Meta-regressions of exposure parameters used in spatial road safety analyses

Apostolos Ziakopoulos[1*] & George Yannis[1]

[1]*National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Heroon Polytechniou Str., GR-15773, Athens, Greece*
**Corresponding author, e-mail: apziak@central.ntua.gr*

Abstract

In road safety studies, spatial analyses are often implemented in order to reveal patterns of spatial variation and to examine the presence of spatial autocorrelation in road crash occurrences. Spatial analyses frequently employ exposure parameters in order to allow for comparisons between models. In this paper, meta-regression techniques are applied to three common exposure parameters (traffic volume/AADT, roadway length and vehicle distance traveled), to obtain quantitative estimates that several study characteristics impose on the values of their coefficients. Results indicate that the impact of traffic volume on crash counts was positively correlated with taking speed limit and road user age into consideration in spatial analyses, while the impact of road length on crash counts in spatial analyses was found to be higher in studies considering only fatal crashes. Finally, the impact of vehicle distance travelled on crash counts was found to be higher in county-level approaches as opposed to Traffic Analysis Zone-level approaches. The implication of findings is the quantification of coefficient discrepancies of exposure parameters between studies when including certain parameters as opposed to excluding them.

Keywords – Road safety, spatial analysis, meta-regression, exposure parameters, area units

1. Introduction

Despite considerable research and engineering effort, road crashes continue to incur considerable human and material losses to modern societies, with the rates of fatalities from road crashes remaining consistent during the previous decades [47]. In order to further suppress crash occurrence, it is critical to examine road crashes while taking into account as much information as possible, and proximity and relative position within the road layout, namely the dimension of space, are an important aspect of that information.

To that end, spatial analyses have been implemented in road safety quite frequently during the previous years. Spatial analyses provide insights by revealing crash hotspots and patterns of spatial variation, and by examining the presence of spatial autocorrelation in road crash occurrence. Spatial autocorrelation is the amount of influence that a value of one parameter in one location applies to the values of parameters at neighboring locations. Accordingly, a large variety of geo-spatial and statistical models have been developed in order to capture intricate patterns and to account for spatial effects, and their various intricacies have been explained in depth, e.g. [7, 26, 52].

As past literature has indicated, in order to establish a common baseline for crash risk comparisons between models, it is informative to include at least one exposure parameter [23]. The exposure variable can be either introduced to the statistical model as another independent variable or it can be already included in the data (for instance when analyzing crash rates normalized by vehicle-distance travelled instead of crash counts). When analyzing road crash frequencies, three of the most prevalent exposure parameters are traffic volume, vehicle-distance travelled and roadway length, though other variations have been also utilized, such as road network density (also per road type or speed limit) or trip generation.

The examination of studies that use explanatory parameters as independent variables can offer interesting insights in road safety spatial analyses, as influences on exposure variables can heavily skew study results. By taking influencing parameters into consideration at the process of study design and establishing a common framework, result transferability can be improved.

Apart from the classic modelling approach of independent/dependent variables, rate-based models have been also developed in the literature – [e.g. 10]. This approach incorporates exposure as an independent variable with its parameter estimate constrained to one. This is also achieved in count data models via the inclusion of an offset term, which has a parameter estimate constrained to one. The use of offset terms have been debated by researchers, with some support for the constraint of offset coefficients to one and others adopting a more unconstrained approach, as effects may be inelastic in some circumstances (for instance, congested conditions where crashes increase at decreasing rates with respect to traffic volume). As such, the current research focuses on classic modelling approaches.

This paper includes sections providing the methodological outline of meta-regression, and afterwards three sections are provided, one for each exposure parameter. Therein, a brief overview of the literature results for each exposure parameter is provided, to allow for a brief introduction for each parameter. Afterwards, the respective meta-regression results are shown. These results provide insights on which study characteristics influence the coefficient values of the exposure variables which in turn predict crash outcomes.

2. Meta-regression methodology

The aforementioned common framework can be established by the application of meta-analytic and meta-regression techniques to road safety studies with spatial analyses. The methodology of meta-analysis can be used to qualitatively combine the results of a number of input studies using the inverse-variance technique. A rich theoretical background for meta-analyses and applications in transport studies is available in the literature [8, 14, 17, 18, 35, 36, 38, 39, 51].

In the present research, a meta-analysis application was explored but ultimately was not found to be possible. This is due to the fact that a meta-analysis requires similar sampling frames, comparable methodologies (developed models etc.) and dependent variables. The examined studies display large dissimilarities in sampling frames, as they investigate different regions, and their dependent variables comprise a multitude of crash variations on several crash severities, as listed by Abdel-Aty et al. [1]. The most inhibiting factor, however, is considered to be the different methodologies/models that have been proposed in the literature, as spatial analyses is a fertile field for advanced statistical methods to be used, which unfortunately limits any aggregation attempts in a meta-analysis. Methodological differences can constitute reason for study exclusion as shown by Roshandel et al. [33].

These limitations can be circumvented in the case of meta-regression, which offers further explanation of the heterogeneity in the existing effects reported in the literature. A recent study by Elvik & Goel [13] underlines that meta-regression can be used to identify sources of differences in coefficient estimates of studies. The authors employed this method to identify factors that explain the large heterogeneity of estimates. They determined stronger safety-in-numbers effects for pedestrians than for motor vehicles and cyclists, and stronger safety-in-numbers effects at the macro level (e.g. a city) than at the micro level (e.g. in junctions).

