatial dependences rising from the hierarchical nature and structure of the parameters examined. In particular, a multilevel modeling technique is applied, allowing for the investigation and quantification of significant effects at all levels of the hierarchical structure (Rasbash et al, 2000).

In the present analysis, aggregate accident data are examined, and therefore a Poisson-family distributed response vector (O) of observed cases is assumed

(Lord et al, 2000), by introducing a log link function in the classical two-level model, as:

$O \sim \text{Poisson}\ (\pi_{ij})$

$\log (\pi_{ij}) = \log (E_{ij}) + \beta_{0j} + \beta_{1j}\, x_j + e_j$

$\beta_{0j} = \beta_0 + u_{0j}$

$\beta_{1j} = \beta_1 + u_{1j}$

where E_{ij} represents the expected numbers of cases for each level-1 unit (Rasbach et al., 2000).

The Poisson distribution is therefore used to model the (lowest) level-1 variance and random parameters at higher levels are assumed to be multivariate normal (Rasbash et al. 2001). An efficient estimation procedure for this non-linear model is predictive quasi-likelihood (Langford et al. 1998). However, a dispersion parameter at level-1 can be estimated, so that:

$\text{var}(\pi_{ij}) = \alpha\ \pi_{ij}$

If $\alpha = 1$, then variation is Poisson-distributed, if $\alpha > 1$ then there is extra-Poisson variation present, and if $\alpha < 1$ the model is underdispersed as can happen when many of the counts are zero. However, quite often there are theoretical reasons to assume that extra-Poisson variation may be present in the data (Dean, 1992, Hauer, 2001). For instance, this is the case when accidents counts come from significantly heterogeneous populations. This situation may be further described by a Negative Binomial distribution of counts, so that the variance of (O) is a quadratic function of (π_{ij}):

$O \sim$ Negative Binomial (π_{ij})
$\log(\pi_{ij}) = \beta_{0j} + \beta_{1j}\, x_j + e_j$
$\mathrm{var}(\pi_{ij}) = \pi_{ij} + \pi_{ij}^2 / v$

It should be noted that, ignoring extra-Poisson variation may not significantly affect parameter estimates; however the related statistical significances may be slightly affected (Dean, 1992).

Another issue concerns the higher level parameter estimates of multilevel models; as the data is used to support the parameter estimates, the residuals at higher levels are shrunken towards the overall mean, and therefore conditional estimates are produced. In order to overcome this effect and produce unconditional estimates, it is a common practice to reduce the number of levels in the model by introducing a categorical explanatory variable instead (Langford et al. 1998, Manton et al. 1987).

In the present research, negative binomial distributional assumptions were tested and are considered for the number of road accidents at the lowest (county) level of the hierarchical structure. Alcohol enforcement is selected as the most representative of the overall Police enforcement activity.

Moreover, two approaches are examined as far as the spatial hierarchy is concerned:

- An ad-hoc geographical hierarchy (counties nested within regions) based on the standard European NUTS (Nomenclature of Territorial Units for Statistics) classification of Greece, also reflecting the regional structure of the Greek Police,
- A qualitative hierarchy (counties nested within clusters) based on the results of mathematical cluster analysis of different spatially varying characteristics, reflecting the regional heterogeneity of Greece.

This type of multilevel modeling can be regarded as an aggregate spatial analysis, where the information on spatial variability is available in "zones" (i.e. counties).

3. Dataset

The dataset used in the framework of this research concerns aggregate data from 49 counties of Greece (245 observations in total) in the period 1998-2002. The response variable is the number of road accidents with casualties, and the considered explanatory variables include the number of alcohol controls, the number of speed infringements, as well as socioeconomic parameters such as population, vehicle ownership, road network type, fuel consumption and GDP (Gross Domestic Product). The dataset variables are summarized in Table 2.

Table 2 to be inserted here

These datasets are produced by the Greek Police and the National Statistics Office since long and their validity has been tested on several occasions, both by the control mechanisms of these authorities and the researchers using them. Accident and casualty under-reporting issues do not affect the validity of the dataset as a representative subset of the whole population. It should also be noted that the Athens and Thessaloniki metropolitan areas, where a disproportional high number of accidents and Police controls are observed, were not included in the dataset.

