The Supplementary for "The data-driven pattern for healthy behaviors of car drivers based on daily records of Traffic Count data from 2018 to 2019 near airports. The Functional Data Analysis"

Mohammad Fayaz[1],Alireza Abadi[2,3*] , Soheila Khodakarim, Mohammadreza Hoseini[5], Alireza Razzaghi[6]

[1] PhD Student of Biostatistics, Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

[2] Professor of Biostatistics, Department of Community Medicine, Faculty of medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

[3] Social Determinants of Health Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

[4]Associate Professor, School of Allied Medical Sciences, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[5]Master Student of Transportation, Tehran North Branch, Islamic Azad University, Tehran, Iran

[6] Road Traffic Injury Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

* Corresponding Author: Dr. Alireza Abadi (alirezaabadi@gmail.com)

Contents

1 The logit of Driving Offences Probability

1.1 The Bayesian Function-on-Scalar Regression

1.1.1 Statistical Formula

We use fully Bayesian Function-on-scalar regression for high dimensional data to assess the following relationships:

$$Y_i(t) = log(\frac{P(t)_{offences\ occurens}}{1 - P(t)_{offences\ Occurens}})_i = \sum_{k=1}^{K} f_k(t)\beta_{k,i} + \epsilon_i(t)$$

$\epsilon_i(t) \sim N(0, \sigma_\epsilon^2), \qquad t\epsilon[0,23]$

$$\beta_{k,i} = \mu_k + Station_i \times \alpha_{1,k} + Direction_i \times \alpha_{2,k} + Weekend_i \times \alpha_{3,k} + (Station_i \times Direction_i) \times \alpha_{4,k} + (Station_i \times Weekend_i) \times \alpha_{5,k} + (Direction_i \times Weekden_i) \times \alpha_{6,k} + (Station_i \times Direction_i \times Weekden_i) \times \alpha_{7,k} + \gamma_{k,i}$$

$\mu_k \sim N(0, \sigma_{\mu_k}^2)$,

$\alpha_{j,k} \sim N\left(0, \sigma_{\alpha_{j.k}}^2\right)$,

$\gamma_{k,i} \sim N(0, \sigma_{\gamma_{k,i}}^2)$,

The $f_k(t)$ are basis functions with $k\epsilon[0,23]$ number of basis for within-curve dependence of functional data, and $\{\beta_{k,i}\}$ are coefficients for modeling the between-curve dependence. The priors for the intercepts$\{\mu_k\}$, the regression coefficients $\{\alpha_{j,k}\}$ and the subject-specific errors $\{\gamma_{k,i}\}$ are Gaussian and for observation error variance σ_ϵ^2 is Jeffrey. The shrinkage priors for the regression coefficients and the MCMC sampling algorithm is used to estimate the regression coefficients. The variable selection is based on the decouples shrinkage and selection (DSS). (32)

1.1.2 The estimated regression coefficient plot

The estimated regression coefficient functions are presented in the following figures:

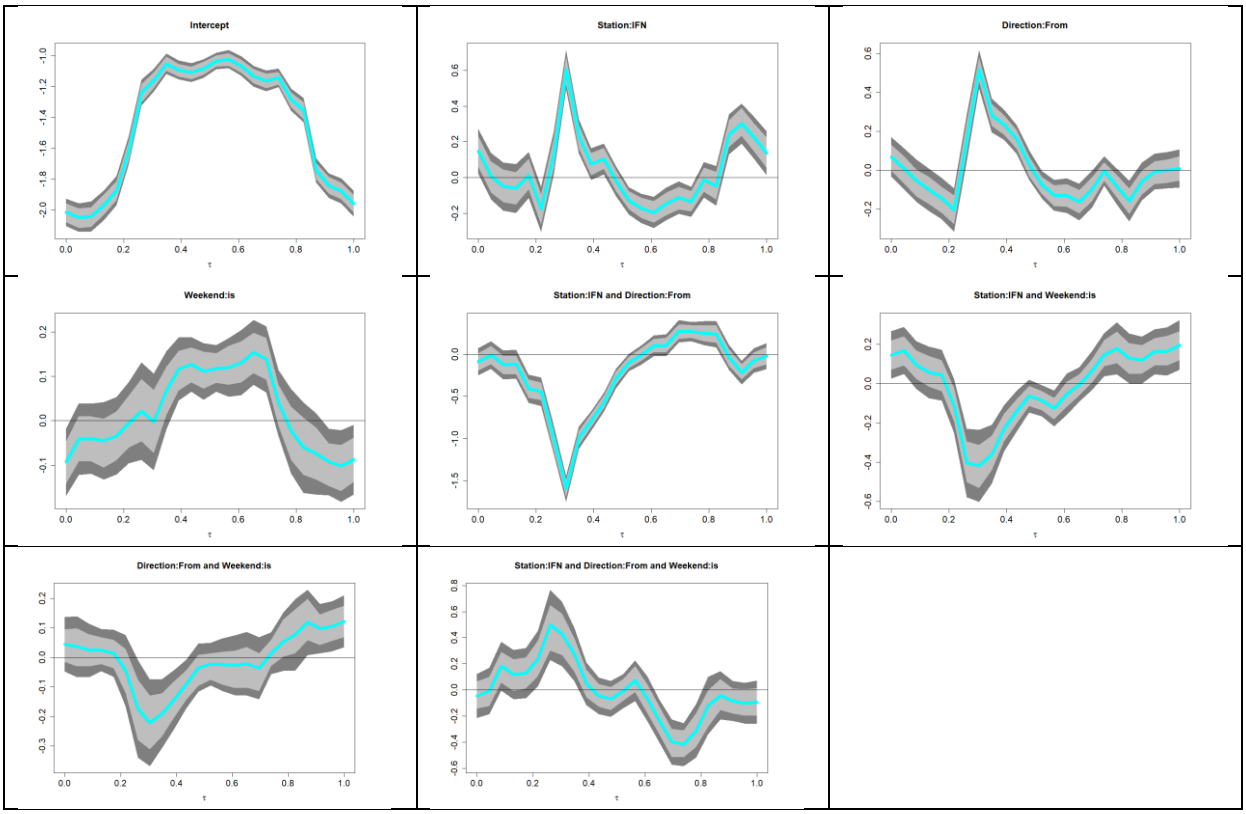


Figure 1. Estimated regression coefficient function, 95% pointwise credible intervals (light gray), and 95% simultaneous credible bands (dark gray). A horizontal line denotes zero change in the odds ratio.