In a meta-regression, effects such as study characteristics are assessed for their influence on coefficient estimates, as aggregated information can describe the differences between studies [38]. However, this method has the drawback of not providing direct model estimators, but rather outline the effects that influence existing the estimates of existing models. The inverse variance technique is utilized, which considers an overall estimate (summary mean, $\bar{Y}$) of effects based on n input estimates as proposed by Elvik [17]:

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i * W_i}{\sum_{i=1}^{n} W_i} \tag{1}$$

Where i is the number of input studies (i=1,2,…,n) and the statistical weights W_i are as follows (SE denotes standard error of a coefficient):

$$W = \frac{1}{SE_i^2} \tag{2}$$

In the present study, Y, which is the dependent variable, expresses the overall estimate of the coefficient of each exposure parameter. If we consider the Y_i as the observed effects in the i-th study, θ_i as the corresponding true effects and ε_i as the corresponding sampling error following a normal distribution, then:

$$Y_i = \theta_i + \varepsilon_i \tag{3}$$

The inverse variance technique allows two model specifications: (i) the fixed effects model and (ii) the random effects model. Fixed effects models provide results as an overall estimate of the included study sample, while random/mixed effects models assume the used sample of studies are a random part of a greater group of effects. In other words, the main target of fixed effects analysis is to provide a conditional estimate exclusively from those studies provided in the meta-analysis. On the contrary, mixed effects are considered as random samples of a greater set, therefore inferences made from them are unconditional [39].

For meta-regression, fixed effects models use the following structure:

$$\theta_i = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k} \tag{4}$$

Where β_i are potential influencers of the true effect of model coefficients θ_i and $x_{i,1}$ is the value of the independent variable $j, (j = 1, 2, \ldots, k)$ in the study i. In this case, the independent variables also known as moderator variables, are the different study characteristics of each study such as the areal unit of analysis.

Mixed effects models can better account for potential heterogeneity between studies, using the previous structure while adding a representative random effects term u_i ($u_i \sim N(\mu, \tau^2)$).

$$\theta_i = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k} + u_i \tag{5}$$

As Theofilatos et al. [36] state, τ^2 is the amount of residual heterogeneity (the variability among true effects that cannot be explained by the moderators entered in the meta-regression model). Obviously, if $\tau^2 = 0$, then the θ_i are homogenous and there is a reversion to a fixed effects model.

In order to determine the proper model specification, the Q-test is used to verify the presence of systematic variation between results provided by studies.

The Q-test for meta-analysis/regression is a form of Cochran's Q-test. The Q-test is a non-parametric statistical test to verify whether a number of factors have identical effects. In other words, the null hypothesis is that there is no systematic variance in the selected group of studies, and fixed-effects models can be used. As Elvik [16] states, the Q-test follows a chi-squared distribution with g-1 degrees of freedom, where g is the combined number of estimates, which are used to determine its significance. Furthermore, Viechtbauer [39] notes that the Q-test usually keeps better control of the Type I error rate and therefore should be preferred for hypothesis testing over likelihood ratio tests.

If the Q test is significant, the variance between studies is larger than would be expected on the basis of the within-study variation, and the use of mixed-effects models over fixed-effects models is warranted. The utility of the Q-test extends to within-study heterogeneity; namely the possibility that several of the effects reported in the same study are strongly heterogeneous with each other. In that case, random effects are included in the equations to allow for meta-regression.

$$Q = \sum_{i=1}^{n} W_i * Y_i^2 - \frac{(\sum_{i=1}^{n} W_i * Y_i)^2}{\sum_{i=1}^{n} W_i} \quad (6)$$

Lastly, a funnel plot can be used to visualize results of meta-regressions by showing the symmetry of the estimate value on the horizontal axis vs. the reported standard errors on the vertical axis, and can aid in detecting possible publication bias [14]. The term publication bias refers to the exclusion of relevant studies from meta-analyses, which reduces their robustness. These studies might have not been published or have counterintuitive effects [21].

On the processing part, meta-regressions in this study were conducted in R-studio using the metafor package and following Viechtbauer [39]. From the value of the t-statistic standard error values could be obtained, provided that the beta-coefficients are known, using the common conversion for regression testing:

$$t = \frac{\hat{\beta}_i - \beta_{i,0}}{se(\hat{\beta}_i)} = \frac{\hat{\beta}_i}{se(\hat{\beta}_i)} \quad (7)$$

A common reference framework was also established, transforming for common units of roadway length (miles), logarithmic and non-logarithmic estimates were transformed on the same scale and similar adjustments were made (e.g. AADT was examined per lane). If there were more than one suitable models reported from each study, only the one with the best reported fit was included (by assessment of AIC, AICc or similar indicators). This is because a particular participant should only contribute data once when calculating the observed outcomes [39]. Thus the funnel plots display the adjusted coefficient estimates plotted by the respective adjusted standard errors.

An equally important decision was the exclusion of results of studies conducting Bayesian modelling, as they utilize fundamentally non-frequentist approaches (posterior distributions, rather than parameter estimates, Bayesian credibility intervals rather than frequentist confidence intervals and so on). The meta-analysis and meta-regression methods considered do not currently offer robust ways of integrating Bayesian and non-Bayesian study results. This decision was supported by the process reported in the study of Roshandel et al. [33] as well.

Similarly to study assessment, meta-regression models with the lowest AICc values are considered to accrue minimum information losses and they are the ones selected. Meta-regression attempts were made on all detected studies for combinations of the moderator variables (study characteristics) that were reported in each case.

Therefore the established criteria for inclusion of a study in the meta-regression are:

1. The study is published in a scientific source (journal or conference, in English)
2. The examination of the considered parameter by the study with functional-econometric statistical models.
3. Correlation of the parameter with road crashes by the study (as opposed to injury severity). This is completed by the reporting of beta coefficient.
4. Reporting of the respective standard error in order to acquire the corresponding sampling variances (essential as per Viechtbauer [39]).

Furthermore, the present study is a quantitative continuation of a rigorous and extensive literature review of 132 road safety studies with a multitude of spatial approaches published and thoroughly explained in earlier research [52] – the reader is referred to that review paper for further details on study content and results.

3. Meta-regression on traffic volume/AADT estimators

In studies conducting spatial analyses, traffic volume (often used as AADT) has been found to be positively correlated with higher overall crash risk [41] and with higher non-local driver crash risk. Higher traffic volume has been found to be positively correlated with both severe and property damage only collisions [6]. Interestingly, AADT was found to be negatively correlated with local driver crash risk [40], with the authors of the study suggesting that local drivers cope better with higher driver conditions compared to foreign ones.