4. Results

4.1. Ad-hoc geographical clustering

In Figure 1, the official geographical classification of Greece is presented. In particular, 49 NUTS-3 counties (not including Athens and Thessaloniki) are nested within 12 NUTS-2 regions. In the following analysis, the assumption that geographical adjacency may be a determinant of the spatial effect of the intensification of alcohol enforcement, is tested.

Figure 1 to be inserted here

Table 3 shows the steps of the model building, and the related parameter estimates and respective standard errors. A null single-level model is initially estimated (Model 1), as a measure for the assessment of fit of the nested models. The first stage concerns a variance components model (Model 2), describing the regional dependences in the accident data, i.e. the random variation of accidents among regions, through a random intercept term. Then, the regional effect of alcohol controls is examined as the main explanatory variable, randomly varying among regions (Model 3). Finally, additional effects are incorporated in the model (Model 4).

Table 3 to be inserted here

From the results of the variance components model, a statistically significant variation of road accidents among different geographical regions is observed. As regards models fit, this model is significantly improved compared to the null model, with a residual deviance equal to 64.7 with 11 degrees of freedom. The variance of the log-accidents among counties (level-1) was calculated and was found equal to 0.541; consequently, it can be deduced that around 10% of the total variation of accidents counts is due to the regional classification, which adequately justifies the multilevel structure. However, that this is only a rough assessment of the variance partitioning; unlike the Normal case, here the level-1 variance depends on the expected value, therefore a simple variance partitioning among levels is not available (Goldstein et al., 2002).

An intuitive fixed effect of alcohol enforcement is obtained in Model 3, indicating a decrease of road accidents when the number of alcohol controls increases. Moreover, a statistically significant regional variation of this effect is obtained. Model fit is further improved compared to the variance components model, with residual deviance equal to 110.08 with 12 degrees of freedom. However, still the main part of the random variation is due to the geographical component of the model. This variation of the regional effect of alcohol enforcement is highlighted in Figure 2, where the different intercepts (constant term) and slopes (effect of alcohol enforcement) are plotted for each region. Additionally, Table 4 shows the fixed (unconditional) level-2 estimates of the effect of alcohol control.

After the incorporation of more explanatory variables in the model, the more detailed converging model is Model 4, including random effects of both alcohol controls and speed infringements, as well as a fixed effect of the percentage of national road. No random effect was found for this variable, which is not surprising, as the upgrade of the road infrastructure presented no impressive increase within counties in the examined period.

However, Model 4 including both the number of alcohol controls and the number of speed infringements results to lower and less significant regional variations. Moreover, the parameter estimates for each region presented some counter-intuitive results. This is probably due to the fact that both variables may be seen practically as measurements of one parameter (i.e. Police enforcement). The correlation between speed infringements and alcohol controls was examined, resulting to a positive correlation of 0.729. In this case (multicollinearity), a redundancy of variables is exposed, causing both logical and statistical problems (Washinghton et al. 2003). As far as multilevel models are concerned, the results of a recent study show that, with multicollinearity presented at level-1 of a two-level model, the fixed-effect parameter estimates produce relatively unbiased values; however, the variance and covariance estimates produce downwardly biased values (Shieh, Fouladi, 2003). According to the above, Model 4 is rejected against Model 3.

Table 4 to be inserted here

Figure 2 to be inserted here

The above modeling process provides some useful results, indicating a significant spatial variation of the effect of enforcement on road accidents. However, the interpretation of this variation is not straightforward. In particular, no consistent geographical pattern (e.g. higher effect in the North regions or in the most decentralized regions) is obtained. For instance, there would be no obvious explanation of the maximum effect of enforcement in regions 7 and 11, which are situated in the North West and East of Greece respectively and each one includes counties of different size and characteristics.

One could see an equivalence of this problem to what is known in spatial analysis as the "Modifiable Areal Unit Problem (MAUP)", which may occur when "zones" are arbitrary in nature and different spatial units (i.e. counties) could be just as meaningful in displaying the same base level data (i.e. accidents counts) (Openshaw, 1984). Although most spatial studies tend towards aggregating units which have adjacent geographical boundaries, it is possible (and often more meaningful) to aggregate spatial units which are spatially distinct.

4.2. Mathematical clustering

On the basis of the above, a different clustering approach was attempted. In particular, the 49 counties of Greece were clustered into groups on the basis of demographic, transport and road safety criteria, through a k-means algorithm.