1.1.3 The estimated regression coefficient table

The estimated coefficient of Bayesian Function on Scalar Regression is presented in the following table.

Table 1: The Estimated Coefficient of Bayesian, Mean (Standard Deviation)

	Intercept	b1	b2	b3	b4	b5	b6	b7
1	-7.3664 (0.107)	0.2006 (0.159)	-0.0515 (0.105)	0.0105 (0.05)	-1.0291 (0.177)	0.0415 (0.089)	-0.002 (0.041)	0.0376 (0.109)
2	0.28 (0.03)	0.2546 (0.054)	0.5096 (0.044)	0.2277 (0.059)	-2.035 (0.072)	-0.8866 (0.097)	-0.4481 (0.085)	0.696 (0.134)
3	-0.0626 (0.033)	-0.0216 (0.036)	-0.0299 (0.036)	0.1408 (0.052)	0.4291 (0.064)	0.0104 (0.041)	0.005 (0.03)	-0.5581 (0.104)
4	0.0883 (0.036)	-0.4047 (0.044)	-0.1178 (0.058)	0.3132 (0.052)	0.4921 (0.071)	-0.0084 (0.046)	-0.0566 (0.09)	-0.4205 (0.123)
5	-0.3533 (0.025)	0.7489 (0.033)	0.5144 (0.031)	-0.1159 (0.038)	-0.8017 (0.044)	-0.0435 (0.051)	0.0003 (0.023)	0.0847 (0.064)
6	-0.1296 (0.018)	0.0548 (0.025)	0.265 (0.026)	-0.0005 (0.019)	-0.0061 (0.03)	0.0011 (0.022)	-0.0965 (0.053)	-0.0515 (0.058)
7	-0.0282 (0.017)	0.0507 (0.026)	0.131 (0.024)	-0.0033 (0.015)	-0.199 (0.036)	-0.0207 (0.03)	-0.0007 (0.015)	0.0488 (0.046)
8	-0.0396 (0.016)	0.0516 (0.024)	0.06 (0.021)	-0.0262 (0.026)	-0.0052 (0.027)	0.157 (0.049)	-0.0021 (0.018)	-0.3596 (0.056)
9	-0.0107 (0.015)	0.0159 (0.021)	0.0575 (0.022)	-0.0197 (0.021)	-0.0485 (0.034)	0.0155 (0.029)	0.0009 (0.017)	-0.0449 (0.046)
10	-0.0287 (0.014)	0.0888 (0.021)	-0.0203 (0.018)	-0.0012 (0.013)	-0.0025 (0.025)	-0.0147 (0.024)	-0.0048 (0.016)	-0.0366 (0.038)
11	-0.009 (0.014)	0.0936 (0.02)	-0.0326 (0.02)	-0.0101 (0.015)	-0.0933 (0.032)	-0.0018 (0.018)	-0.001 (0.015)	0.0556 (0.042)
12	-0.0068 (0.01)	0.0002 (0.013)	0.0206 (0.017)	-0.0119 (0.016)	-0.0236 (0.025)	-0.0103 (0.022)	0.0063 (0.018)	0.085 (0.044)
13	-0.0163 (0.011)	0.0958 (0.016)	-0.0171 (0.016)	0.0267 (0.017)	-0.1197 (0.025)	0.0044 (0.017)	0.0086 (0.017)	0.0166 (0.027)
14	-0.021 (0.012)	0.0366 (0.018)	0.0049 (0.012)	-0.0034 (0.013)	-0.0029 (0.019)	0.0167 (0.024)	-0.0054 (0.015)	-0.0385 (0.036)
15	-0.007 (0.01)	0.0017 (0.011)	-0.0241 (0.015)	0.0108 (0.015)	0.0037 (0.016)	0.0147 (0.022)	-0.0055 (0.015)	0.0125 (0.027)
16	-0.0118 (0.01)	0.0336 (0.017)	0.0127 (0.015)	-0.0053 (0.012)	-0.0244 (0.023)	-0.0102 (0.018)	-0.0041 (0.014)	0.001 (0.024)
17	-0.0075 (0.01)	-0.0121 (0.014)	0.0453 (0.016)	0.0237 (0.019)	-0.0195 (0.022)	-0.0038 (0.017)	0.0145 (0.021)	0.0104 (0.026)
18	0.0194 (0.011)	-0.0426 (0.016)	0.0035 (0.012)	-0.0277 (0.017)	0.0368 (0.022)	-0.0134 (0.023)	-0.0038 (0.013)	0.0251 (0.03)
19	-0.0372 (0.015)	0.0426 (0.02)	0.0495 (0.02)	0.0286 (0.018)	-0.0534 (0.03)	0.0587 (0.03)	-0.0021 (0.014)	-0.1591 (0.035)
20	-0.007 (0.009)	-0.0053 (0.011)	0.012 (0.011)	0.0243 (0.015)	-0.0048 (0.015)	-0.0131 (0.019)	0.0001 (0.013)	-0.0348 (0.028)
21	-0.0002 (0.008)	0.0108 (0.012)	0.0182 (0.013)	0.0022 (0.01)	-0.0929 (0.021)	-0.0071 (0.014)	0.0123 (0.016)	0.0113 (0.022)
22	0.03 (0.01)	-0.0161 (0.014)	-0.053 (0.013)	-0.0055 (0.011)	0.0122 (0.017)	-0.0605 (0.024)	0.0027 (0.012)	0.1214 (0.031)
23	-0.0046 (0.01)	0.041 (0.013)	-0.0263 (0.013)	0.0029 (0.009)	0.0359 (0.022)	-0.0031 (0.012)	0.0007 (0.01)	-0.0018 (0.019)

** b1= Station: IFN, b2= Direction: From, b3= Weekend: is, b4= Station: IFN and Direction: From, b5= Station: IFN and Weekend: is, b6= Direction: From and Weekend: is, b7= Station: IFN and Direction: From and Weekend:is

2 The Behavioral Factor

2.1 The average speed in airports

2.1.1 - The B-Spline with Smoothing Parameter

We estimate the average speed function by B-Spline using smoothing parameter by GCV for IKA and IFN airport and in both directions. In the IKA airport, the two curves to the airport and from the airport are almost the same in most hours, but in the from the airport has one peak at 9:00. In the other hand, in IFN airport the average speed to the airport is higher in all hours than from the airport, and it has also one peak at 8:00.