In intersections, Huang et al. [22] found that major road AADT positively contributes to crash occurrence at a significant level for motor vehicle, bicycle and pedestrian crashes (minor road AADT was found positive as well, but not statistically significant).

After examining the literature with the established 4 criteria, 4 spatial analysis studies contributed to the meta-regression for AADT with a total of 8 effects [3, 24, 44, 46]. The transformed coefficient values used as input for the meta-regression are provided on Figure 1.

The Q-test for Residual Heterogeneity was not found to be statistically significant ($Q_{[df = 6]}$ = 4.1614; p-value = 0.6548), suggesting no considerable heterogeneity among the true effects. Therefore, there is justification for using the fixed-effects meta-regression model. The outputs of the fixed-effects meta-regression appear on Table 1. Both study characteristics are treated as binary variables, based on whether they were considered in an input study or not, the latter being reference categories. For clarification, it is again noted that the estimates provided here are moderator variable (independent variable) impacts on regression coefficients and not the direct effect of AADT/traffic on crash occurrence.

The respective funnel plot is shown on Figure 2. The test for funnel plot asymmetry was not statistically significant (z = 1.9580; p-value = 0.0502), which suggests no indication of publication bias amongst the studies, though with a small statistical margin. Results indicate that from all considered study characteristics, the main moderator variables (study characteristics) affecting the overall estimate of traffic volume on crash occurrence are the examination of the presence of a speed limit and road user age. More specifically, the impact that AADT has on crash occurrence is increased if researchers consider the speed limits present in the study areas (as opposed to not considering them). A slightly higher impact of AADT on crash occurrence is found if researchers consider the age categories of road users present in the study areas (as opposed to not considering them).

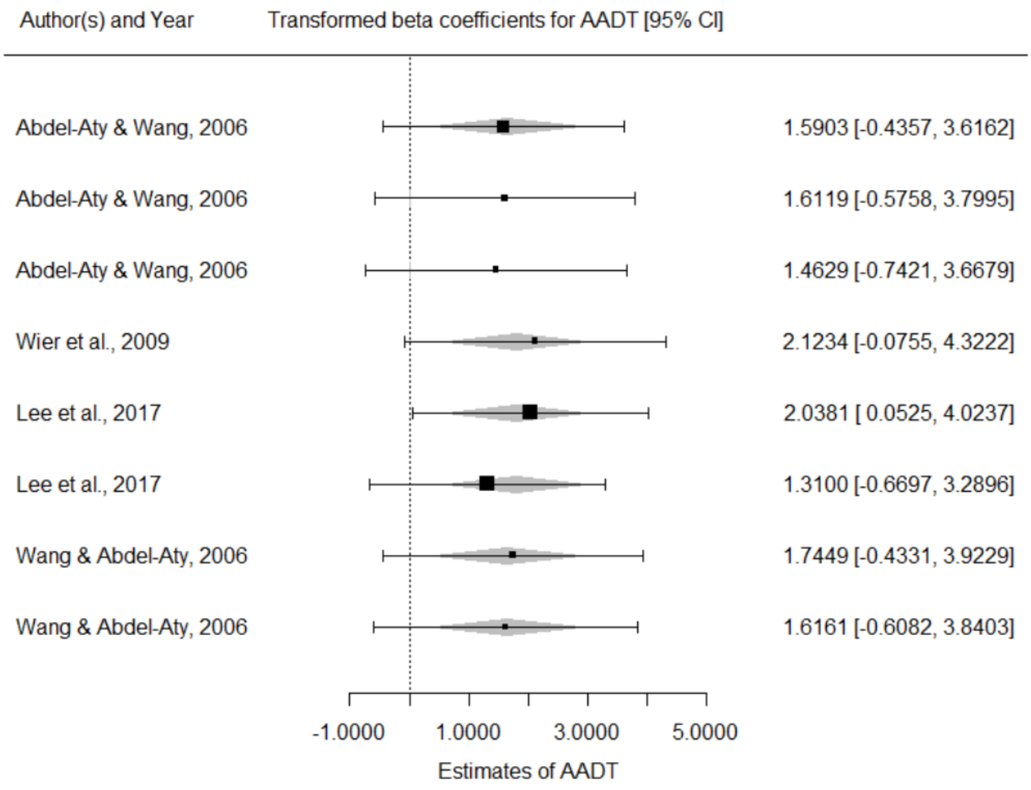


Figure 1: Funnel plot of beta coefficients of traffic volume on crash occurrence

Additional study characteristics that were considered but where not found to be statistically significant for AADT were the types of dependent variable (with categories: total crashes, pedestrian crashes and rear-end crashes), modal distinction (with categories: total crashes, motorized vehicle crashes, vulnerable road user-vehicle crashes and pedestrian-vehicle only crashes), regional approach (with categories: intersections and census tracts) and the examination of the number of lanes, all as defined in the respective studies.

Table 1: Parameter estimates of meta-regressions coefficients for the effect of traffic volume estimators on crash occurrence

Moderator Variable	**Estimate**	**Standard Error**	**p-value**
Speed Limit [ref. cat.: 0]	1.6479	0.6297	0.0089
Age [ref. cat.: 0]	1.8031	0.6031	0.0028

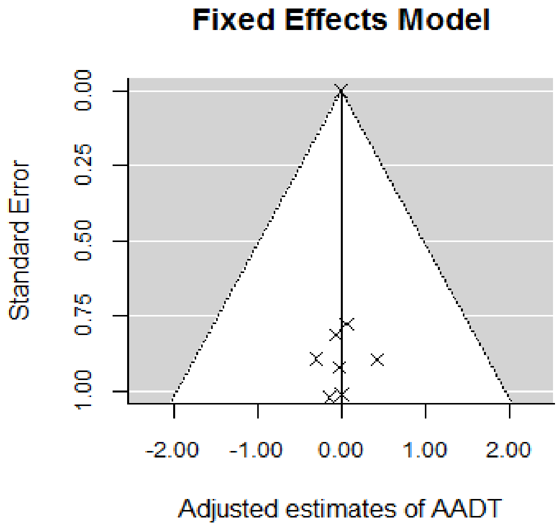


Figure 2: Funnel plot of beta coefficients of traffic volume on crash occurrence

4. Meta-regression on roadway length estimators

In studies conducting spatial analyses, roadway length is one of the most traditional exposure variables. Increased roadway length has been found to significantly and positively contribute to slight, serious and fatal crashes (consistently with different models) in segments of an English motorway [43]. Noland & Quddus et al. [28] reported that minor road length did not have an effect on serious injuries (and even decreased slight injuries), while it increased serious injury occurrence in roads of a higher category ("B" roads). Abdel-Aty et al. [2] developed spatial models for 1349 Traffic Analysis Zones (TAZs) in Florida and determined that roadway lengths with higher speed limits (e.g. 45 & 65 mph) were positively correlated with increased crash frequency and severity in general, while lower speed limits (e.g. 25 mph) were negatively associated with crash frequency during peak hours.