A clustering (Q) defines the partitioning of the data in a set of clusters {Qi, i=1,2,…,C}. In a k-means type of clustering a popular approach of finding the optimum number of clusters is the validity index of Davies-Bouldin (Davies, Bouldin, 1979). According to the index, the optimum clustering is the one that minimizes the following quantity:

$$\frac{1}{C}\sum_{k=1}^{c}\max\left\{\frac{S_c(Q_k)+S_c(Q_l)}{d_{ce}(Q_k,Q_l)}\right\}$$

where C is the number of clusters, S_c is the mean within-cluster centroid distance, and d_{ce} the centroid distance between clusters. The smaller the value of the validity index, the better the clustering (Davies, Bouldin, 1979).

Data for year 1998 were used for the clustering, as this year reflected the beginning of the enforcement intensification period, but also according to data availability (i.e. some demographic and transport data were available only for 1998). A range of cluster numbers was tested, in accordance to the needs of multilevel analysis (i.e. minimum number of higher level units), but also in order to achieve an adequate number of within-cluster units. Moreover, a number of combinations of variables were examined for the clustering. These variables include socio-economic characteristics (population, gross domestic product), transport parameters (national road network length, vehicle ownership, fuel consumption), as well as road safety parameters. In particular, the number of speed infringements was also incorporated in the analysis, being representative of road safety attitudes and drivers behavior. Preliminary analysis indicated some correlations among the variables; consequently, only combinations of uncorrelated variables were examined.

Table 5 summarizes the results of the various cluster analyses. It is shown that the best clustering is obtained for eight clusters of counties, separated on the basis of vehicle ownership, fuel consumption and speed violations.

Table 5 to be inserted here

The final cluster centers are presented in Table 6. It is noted that clusters A, B, and C include only one county each. In particular:

- Cluster A concerns the county of Heraklion on the island of Crete, which is the 5th largest urban area of Greece, as well as a very popular tourist destination, characterized by the highest vehicle ownership and fuel consumption, as well as a high number of speed violations.
- Cluster B concerns the county of Achaia, which includes the 3rd largest urban area of Greece i.e. the city and port of Patras. This county presents a very high vehicle ownership and fuel consumption, but a medium number of speed violations.
- Cluster C concerns the county of Larissa in central Greece, which includes the 4th largest urban area of Greece. This county also presents high vehicle ownership and fuel consumption, and a medium number of speed violations.
- Cluster D includes nine counties, which are mainly located close to or including large urban areas, and present medium vehicle ownership and fuel consumption, but a disproportional high number of speed violations.
- Cluster E includes eight counties presenting medium vehicle ownership and fuel consumption (although lower than cluster D), as well as medium (although higher than clusters B and C) number of speed violations.
- Cluster F includes ten counties presenting the lowest vehicle ownership, medium (but lower compared to cluster E) fuel consumption and medium (but lower than clusters B and C) number of speed violations.
- Cluster G includes eleven counties presenting medium vehicle ownership, low fuel consumption and low number of speed violations.
- Cluster H includes eight counties presenting low vehicle ownership, low fuel consumption and low number of speed violations.

Table 6 to be inserted here

Figure 3 to be inserted here

Figure 3 highlights the results of the mathematical clustering of Greek counties. The next step concerns the building of a multilevel model, under negative binomial distributional assumptions, and under this new spatial hierarchy. Results are presented in Table 7. The variance components model (Model 5), although presenting a satisfactory overall fit (i.e. residual deviance equal to 35.5 for 7 residual degrees of freedom, compared to the null model), does not fully justify the multilevel structure, as a low and less significant level-2 variance is obtained. This can be attributed to the fact that three higher level units coincide with three lower level units (clusters A, B and C), weakening significantly the multilevel structure. Moreover, the clustering of counties itself being based on qualitative characteristics may have resulted in reduced within-cluster variability.

Adding the effect of alcohol controls, however, improves the model. More specifically, Model 6 presents a significantly improved fit compared to Model 5, with residual deviance equal to 83.65 for 8 degrees of freedom. Additionally, a significant random variation of the effect of alcohol enforcement is obtained. Furthermore, almost all fixed (unconditional) parameter estimates at the highest level are statistically significant, as shown in Table 8.