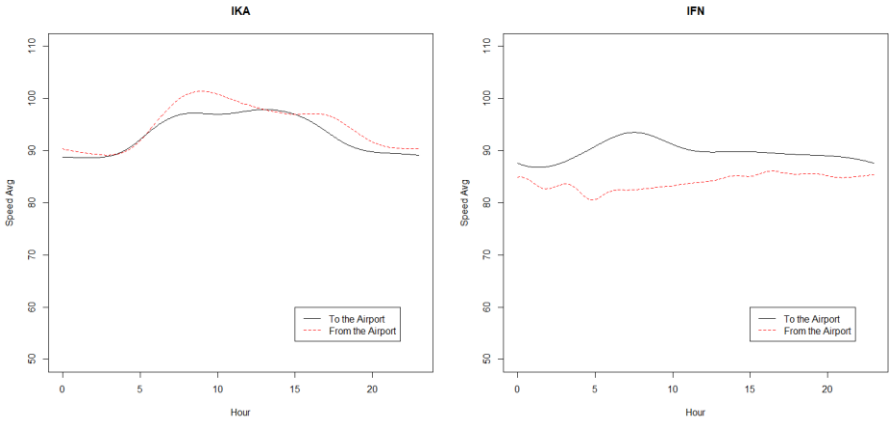


Figure 2: The functional average speed with using B-Spline in both airports and both directions.

2.2 The Bayesian Function on Regression

2.2.1 The estimated regression coefficient plot

The Bayesian FOSR is used to estimate the effects of total vehicle types from 1 to 5 on the average speed function.

The intercept shows that the average speed from 10:00 to 20:00 is the highest. The type 1 vehicle from 20:00 to 21:00 has statistically negative effect. In the type 2 vehicle from 19:00 to 20:00 has statistically positive effect. In the type 3 at 6:00 and 20:00 has statistically positive effect. In type 4, at 20:00 has statistically negative effect and in type 5 has at 6:00 and 20:00 statistically negative effect. (Figure 1_6)

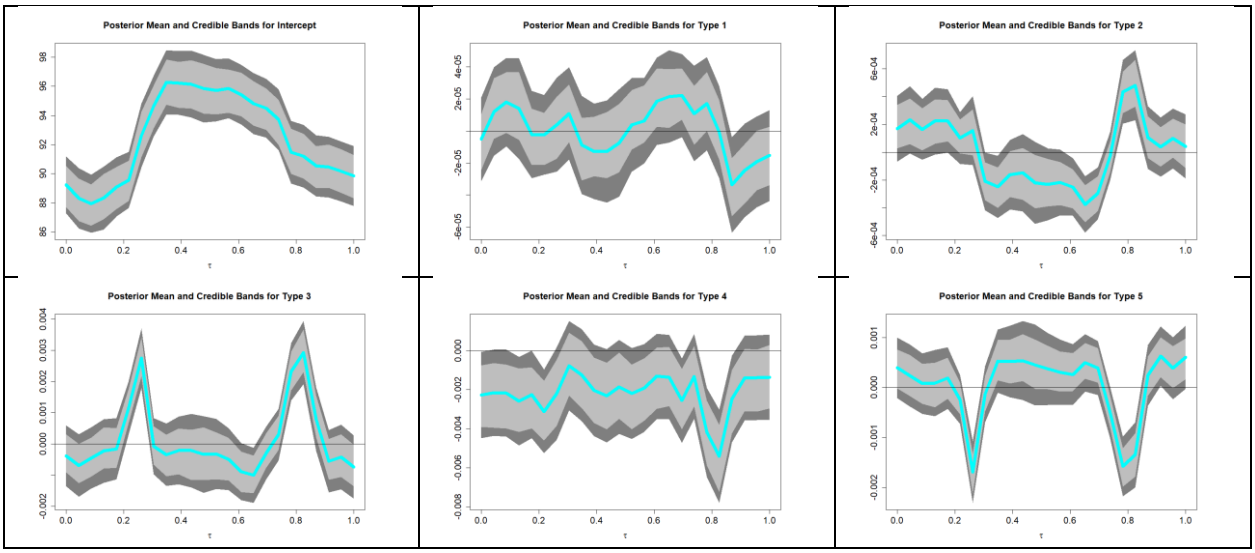


Figure 3: The Posterior Mean and Credible Bands for estimated effects of Vehicle Types from 1 to 5.