Considering vulnerable road users (VRUs), Nashad et al. [27] note an increase to crash likelihood involving VRUs if sidewalk lengths are increased in a zone, indicating a transfer of effect across transport modes. A similar result is reported in Wang & Kockelman et al. [42], albeit via a highly non-linear, two-stage relationship. Lastly, a study in Canada developed advanced urban models that revealed that bicycle-car collisions are directly associated with total lane and bicycle lane kilometers [45].

After examining the literature with the established 4 criteria, 7 spatial analysis studies contributed to the meta-regression for road length with a total of 29 effects [5, 10, 19, 20, 29, 31, 50]. Transformed coefficient values used as input for the meta-regression are provided on Figure 3.

The Q-test for Residual Heterogeneity was statistically significant ($Q_{[df = 20]}$ = 39.2066; p-value = 0.0063), suggesting considerable heterogeneity among the true effects. Therefore, there is justification for using the mixed-effects meta-regression model. The outputs of the mixed-effects meta-regression appear on Table 2. It is noted that the estimates provided here are moderator impacts on regression coefficients and not the direct effect of roadway length on crash occurrence. (KSI crashes: Killed and Serious injury crashes).

Figure 3: Funnel plot of beta coefficients of road length on crash occurrence

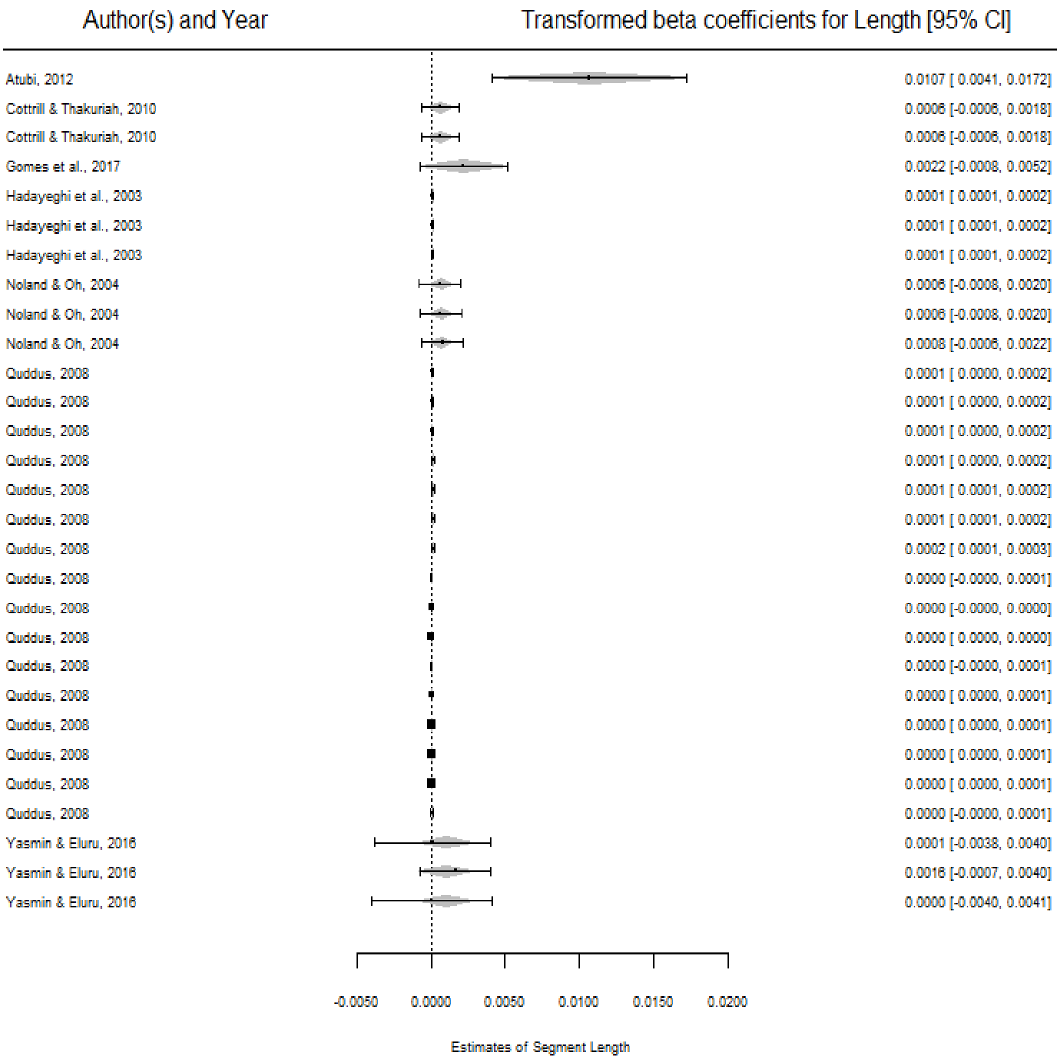


Table 2: Parameter estimates of meta-regressions coefficients for the effect of roadway length estimators on crash occurrence

Moderator Variable	**Estimate**	**Standard Error**	**p-value**
Constant term	0.0024	0.0015	0.1204
Unit of analysis: KSI crashes [ref. cat.: fatal crashes]	-0.0001	0.0000	0.0022
Unit of analysis: Serious injury crashes [ref. cat.: fatal crashes]	-0.0001	0.0000	0.0067
Unit of analysis: Slight injury crashes [ref. cat.: fatal crashes]	-0.0001	0.0000	0.0052
Unit of analysis: Total crashes [ref. cat.: fatal crashes]	-0.0001	0.0001	0.0127

The respective funnel plot is shown on Figure 4. The test for funnel plot asymmetry was not statistically significant (z = 3.9332; p-value <0.0001), which suggests no indication of publication bias amongst the studies.