Table 7 to be inserted here

Table 8 to be inserted here

The variation of the effect of alcohol enforcement among clusters is presented in Figure 4, where the different intercepts (constant term) and slopes (effect of alcohol enforcement) are plotted for each cluster.

Figure 4 to be inserted here

First of all, it is interesting to note the non-significant effect of alcohol enforcement in cluster E. It is reminded that cluster E can be ranked as medium, given that it has average values in all examined characteristics, presenting thus no particularity.

A relatively low effect of alcohol enforcement is also obtained in cluster D, which presents the highest number of speed violations, all other characteristics having medium values. It can be deduced that the non-compliant road safety attitudes and more reckless driver behaviors, expressed by the very high number of speed violations, correspond to (or even result in) a limited effect of Police enforcement.

A relatively low effect is also observed in clusters A and B. These clusters are characterized by the highest vehicle ownership and fuel consumption, as well as an important number of speed violations. Furthermore, cluster B has somewhat fewer speed violations and somewhat higher effect of enforcement on road accidents. It can be said that in large urban areas, with increased level of motorization and mobility, a significant yet not impressive effect of enforcement is observed.

The highest effect of alcohol enforcement is obtained in clusters C, F, G and H. Clusters F, G and H concern the less urban and less active counties of Greece, as explained above. Moreover, these clusters present the lowest rates of speed violations. It is thereby demonstrated that a low level of motorization and a limited mobility, also implying a less experienced drivers' population, corresponds to a maximum effect of enforcement, due to a more compliant driving behavior.

The increased effect in cluster C, which includes a single county, would deserve some further discussion. This cluster may be considered to be quite similar to clusters A and B as far as the level of urbanization and motorization is concerned, and also presents similar speed violations rates. However, a lower mobility, as expressed by fuel consumption, results to an increased effect of enforcement. In practice, this reduced mobility can be explained when considering that cluster A (located on the popular island of Crete) and cluster B (including the port of Patras) accommodate a significant traffic during the summer, which is the main factor differencing cluster C from these two clusters.

From the above analysis it can be seen that the interpretation of results, although not always straightforward, can be based on specific determinants.

Accordingly, the magnitude of the effect of alcohol enforcement can be further attributed to behavioral parameters, reflected in the examined characteristics. Unlike the simple geographical clustering, the mathematical clustering allowed defining regions on the basis of some spatial homogeneity in demographic and transport characteristics and providing a comprehensive and meaningful ranking of the efficiency of enforcement.

5. Discussion

In the present research, the regional effect of enforcement on injury road accidents is examined, resulting from a nationwide intensification of Police activity in Greece in the period 1998-2002. Alcohol enforcement is selected among the various types of enforcement as representative of the overall Police activity, which also included speeding, seat belt and helmet enforcement.

Multilevel models were then developed, in order to capture the spatial variation of the effect of alcohol enforcement, which was found to be highly significant. It is noted that no other variables were found to add explanatory effect in the reduction of road accidents in Greece. This was not surprising, as no other parameter (e.g. vehicle ownership, road network length etc.) presented a significant variation, comparable to the increase of enforcement, in the examined period. Consequently, as in other studies (Goldenbeld, Van Schagen, 2005), the intensification of enforcement is considered to be the main cause of the improvement of road safety in Greece. Although additional explanatory variables (for which data was not available) could contribute to the description of this trend, the present models are efficient in describing the regional variation of this trend.

The results of the ad-hoc geographical clustering confirmed the significant regional variation of the effect of enforcement, however the interpretation of results was proved to be quite complex. It might be reasonable to assume that the regional variation of the effect is mainly a result of different regional practices in the implementation of enforcement. However, it was necessary to address the problem of the significant within-region heterogeneity, which affected the results.

The mathematical clustering based on demographic, transport and road safety parameters allowed for a more efficient interpretation of results and revealed some specific causal relationships. More specifically, a consistent effect was identified, according to which, the lower the number of speed infringements, the higher the effect of enforcement. Moreover, regions with varying levels of motorization and drivers exposure, but with very high number of speed infringements were found to have the lowest effect of enforcement.

It was shown that, the number of speed infringements reflects drivers' behavior and attitude towards road safety, as also suggested by previous research (SARTRE, 2004). According to the above, the prevalence of more risk-taking behavior of the drivers' population results in limited effect of enforcement.