3 Type of Driving Offences

3.1 Speeding

3.1.1 The Descriptive Statistics

The proportion of traffic offences type 1: Speeding

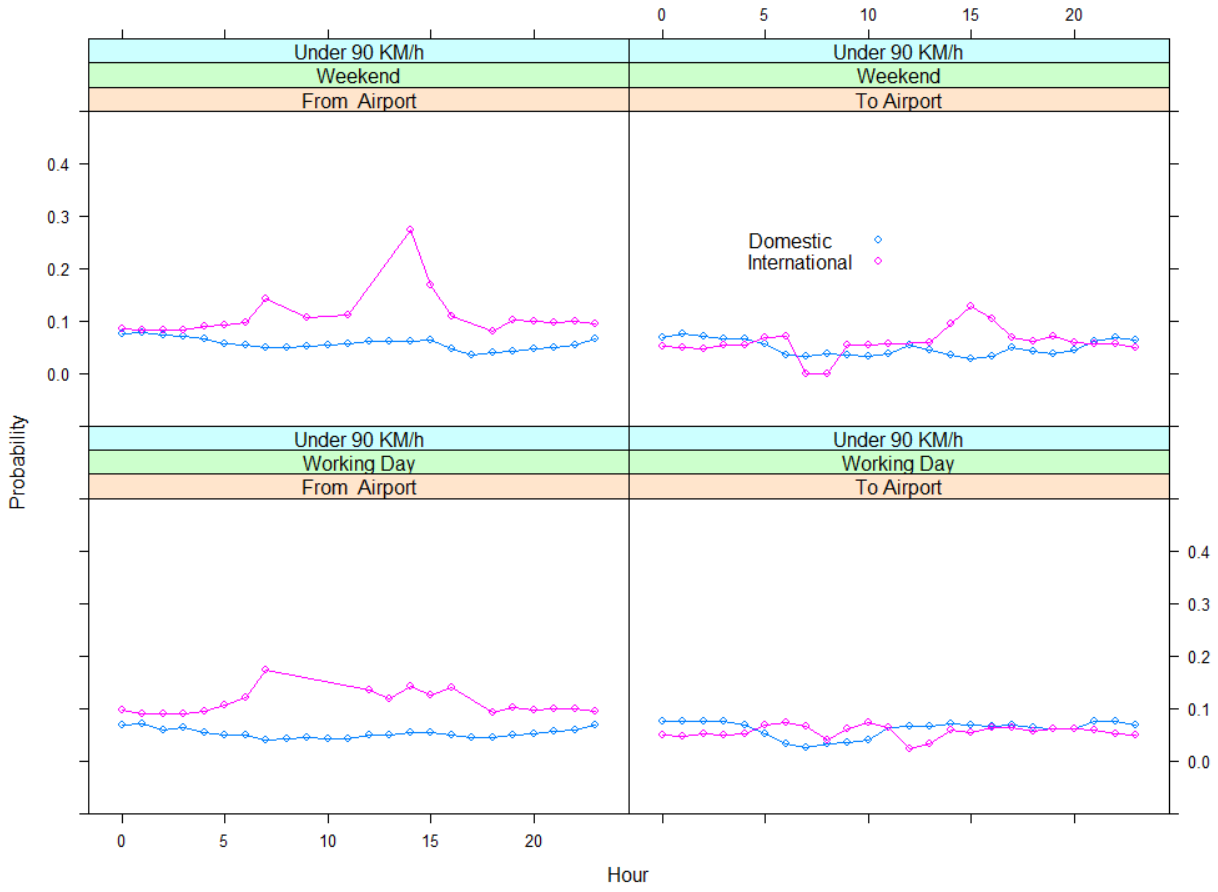


Figure 4: The proportion of traffic offences type 1, Speeding. For under 90 KM/h average speed, working day and direction.

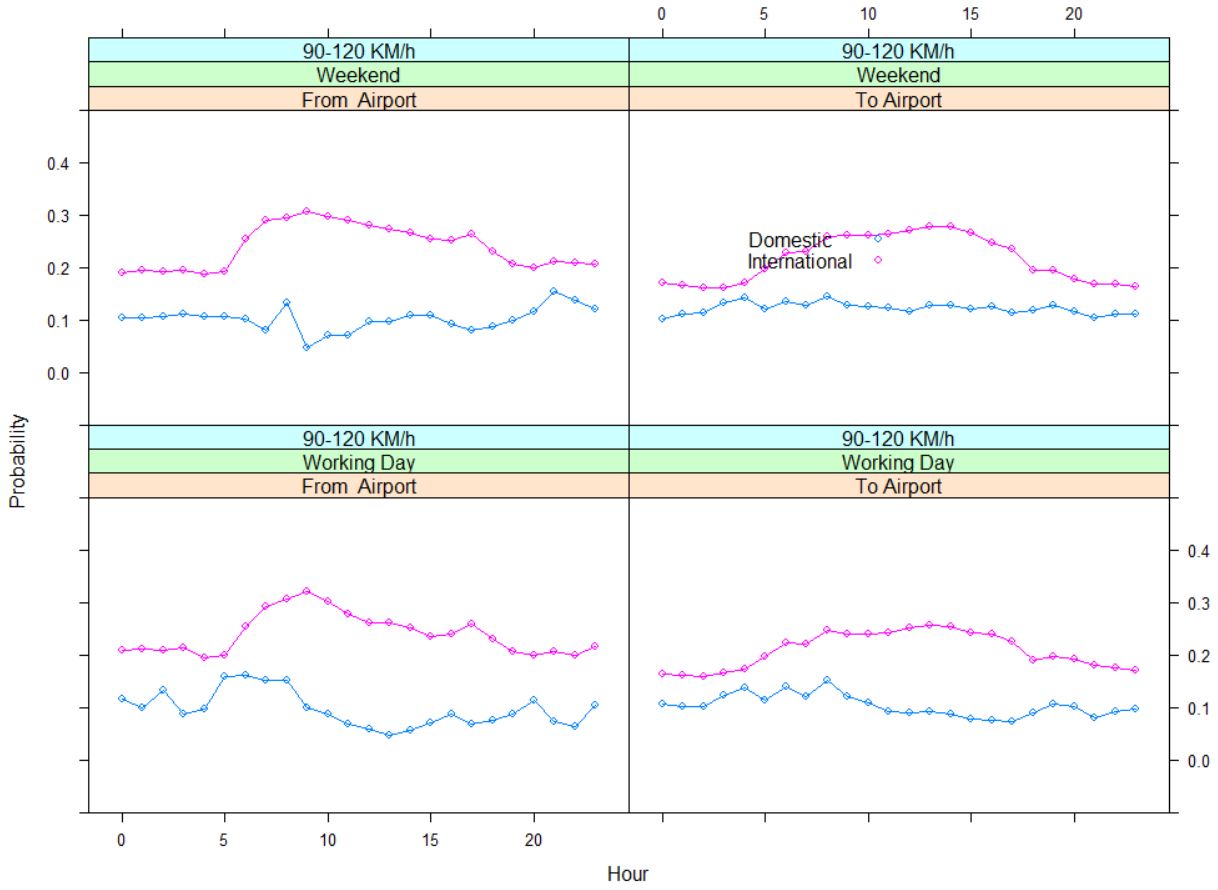


Figure 5: The proportion of traffic offences type 1, Speeding. For 90-120 KM/h average speed, working day and direction.