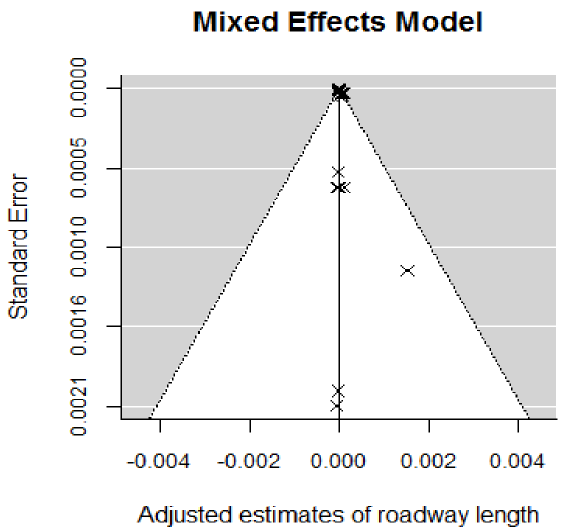


Figure 4: Funnel plot of beta coefficients of roadway length on crash occurrence

Results indicate that from all considered study characteristics, the main moderator variable (study characteristic) affecting the overall estimate of roadway length on crash occurrence is the severity of considered crashes (with fatal crashes as reference category). More specifically, analyzing crashes in any other severity level than fatal crashes decreases the impact of roadway length on crash occurrence in comparable levels across severities.

Additional study characteristics that were considered but where not found to be statistically significant for roadway length were traffic speed, road user age, modal distinction (with categories: total crashes, motorized vehicle crashes, VRU-vehicle crashes and pedestrian-vehicle only crashes), the presence of intersections, regional approach (with categories: County/State, TAZ/Census Tract (CT)/Census Ward) and road type (with categories: Aggregate, A-road, B-road, Motorways, Minor-road, Low-risk road), all as defined in the respective studies.

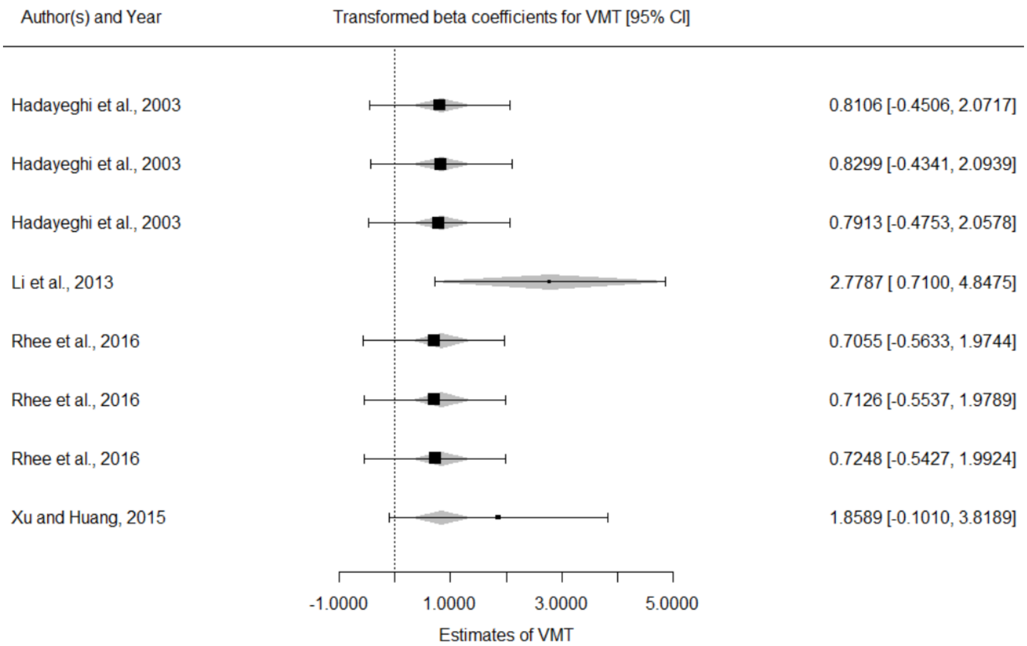
5. Meta-regression on vehicle distance travelled estimators

Vehicle distance travelled, usually expressed in miles (VMT) or kilometers (VKT), is another classic road safety exposure indicator. In studies conducting spatial analyses, some controversial results have been found for vehicle distance travelled. Lee et al. [24] conducted univariate and multivariate CAR analyses and found a positive correlation of VMT at a significant level for the occurrence of motor vehicle, bicycle and pedestrian crashes. An interesting result is reported by Cai et al. [7], who presented several crash models for at CT, TAZ and Traffic Analysis District (TAD) levels for Florida, USA, for total, severe and non-motorized crashes. VMT, when significant, was found to be positively correlated with crash occurrence. However, heavy vehicle mileage in VMT was found to reduce crash occurrence across all severity levels.

On the other hand, results of a spatial analysis in Florida, USA showed that the crash rate decreased with daily VMT increases, which the authors attribute to higher levels of traffic density and lower travel speed with higher daily VMT, or better maintained and safer road environments [12]; a similar result was reported by another study as well [4]. A study in Belgium showed that the crash occurrence contributions of non-motorway VKTs were more than twice of motorway VKTs for crashes between cars amongst all severity levels. However, the sign of motorway VKTs was reversed when examining crashes between cars and VRUs, which can be explained from lack of intermodal interaction [30].