The exposure of drivers was found to be another significant, yet secondary, determinant of the effect of enforcement. Lower mobility was found to correspond to increased effect of enforcement, often regardless of the level of motorization, confirming thus existing findings (SARTRE, 2004).

In general, it was shown that less urbanized and more road safety compliant regions are an easy target as far as enforcement is concerned. On the contrary, a more systematic effort would be required to achieve a more significant effect in the more urbanized and - consequently - less road safety compliant regions (offenders may believe that they can better escape controls in areas of more dense traffic and population).

The intensification of enforcement was proved to be efficient in reducing the number of road accidents at both national and regional level in Greece indicating thus that an efficient national road safety strategy should comprise enforcement as a primary component. Furthermore, it is indicated that enforcement strategies may be more efficient when targeting specific patterns and groups, allowing thus for the better adaptation of Police practices and the optimization of resources allocation for road safety enforcement.

References

Bjørnskau, T., Elvik, R., (1992). Can road traffic law enforcement permanently reduce the number of accidents? Accident Analysis & Prevention, Volume 24, Issue 5, Pages 507-520.

Chen, G., Meckle, W., Wilson, J. (2002). Speed and safety effect of photo radar enforcement on a highway corridor in British Columbia. Accident Analysis & Prevention, Volume 34, Issue 2, Pages 129-138.

Davies D.L., Bouldin D.W. (1979). A cluster separation measure. IEEE Transactions on Pattern Recognition and Machine Intelligence, 1, 224-227.

De Waard, D., Rooijers, T. (1994). An experimental study to evaluate the effectiveness of different methods and intensities of law enforcement on driving speed on motorways. Accident Analysis & Prevention, Volume 26, Issue 6, Pages 751-765.

Dean, C. (1992). Testing for overdispersion in Poisson and binomial regression models. J. Amer. Statist. Assoc. 87, 451-457.

ESCAPE Consortium (2003). Traffic enforcement in Europe: Effects, measures, needs and future. The “Escape” Project Final report, Contract Nº: RO-98-RS.3047, 4th RTD Framework Programme.

Goldenbeld, C., van Schagen, I. (2005). The effects of speed enforcement with mobile radar on speed and accidents: An evaluation study on rural roads in the Dutch province Friesland. Accident Analysis & Prevention, Volume 37, Issue 6, Pages 1135-1144.

Goldstein, H., Browne, W. and Rasbash, J. (2002). Partitioning variation in multilevel models. Understanding Statistics 2.

Hakkert, A. S., Gitelman, V., Cohen, A., Doveh, E., Umansky, T. (2001). The evaluation of effects on driver behavior and accidents of concentrated general enforcement on interurban roads in Israel. Accident Analysis & Prevention, Volume 33, Issue 1, Pages 43-63.

Hauer, E. (2001). Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. Accid. Anal. Prev. 33. p. 799-808.

Hauer, E., Ahlin, E.J., Bowser, J.S. (1982). Speed enforcement and speed choice. Accident Analysis & Prevention, Volume 14, Issue 4, Pages 267-278.

Langford, I, Bentham, G, McDonald, A. (1998). Multi-level modelling of geographically aggregated health data: a case study on malignant melanoma mortality and UV exposure in the European Community. Statistics in medicine, vol. 17, 41–57.

Langford, I.H., Leyland, A.H., Rashbach, j., Goldstein, H. (1999). Multilevel modelling of the geographical distribution of diseases. Applied Statistics 48 part 2, pp 253-268.

Lord, D., Washington, S., Ivan, J. (2005). Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accid. Anal. Prev. 37. p. 35-46.

Manton, K. G., Stallard, E., Woodbury, M. A., Riggan, W. B., Creason, J. P. and Mason, T. J. (1987). Statistically adjusted estimates of geographic mortality profiles. Journal of National Cancer Institute 78.5, 805-815.

Mathijssen, M.P.M.(2005). Drink driving policy and road safety in the Netherlands: a retrospective analysis. Transportation Research Part E: Logistics and Transportation Review, Volume 41, Issue 5, Pages 395-408.

Openshaw, S. (1984). The modifiable areal unit problem. Concepts and Techniques in Modern Geography 38, p. 41.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I., Lewis, T. (2000). A User’s Guide to MLwiN, Version 2.1c. Centre for Multilevel Modelling, Institute of Education, University of London, UK.