3.1.2 The estimated effects

The GAM (Generalized Additive Model) Result

Type 1 Driving Offences:

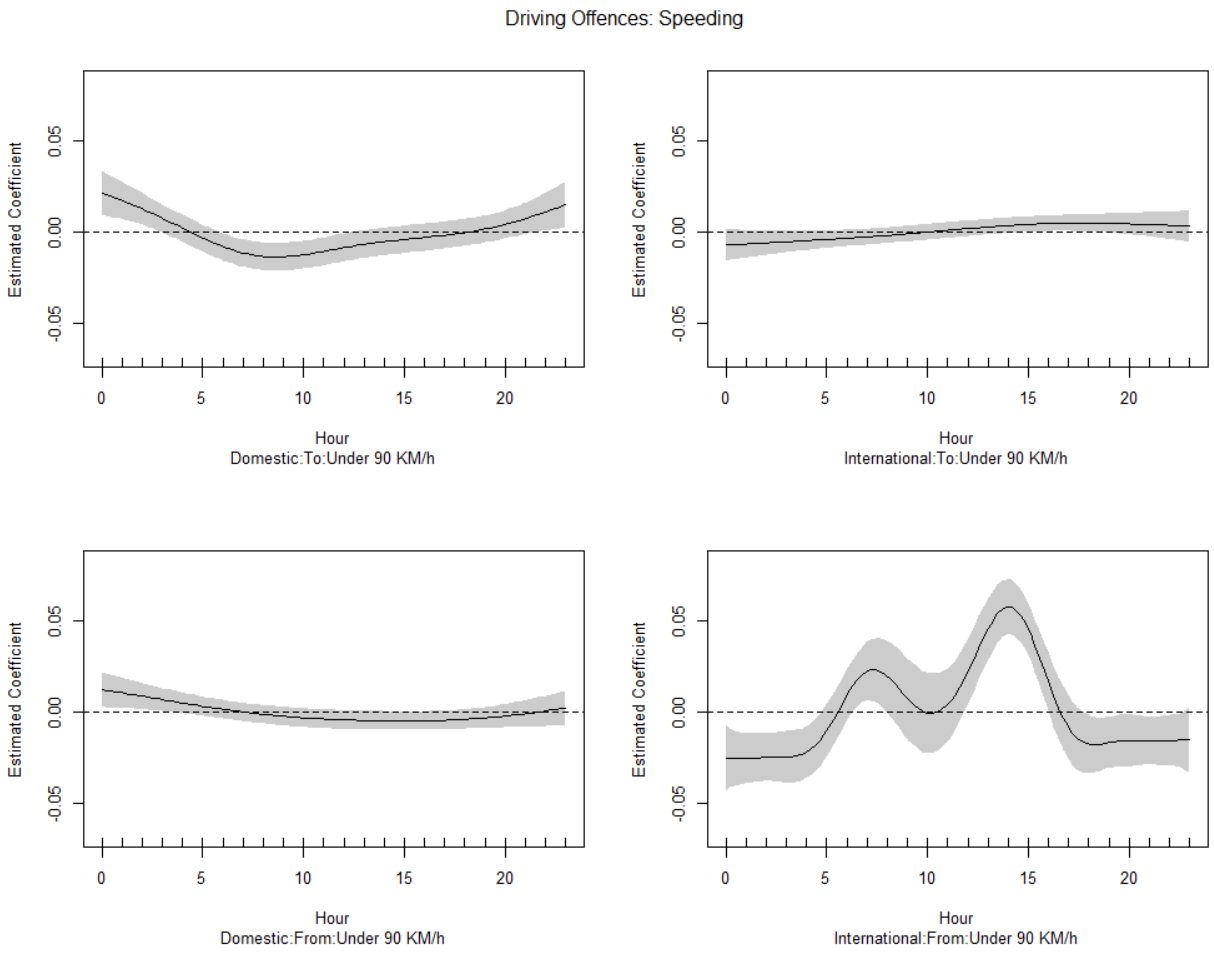
Table 2: The ANOVA table of the Parametric Coefficients for type 1 Driving Offences

Terms	Base	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept		0.10	0.0025	40.5	< 2e-16
Days	Working Day	0.00	0.0017	-2.4	0.01511
Direct	To Airport	0.01	0.0033	4.0	8.30E-05
Type	International	0.14	0.0033	41.6	< 2e-16
Speed	Under 90 KM/H	-0.04	0.0033	-13.4	< 2e-16
Direct:Type	To Airport: International	-0.04	0.0047	-8.6	3.29E-16
Direct:Speed	To Airport: Under 90 KM/H	-0.01	0.0047	-2.7	0.00778
Direct:Type:Speed	From Airport : International: Under 90 KM/H	-0.08	0.0050	-16.1	< 2e-16
Direct:Type:Speed	To Airport : International: Under 90 KM/H	-0.10	0.0047	-20.3	< 2e-16

Family: Gaussian, AIC: -1949.267 , BIC: -1733.174 , R-sq.(adj): 0.95, Deviance Explained: 95.6%, n= 373

Smooth terms by Hour	edf	Ref.df	F	p-value
Domestic:From Airport:90-120 KM/h	5.445	13	3.569	3.64E-09
International:From Airport:90-120 KM/h	7.937	13	19.495	< 2e-16
Domestic:To Airport:90-120 KM/h	3.216	13	1.449	0.000158
International:To Airport:90-120 KM/h	5.552	13	19.694	< 2e-16
Domestic:From Airport:Under 90 KM/h	1.762	13	0.487	0.022009
International:From Airport:Under 90 KM/h	8.257	13	7.587	< 2e-16
Domestic:To Airport:Under 90 KM/h	3.16	13	1.624	4.63E-05
International:To Airport:Under 90 KM/h	1.354	13	0.339	0.037615

Number of basis= 14, cubic regression spline, Estimation method= REML (edf: Effective Degrees of Freedom)



١٩١

١٩٢ Figure 6: The estimated coefficients of Speeding driving offences for Under 90 KM/upper left: The
١٩٣ domestic airport and "To" direction is slightly zero in all hours, Upper right: The international airport and
١٩٤ "To" direction is slightly zero in all hours, bottom left: The domestic airport and "From" direction is slightly
١٩٥ zero in all hours, bottom right: The International airport and "From" direction has positive effect between
١٩٦ 13 to 15.