After examining the literature with the established 4 criteria, 7 spatial analysis studies contributed to the meta-regression for road length with a total of 8 effects [20, 25, 32, 49]. It was decided to perform a meta-regression on a VMT level. Transformed coefficient values used as input for the meta-regression are provided on Figure 5.

Figure 5: Funnel plot of beta coefficients of VMT on crash occurrence



The Q-test for Residual Heterogeneity was not found to be statistically significant ($Q_{[df = 6]}$ = 1.1594; p-value = 0.9788), suggesting no considerable heterogeneity among the true effects. Therefore, there is justification for using the fixed-effects meta-regression model. The outputs of the fixed-effects meta-regression appear on Table 3.

Table 3: Parameter estimates of meta-regressions coefficients for the effect of VMT estimators on crash occurrence

Moderator Variable	Estimate	Standard Error	p-value
Constant term	2.7787	1.0555	0.0085
Regional approach: TAZ [ref. cat.: County]	-1.9449	1.0858	0.0733

The respective funnel plot is shown on Figure 6. The test for funnel plot asymmetry was statistically significant (z = 1.0579; p-value = 0.2901), which suggests a degree of publication bias amongst the studies. Results that are published cause the plot to appear asymmetrical; consequently there are results from studies that are unpublished, missing or have counterintuitive effects [21]. The trim-and-fill method cannot be applied to improve meta-regression models with moderators [36, 39].

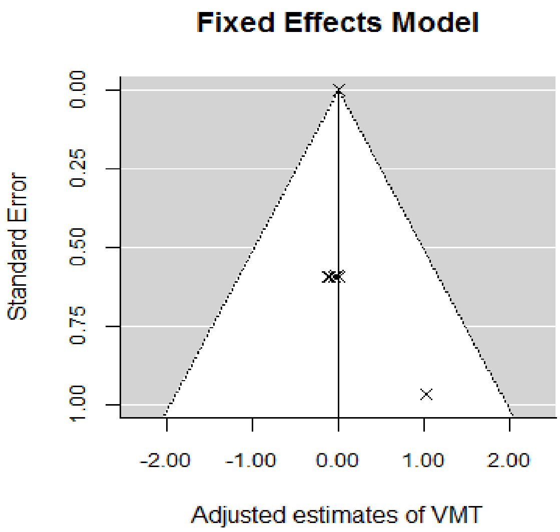


Figure 6: Funnel plot of beta coefficients of VMT on crash occurrence

Results indicate that from all considered study characteristics, the main moderator variable (study characteristic) affecting the overall estimate of the effect of VMT estimators on crash occurrence is the level of regional approach. More specifically, the impact that VMT has on crash occurrence decreases in studies considering a TAZ-level approach as contrasted to a county-level approach. This finding is an indicator of how the levels of units when conducting spatial analysis might influence the final outcomes, though more studies are needed to verify it. Its p-value lies between the thresholds of 0.05 and 0.10, and the correlation is also largely described by the constant term.

This is a particularly interesting result, because it hints at the effect of the modifiable areal unit problem (MAUP). MAUP occurs when boundaries are changed inside the study areas, affecting the coefficients of statistical models, a problem that has been manifesting and studied in road safety as well [1, 37, 48] The manifestation of MAUP in VMT analysis indicates a particular sensitivity of the results obtained for this parameter based on the size of the boundary area for the study. In areas with high VMT influences on crashes, such as urban road networks, these effects may be exacerbated. Therefore researchers and authorities need to establish common zonal frameworks before comparing results of spatial analysis from different areas.

Additional study characteristics that were considered but where not found to be statistically significant for VMT were the presence of intersections, regional approach (with categories: County, TAZ), unit of analysis (with categories: Total, Fatal or Serious crashes) and road user age, all as defined in the respective studies.

6. Discussion

The meta-regression approach followed in this research indicated evidence of how the various study characteristics can affect the calculated coefficient values, especially for exposure variables. The impact of traffic volume on crash counts was found to be positively correlated with taking speed limit and road user age into consideration in spatial analyses, while the impact of road length on crash counts in spatial analyses was found to be higher in studies considering only fatal crashes. Similarly, the impact of vehicle distance travelled on crash counts was found to be more important in county-level approaches as opposed to TAZ-level approaches, indicating vehicle distance travelled as more prone to statistical bias from fluctuations in boundary definition.

The findings presented in the present study are meant as an indicator of how study design affects outcomes of spatial analyses. It is not suggested that researchers who do not take into account, speed limits, for example, in spatial road safety analyses produce biased results. Rather, the implication of these findings is the quantification of those discrepancies between the studies that do include speed limits and the studies that do not, as well as an identification of the most significant factors that cause the discrepancy. The current results might serve to bridge differences between outcomes of studies of different designs in the future.

Apart from the academic exercise, this research has value for practitioners and authorities as well. Road management authorities commission several road safety studies in order to detect problematic spots. Road safety measures have different costs and effectiveness [11], thus the allocation of limited administrative funds requires scientifically informed decisions to achieve the maximum possible benefits. This research highlights the circumstances under which estimates change, and therefore aid in prioritizing road safety measures and interventions. As an example, results imply that traffic management measures aimed to reduce AADT will have increased impacts in reducing crash frequency an area determined as hotspot without accounting for speed limits, all other parameters held constant.

Naturally, the followed approach has some limitations. Literature findings indicate that Bayesian models frequently appear to offer more precise predictions and to have better performance overall. As discussed previously, the meta-analysis and meta-regression methods considered do not currently offer proper ways of including Bayesian study results. However, since the meta-regression results provided herein are indicative of the effects of study design and environment in spatial exposure parameters, they can provide insights to future studies, even in Bayesian frameworks, for instance by influencing priors.

Furthermore, models may consider exposure parameters on crash frequencies but not the influence of all other variables, which then become unobserved factors. Elvik [16] mentions that if studies do not account for all contributing factors, estimates of risk likely reflected not just the examined risk factor but the confounding ones as well. Such differences, along with differences in sampling frame, were an added reason why no meta-analysis was possible in this study.