SARTRE Consortium (2004). European drivers and road risk. Part 1: Report on principal results. Part 2: Report on in-depth analyses. INRETS, Arcueil, 2004.

Shieh, Y, Fouladi, R. (2003). The Effect of Multicollinearity on Multilevel Modeling Parameter Estimates and Standard Errors. Educational and Psychological Measurement, Vol. 63, No. 6, 951-985.

Tay, R. (2005). Drink driving enforcement and publicity campaigns: are the policy recommendations sensitive to model specification? Accident Analysis & Prevention, Volume 37, Issue 2, Pages 259-266.

Vaa, T. (1997). Increased police enforcement: Effects on speed. Accident Analysis & Prevention, Volume 29, Issue 3, Pages 373-385.

Vanlaar, W. (2005). Drink driving in Belgium: results from the third and improved roadside survey. Accident Analysis and Prevention 37, 391–397.

Washington, S. P., Karlaftis, M. G., Mannering, F. L. (2003). Statistical and Econometric Models for Transportation Data Analysis. Chapman & Hall/CRC.

Wechsler, H., Eun Lee, J., Nelson, T.F., Lee, H. (2003). Drinking and driving among college students: The influence of alcohol-control policies. American Journal of Preventive Medicine, Volume 25, Issue 3, Pages 212-218.

Welki, A.M., Zlatoper, T.J. (2005). The impact of highway safety regulation enforcement activities on motor vehicle fatalities. Transportation Research Part E: Logistics and Transportation Review, Article In Press.

Table 1. Basic trends of road safety and enforcement in Greece (1998 - 2002)

	1998	1999	2000	2001	2002	5-year change
injury road accidents	24,819	24,231	23,127	19,710	16,852	-32%
persons killed	2,182	2,116	2,088	1,895	1,654	-24%
vehicles (x1000)	4,323	4,690	5,061	5,390	5,741	33%
speed infringements	92,122	97,947	175,075	316,451	418,421	354%
drink & drive infringements	13,996	17,665	30,507	49,464	48,947	250%
drink & drive controls	202,161	246,611	365,388	710,998	1,034,502	412%

Table 2. Variables and values considered in the analysis

Region	The NUTS-2 regions of Greece (1-12)
County	The NUTS-3 counties of Greece (1-49)
Accs	The annual number of accidents per county (1998-2002)
Alcontrol	The annual number of 1000 alcohol controls per county (1998-2002)
speedinf	The annual number of 1000 speed violations per county (1998-2002)
Pop	The annual population per county (1998-2002)
vehown	The annual number of vehicles per 1000 inhabitants per county (1998-2002)
natroad	The annual percentage of national road network per county (1998-2002)
Fuel	The fuel consumption per county (year 1998)
Gdp	The gross domestic product per county (year 1998)

Table 3. Negative binomial multilevel models for the regional effect of enforcement on road accidents - Ad hoc geographical clustering

	Model 1 (Null model)	Model 2 (variance components)	Model 3 (plus effect of alcohol controls)	Model 4 (plus various effects)
Fixed effects				
constant	-6.422 (0.027)	-6.477 (0.075)	-6.599 (0.098)	-6.583 (0.078)
alcontrols			-0.052 (0.013)	-0.032 (0.007)
Speedinf				-0.056 (0.020)
natroad				0.418 (0.174)
Random effects				
Level 2 - Regions				
σ_{u0}^2 (constant)		0.060 (0.027)	0.105 (0.046)	0.085 (0.039)
σ_{u1}^2 (alcontrols)			0.002 (0.001)	0.0004 (0.0002)
σ_{u2}^2 (speedinf)				0.003 (0.002)
σ_{u01}^2 (covariance)			0.009 (0.005)	0.004 (0.002)
σ_{u02}^2 (covariance)				0.008 (0.006)
σ_{u12}^2 (covariance)				0.0009 (0.0005)
Degrees of freedom	244	233	221	208
-2*loglikelihood	2,807.04	2,742.27	2,632.19	2,614.49