١٩٧

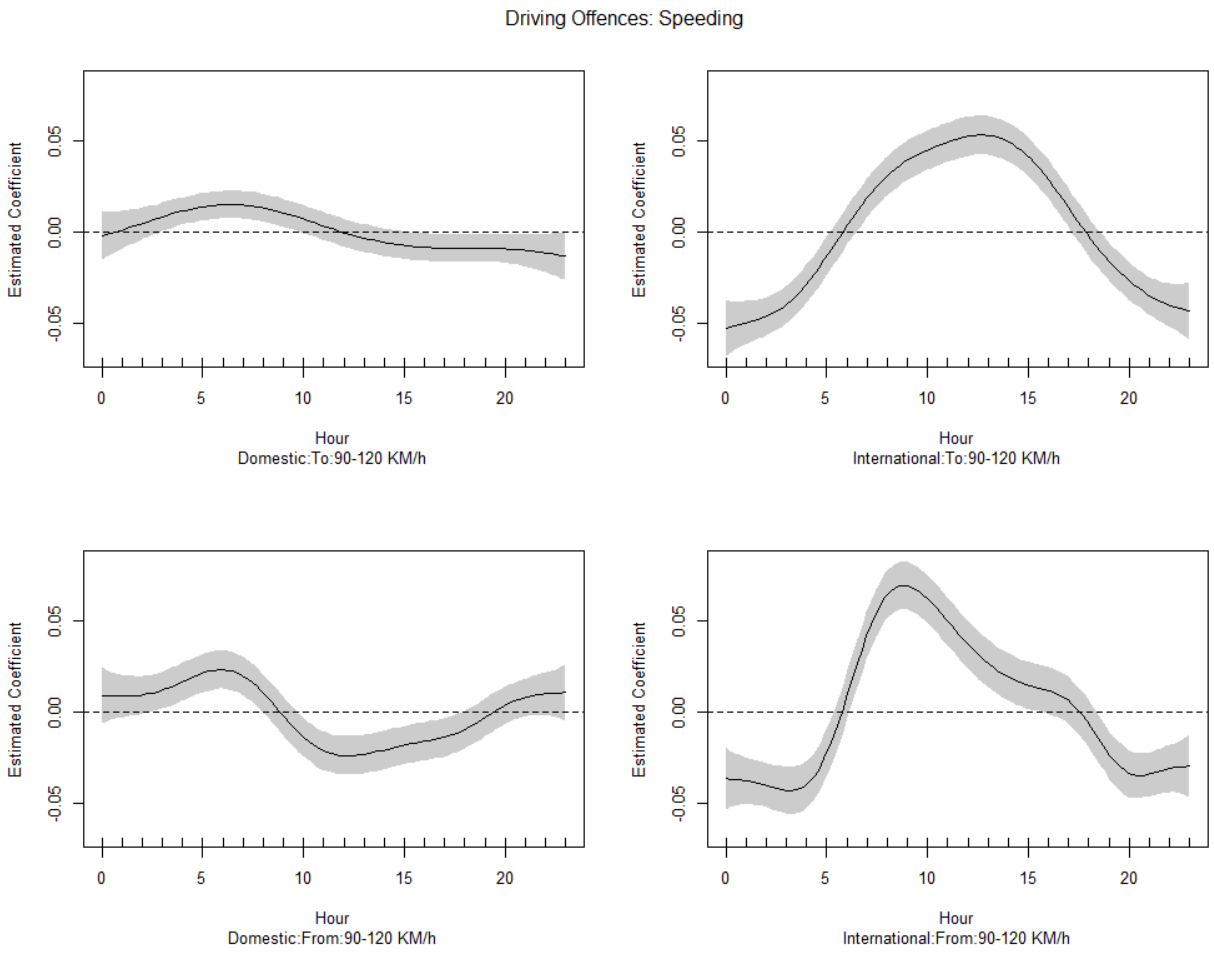


Figure 7: – The estimated coefficients of Speeding driving offences for 90-120 KM/upper left: The domestic airport and "To" direction has positive effects in early morning (4-8), Upper right: The international airport and "To" direction has positive effect in working hours (7-18), bottom left: The domestic airport and "From" direction has positive effect in early morning (4-8), bottom right: The International airport and "From" direction has positive effect in working hours (7-15).