However, in spatial analyses, the majority of unobserved heterogeneity can be accounted for: this is achieved by introducing spatially structured and unstructured random effects in the spatial models. Furthermore, by conducting the Q-test, bias related to unpublished/counterintuitive results is detected and reported (which was the case for vehicle distance travelled). Therefore, the authors are confident that while imperfect, the results of these paper are fruitful, especially regarding their qualitative aspect: there is clear indication, for instance, that accounting for speed limit alters predictions for the influence of AADT on crash occurrence. Realistically, no study will ever flawlessly account for all confounding factors, at least in the foreseeable future, let alone a sufficient number of similar enough studies to conduct a meta-analysis.

The number of studies included in the meta-regressions is indeed small. This is expected as specialization of each study increases, and authors pursue innovative designs and results. In the science of epidemiology, Terrin et al. [34] mention that more than 50% of meta-analyses include 10 or fewer studies. Elvik [16] who has addressed the issue distinguishes two groups: precise (or reliable) studies and imprecise (or unreliable studies). Based on the criteria set in that paper, the studies that are included here are precise/reliable. The funnel plots and tests for funnel plot asymmetry were conducted to test for publication bias; no indication of publication bias was found amongst the studies. The funnel plot symmetry and the overall convergence of more effects towards the central axis further support the reliability of the studies [34], despite their small numbers. Therefore the authors retain the produced results as informative.

Certain research directions can be derived from the meta-regression analyses provided in the present research. The meta-regression results (and the fact that not enough studies were found for a meta-analysis) is an indicator of the strong diversity present in published studies. More studies can contribute to the current knowledge and state-of-the art and further hone the results presented here (especially by consistently reporting standard errors). Dedicated studies that utilize case-control or cross-sectional designs assessing road safety interventions or other measures that extend to the parameters can be used to clarify why the factors found significant in this study influence the exposure parameter coefficients. Additionally, more research is needed to produce studies that do not examine crash frequency, but crash injury severity, and to determine if the influence caused by the same parameters persists for that dependent variable as well.

As previously stated, by establishing a common framework, result transferability can be improved. An important expectation is the introduction of an assessment method that would be akin to meta-regression, for Bayesian studies, which, to the authors' knowledge, is yet to be presented. Furthermore, it is worth noting that as we enter the era of big data is that the large information available to transport and road safety spatial analyses is expected to lead to the convergence of Bayesian and frequentist parameters.

7. Acknowledgements

This research is co-financed by Greece and the European Union (European Social Fund - ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research" (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

References

1. Abdel-Aty, M., Lee, J., Siddiqui, C., & Choi, K. (2013). Geographical unit based analysis in the context of transportation safety planning. Transportation Research Part A: Policy and Practice, 49, 62-75.
2. Abdel-Aty, M., Siddiqui, C., Huang, H., & Wang, X. (2011). Integrating trip and roadway characteristics to manage safety in traffic analysis zones. Transportation Research Record, 2213(1), 20-28.
3. Abdel-Aty, M., & Wang, X. (2006). Crash estimation at signalized intersections along corridors: analyzing spatial effect and identifying significant factors. Transportation Research Record: Journal of the Transportation Research Board, (1953), 98-111.
4. Aguero-Valverde, J., & Jovanis, P. P. (2006). Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis & Prevention, 38(3), 618-625.
5. Atubi, A. O. (2012). Determinants of road traffic accident occurrences in Lagos State: Some lessons for Nigeria. International Journal of Humanities and Social Science, 2(6), 252-259.
6. Barua, S., El-Basyouny, K., & Islam, M. T. (2014). A full Bayesian multivariate count data model of collision severity with spatial correlation. Analytic Methods in Accident Research, 3, 28-43.
7. Cai, Q., Abdel-Aty, M., Lee, J., & Eluru, N. (2017). Comparative analysis of zonal systems for macro-level crash modeling. Journal of safety research, 61, 157-166.
8. Caird, J. K., Johnston, K. A., Willness, C. R., & Asbridge, M. (2014). The use of meta-analysis or research synthesis to combine driving simulation or naturalistic study results on driver distraction. Journal of safety research, 49, 91-e1.
9. Copas, J., & Shi, J. Q. (2000). Meta-analysis, funnel plots and sensitivity analysis. Biostatistics, 1(3), 247-262.
10. Cottrill, C. D., & Thakuriah, P. V. (2010). Evaluating pedestrian crashes in areas with high low-income or minority populations. Accident Analysis & Prevention, 42(6), 1718-1728.
11. Daniels, S., Martensen, H., Schoeters, A., Van den Berghe, W., Papadimitriou, E., Ziakopoulos, A., ... & Weijermars, W. (2019). A systematic cost-benefit analysis of 29 road safety measures. Accident Analysis & Prevention, 133, 105292.
12. Dong, N., Huang, H., Xu, P., Ding, Z., & Wang, D. (2014). Evaluating spatial-proximity structures in crash prediction models at the level of traffic analysis zones. Transportation Research Record: Journal of the Transportation Research Board, (2432), 46-52.
13. Elvik, R., & Goel, R. (2019). Safety-in-numbers: An updated meta-analysis of estimates. Accident Analysis & Prevention, 129, 136-147.
14. Elvik, R., & Bjørnskau, T. (2017). Safety-in-numbers: a systematic review and meta-analysis of evidence. Safety science, 92, 274-282.
15. Elvik, R. (2015). Some implications of an event-based definition of exposure to the risk of road accident. Accident Analysis & Prevention, 76, 15-24.
16. Elvik, R. (2011). Effects of mobile phone use on accident risk: Problems of meta-analysis when studies are few and bad. Transportation research record, 2236(1), 20-26.
17. Elvik, R. (2005). Introductory guide to systematic reviews and meta-analysis. Transportation Research Record: Journal of the Transportation Research Board, (1908), 230-235.
18. Elvik, R. (2001). Area-wide urban traffic calming schemes: a meta-analysis of safety effects. Accident Analysis & Prevention, 33(3), 327-336.
19. Gomes, M. J. T. L., Cunto, F., & da Silva, A. R. (2017). Geographically weighted negative binomial regression applied to zonal level safety performance models. Accident Analysis & Prevention, 106, 254-261.
20. Hadayeghi, A., Shalaby, A., & Persaud, B. (2003). Macrolevel accident prediction models for evaluating safety of urban transportation systems. Transportation Research Record: Journal of the Transportation Research Board, (1840), 87-95.
21. Høye, A., & Elvik, R. (2010). Publication Bias in Road Safety Evaluation: How Can It Be Detected and How Common Is It?. Transportation Research Record: Journal of the Transportation Research Board, (2147), 1-8.