Table 4. Fixed parameter estimates and standard errors for Model 3

	Fixed estimate	**Standard error**
Region 1	-0.026	0.008
Region 2	0.002	0.005
Region 3	-0.011	0.005
Region 4	-0.097	0.030
Region 5	-0.052	0.028
Region 6	-0.069	0.017
Region 7	-0.180	0.051
Region 8	-0.053	0.009
Region 9	-0.053	0.011
Region 10	-0.029	0.011
Region 11	-0.123	0.038
Region 12	-0.014	0.003

Table 5. Davies-Bouldin index for different clusterings

	Number of clusters						
Variables	3	5	6	7	8	9	10
Vehown - natroad	0.68	0.80		0.81	0.64	0.66	0.62
Vehown - natroad - gdp	0.60	0.44	0.41	0.56	0.58	0.55	0.52
Vehown - fuel - speed	0.52	0.42	0.34	0.43	**0.33**	0.40	0.44
Vehown - fuel - speed - pop- natroad	0.55	0.57	0.43	0.45	0.47	0.52	0.44

Table 6. Final cluster centers

Cluster	**Vehicle ownership**	**Fuel consumption**	**Speed infringements**
A	0.240	96,679	1,930
B	0.181	84,938	1,578
C	0.188	74,495	1,581
D	0.160	49,840	2,563
E	0.169	34,175	1,763
F	0.117	25,991	1,292
G	0.164	15,891	822
H	0.138	6,075	253

Table 7. Negative binomial multilevel model for the number of accidents - Mathematical clustering

	Model 1 (null model)	Model 5 (variance components)	Model 6 (plus effect of alcohol controls)
Fixed effects			
constant	-6.422 (0.027)	-6.510 (0.071)	-6.562 (0.066)
alcontrols			-0.043 (0.013)
Random effects			
Level 2 - Clusters			
σ_{u0}^2 (constant)		0.033 (0.020)	0.027 (0.017)
σ_{u1}^2 (alcontrols)			0.0012 (0.0006)
σ_{u01}^2 (covariance)			0.0044 (0.0031)
Degrees of freedom	244	237	229
-2*loglikelihood	2,807.04	2,771.54	2,689.89

Table 8. Fixed parameter estimates (sorted descending) and standard errors for Model 6

	Fixed estimate	**Standard error**
Cluster H	-0.159	0.054
Cluster G	-0.115	0.019
Cluster C	-0.074	0.030
Cluster F	-0.045	0.009
Cluster B	-0.028	0.008
Cluster D	-0.013	0.004
Cluster A	-0.012	0.005
Cluster E	-0.002	0.006

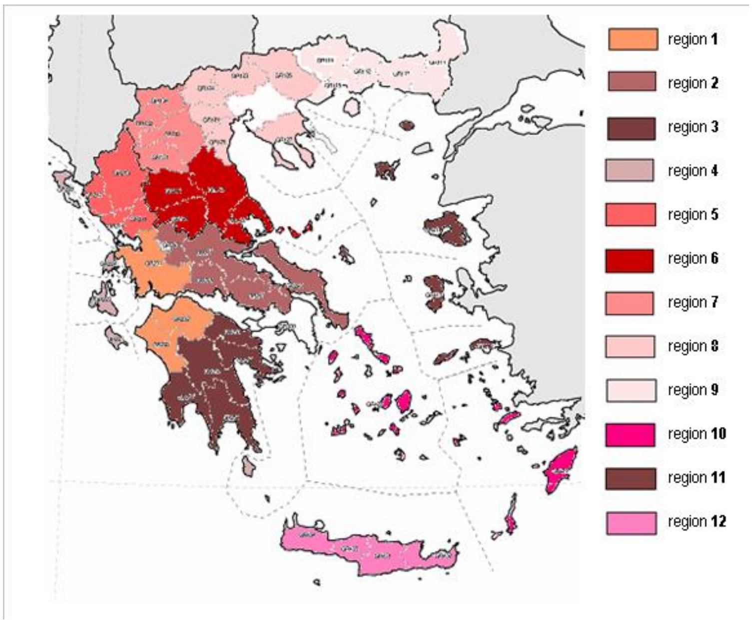

Figure 1. NUTS-2 classification of Greece
(Athens and Thessaloniki excluded)