3.2 Unsafe Distance

3.2.1 The Descriptive Statistics

The proportion of traffic offences type 2: Unsafe Distance

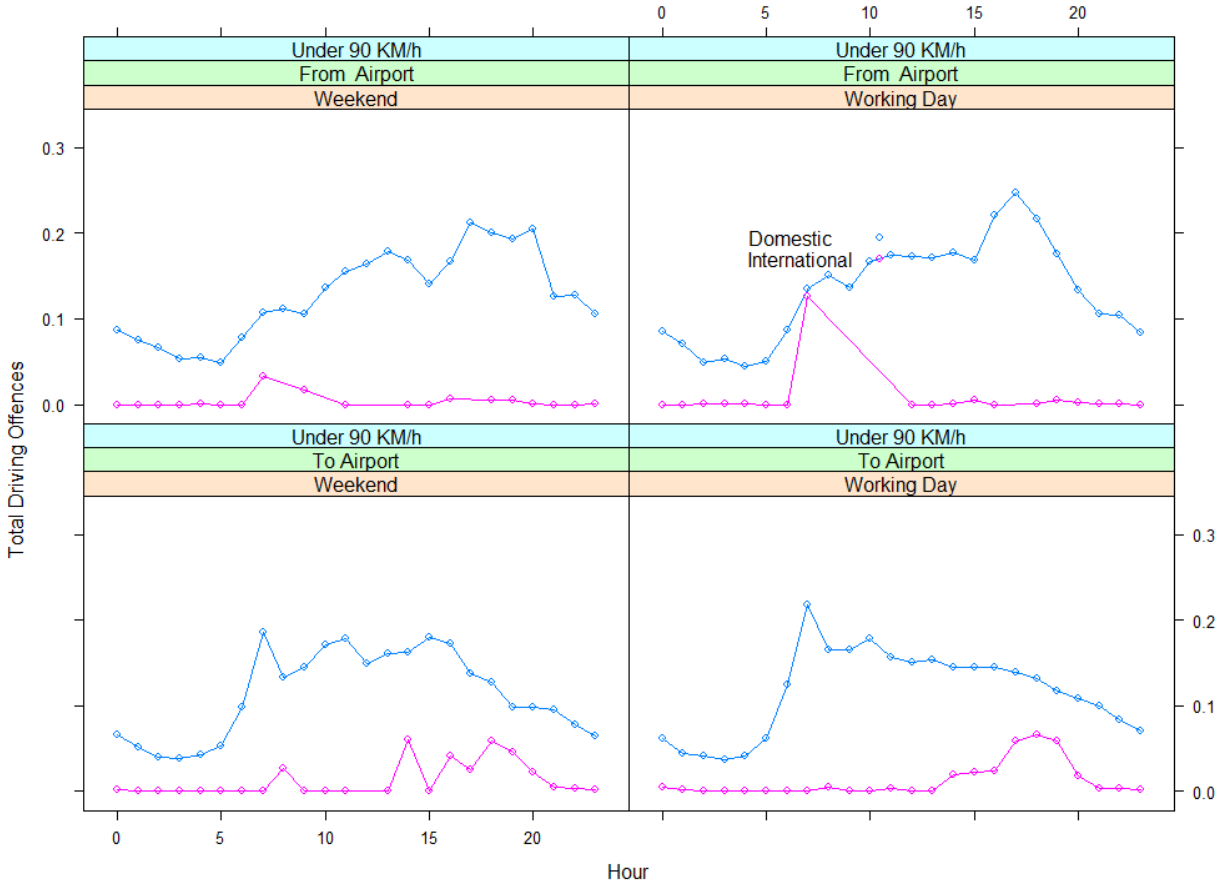


Figure 8: The proportion of traffic offences type 2, Unsafe Distance. For under 90 KM/h average speed, working day and direction.

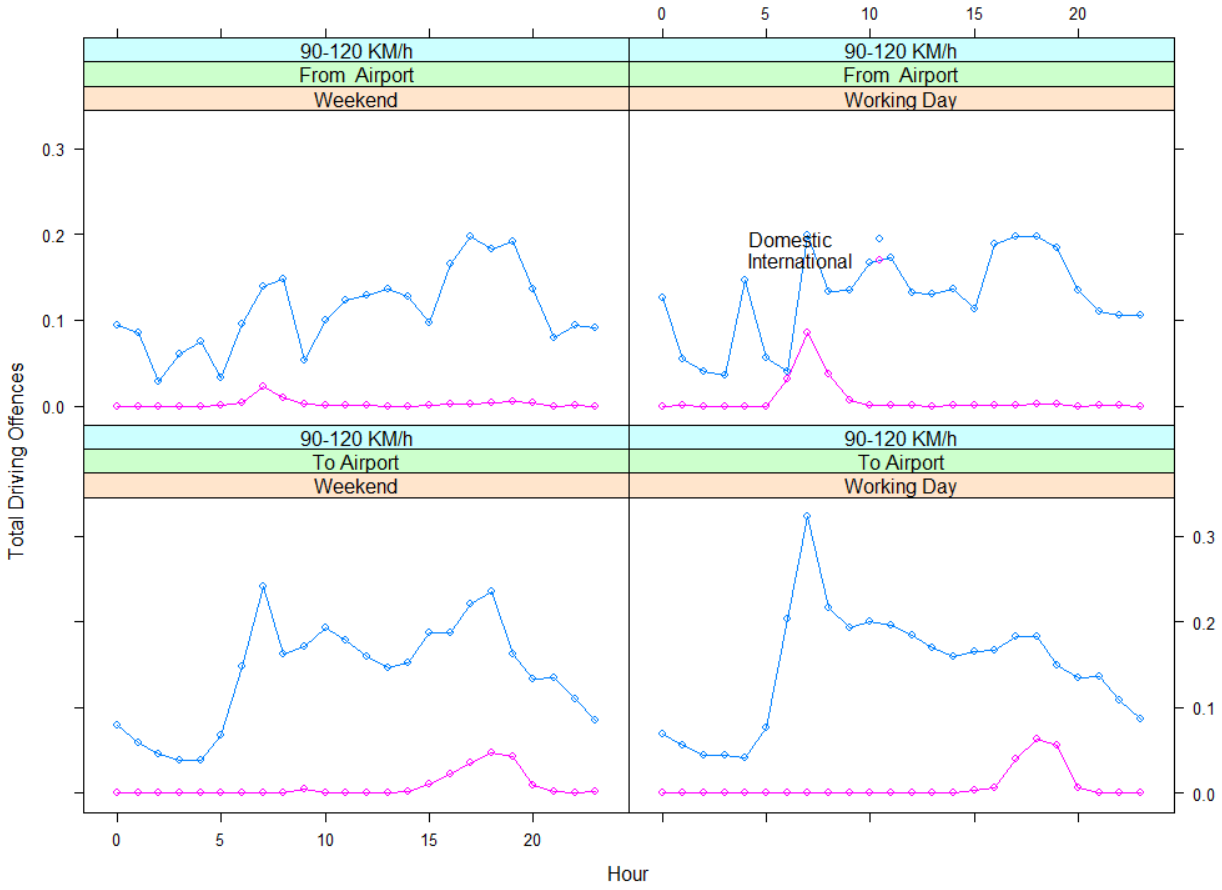


Figure 9: The proportion of traffic offences type 2, Unsafe Distance. For 90-120 KM/h average speed, working day and direction.