22. Huang, H., Zhou, H., Wang, J., Chang, F., & Ma, M. (2017). A multivariate spatial model of crash frequency by transportation modes for urban intersections. Analytic methods in accident research, 14, 10-21.
23. Imprialou, M. I. M., Quddus, M., Pitfield, D. E., & Lord, D. (2016). Re-visiting crash–speed relationships: A new perspective in crash modelling. Accident Analysis & Prevention, 86, 173-185.
24. Lee, J., Abdel-Aty, M., & Cai, Q. (2017). Intersection crash prediction modeling with macro-level data from various geographic units. Accident Analysis & Prevention, 102, 213-226.
25. Li, Z., Wang, W., Liu, P., Bigham, J. M., & Ragland, D. R. (2013). Using geographically weighted Poisson regression for county-level crash modeling in California. Safety science, 58, 89-97.
26. Loo, B. P., & Anderson, T. K. (2015). Spatial Analysis Methods of Road Traffic Collisions. CRC Press.
27. Nashad, T., Yasmin, S., Eluru, N., Lee, J., & Abdel-Aty, M. A. (2016). Joint modeling of pedestrian and bicycle crashes: copula-based approach. Transportation Research Record: Journal of the Transportation Research Board, (2601), 119-127.
28. Noland, R. B., & Quddus, M. A. (2005). Congestion and safety: A spatial analysis of London. Transportation Research Part A: Policy and Practice, 39(7-9), 737-754.
29. Noland, R. B., & Oh, L. (2004). The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis & Prevention, 36(4), 525-532.
30. Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., & Wets, G. (2013). Evaluating the road safety effects of a fuel cost increase measure by means of zonal crash prediction modeling. Accident Analysis & Prevention, 50, 186-195.
31. Quddus, M. A. (2008). Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accident Analysis & Prevention, 40(4), 1486-1497.
32. Rhee, K. A., Kim, J. K., Lee, Y. I., & Ulfarsson, G. F. (2016). Spatial regression analysis of traffic crashes in Seoul. Accident Analysis & Prevention, 91, 190-199.
33. Roshandel, S., Zheng, Z., & Washington, S. (2015). Impact of real-time traffic characteristics on freeway crash occurrence: Systematic review and meta-analysis. Accident Analysis & Prevention, 79, 198-211.
34. Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. Journal of clinical epidemiology, 58(9), 894-901.
35. Theofilatos, A., Ziakopoulos, A., Papadimitriou, E., & Yannis, G. (2018). How many crashes are caused by driver interaction with passengers? A meta-analysis approach. Journal of safety research, 65, 11-20.
36. Theofilatos, A., Ziakopoulos, A., Papadimitriou, E., Yannis, G., & Diamandouros, K. (2017). Meta-analysis of the effect of road work zones on crash occurrence. Accident Analysis & Prevention, 108, 1-8.
37. Ukkusuri, S., Hasan, S., Aziz, H., 2011. Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency. Transportation Research Record: Journal of the Transportation Research Board 2237, 98–106.
38. Van Houwelingen, H. C., Arends, L. R., & Stijnen, T. (2002). Advanced methods in meta-analysis: multivariate approach and meta-regression. Statistics in medicine, 21(4), 589-624.
39. Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. Journal of statistical software, 36(3).
40. Wang, Y., Veneziano, D., Russell, S., & Al-Kaisy, A. (2016). Traffic Safety Along Tourist Routes in Rural Areas. Transportation Research Record: Journal of the Transportation Research Board, (2568), 55-63.
41. Wang, J., & Huang, H. (2016). Road network safety evaluation using Bayesian hierarchical joint model. Accident Analysis & Prevention, 90, 152-158.
42. Wang, Y., & Kockelman, K. M. (2013). A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. Accident Analysis & Prevention, 60, 71-84.
43. Wang, C., Quddus, M. A., & Ison, S. G. (2009). Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. Accident Analysis & Prevention, 41(4), 798-808.
44. Wang, X., & Abdel-Aty, M. (2006). Temporal and spatial analyses of rear-end crashes at signalized intersections. Accident Analysis & Prevention, 38(6), 1137-1150.

45. Wei, F., & Lovegrove, G. (2013). An empirical tool to evaluate the safety of cyclists: Community based, macro-level collision prediction models using negative binomial regression. Accident Analysis & Prevention, 61, 129-137.
46. Wier, M., Weintraub, J., Humphreys, E. H., Seto, E., & Bhatia, R. (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis & Prevention, 41(1), 137-145.
47. World Health Organization – WHO. (2018). Global status report on road safety 2018. Available from: http://www.who.int/violence_injury_prevention/road_safety_status/2018/en/
48. Xu, P., Huang, H., & Dong, N. (2018). The modifiable areal unit problem in traffic safety: basic issue, potential solutions and future research. Journal of traffic and transportation engineering (English edition), 5(1), 73-82.
49. Xu, P., & Huang, H. (2015). Modeling crash spatial heterogeneity: random parameter versus geographically weighting. Accident Analysis & Prevention, 75, 16-25.
50. Yasmin, S., & Eluru, N. (2016). Latent segmentation based count models: analysis of bicycle safety in Montreal and Toronto. Accident Analysis & Prevention, 95, 157-171.
51. Ziakopoulos, A., Theofilatos, A., Papadimitriou, E., & Yannis, G. (2019). A meta-analysis of the impacts of operating in-vehicle information systems on road safety. IATSS Research.
52. Ziakopoulos, A., & Yannis, G. (2019). A review of spatial approaches in road safety. Accident Analysis & Prevention, 105323.