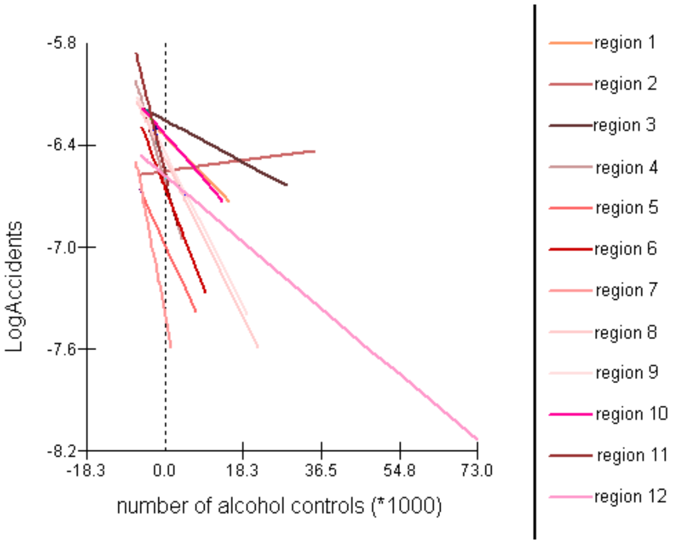

Figure 2. Regional effect of alcohol enforcement - Ad hoc geographical clustering
(Random intercepts and slopes of the negative binomial model)

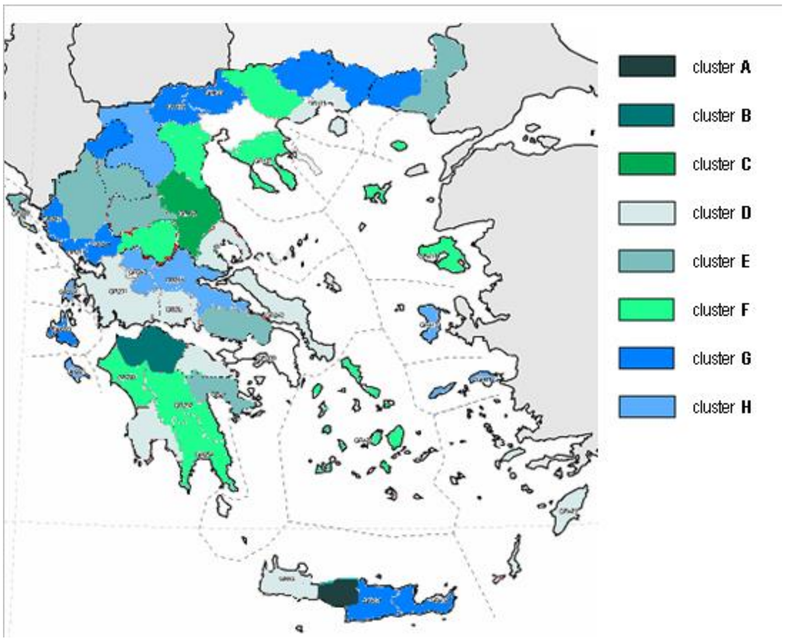

Figure 3. Mathematical clustering of Greece
(Athens and Thessaloniki excluded)

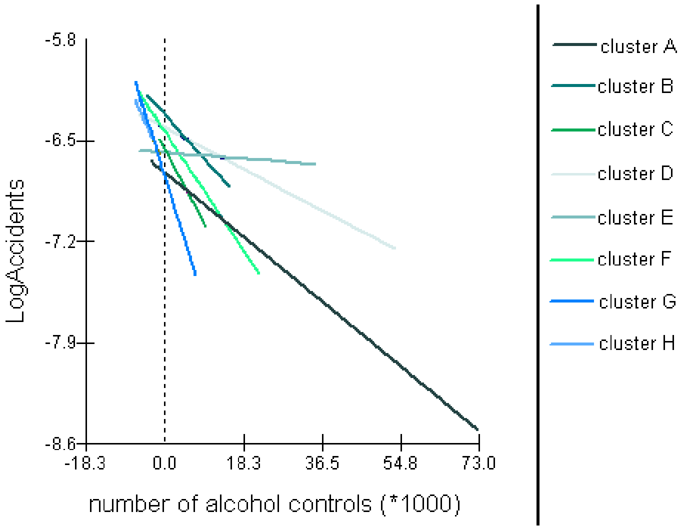

Figure 4. Regional effect of alcohol enforcement - Mathematical clustering
(Random intercepts and slopes of the negative binomial model)