3.2.2 The estimated effects

Type 2 Driving Offences:

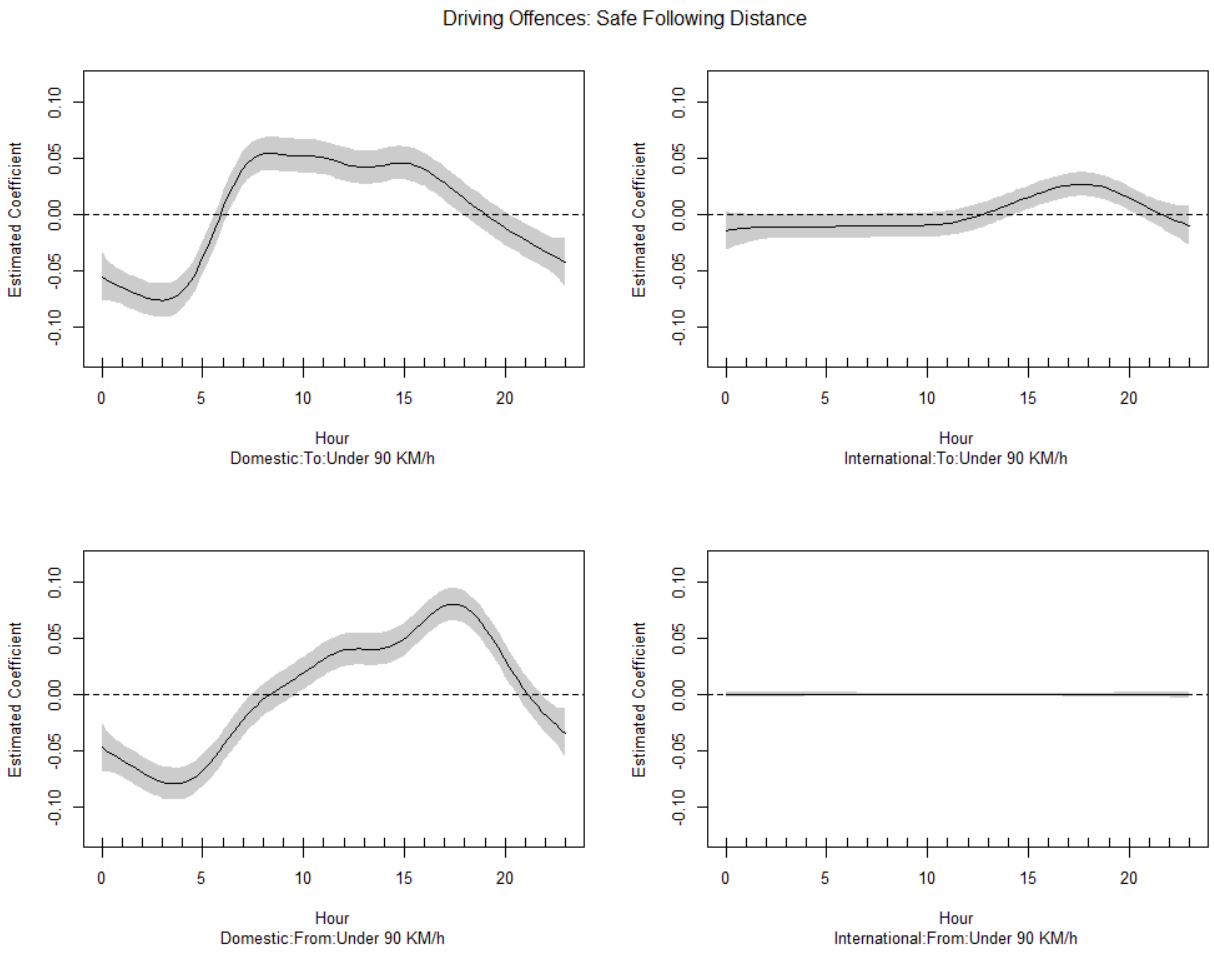
Table 3: The ANOVA table of the Parametric Coefficients for type 2 Driving Offences

Terms	Base	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept		0.116	0.0028	41.002	< 2e-16
Days	Working Day	0.005	0.0019	2.459	0.014475
Direct	To Airport	0.023	0.0038	5.978	6.11E-09
Type	International	-0.114	0.0038	-30.207	< 2e-16
Speed	Under 90 KM/H	0.011	0.0038	2.930	0.003635
Direct:Type	To Airport: International	-0.020	0.0053	-3.689	0.000265
Direct:Speed	To Airport: Under 90 KM/H	-0.038	0.0053	-7.192	4.66E-12
Direct:Type:Speed	From Airport : International: Under 90 KM/H	-0.010	0.0055	-1.841	0.066522
Direct:Type:Speed	To Airport : International: Under 90 KM/H	0.032	0.0053	6.006	5.23E-09

Family: Gaussian, AIC: -1850.26, BIC: -1589.968, R-sq.(adj): 0.936, Deviance Explained: 94.5%, n= 373

Smooth terms by Hour	edf	Ref.df	F	p-value
Domestic:From Airport:90-120 KM/h	10.73	13	17.887	< 2e-16
International:From Airport:90-120 KM/h	0.07	13	0.006	0.31497
Domestic:To Airport:90-120 KM/h	11.61	13	39.296	< 2e-16
International:To Airport:90-120 KM/h	4.43	13	1.501	0.000438
Domestic:From Airport:Under 90 KM/h	8.07	13	28.246	< 2e-16
International:From Airport:Under 90 KM/h	0.10	13	0.008	0.310932
Domestic:To Airport:Under 90 KM/h	8.52	13	23.759	< 2e-16
International:To Airport:Under 90 KM/h	4.58	13	2.357	2.27E-06

Number of basis= 14, cubic regression spline, Estimation method= REML (edf: Effective Degrees of Freedom)



٢٤٤

٢٤٥ Figure 10: The estimated coefficients of Failing to follow a vehicle at a safe distance offences for Under 90
٢٤٦ KM/upper left: The domestic airport and "To" direction has positive effect working hours, Upper right: The
٢٤٧ international airport and "To" direction has only positive effect 15 to 20, bottom left: The domestic airport
٢٤٨ and "From" direction has positive effect between 10 to 20, bottom right: The International airport and
٢٤٩ "From" direction has zero effect.

٢٥٠

٢٥١

٢٥٢

٢٥٣

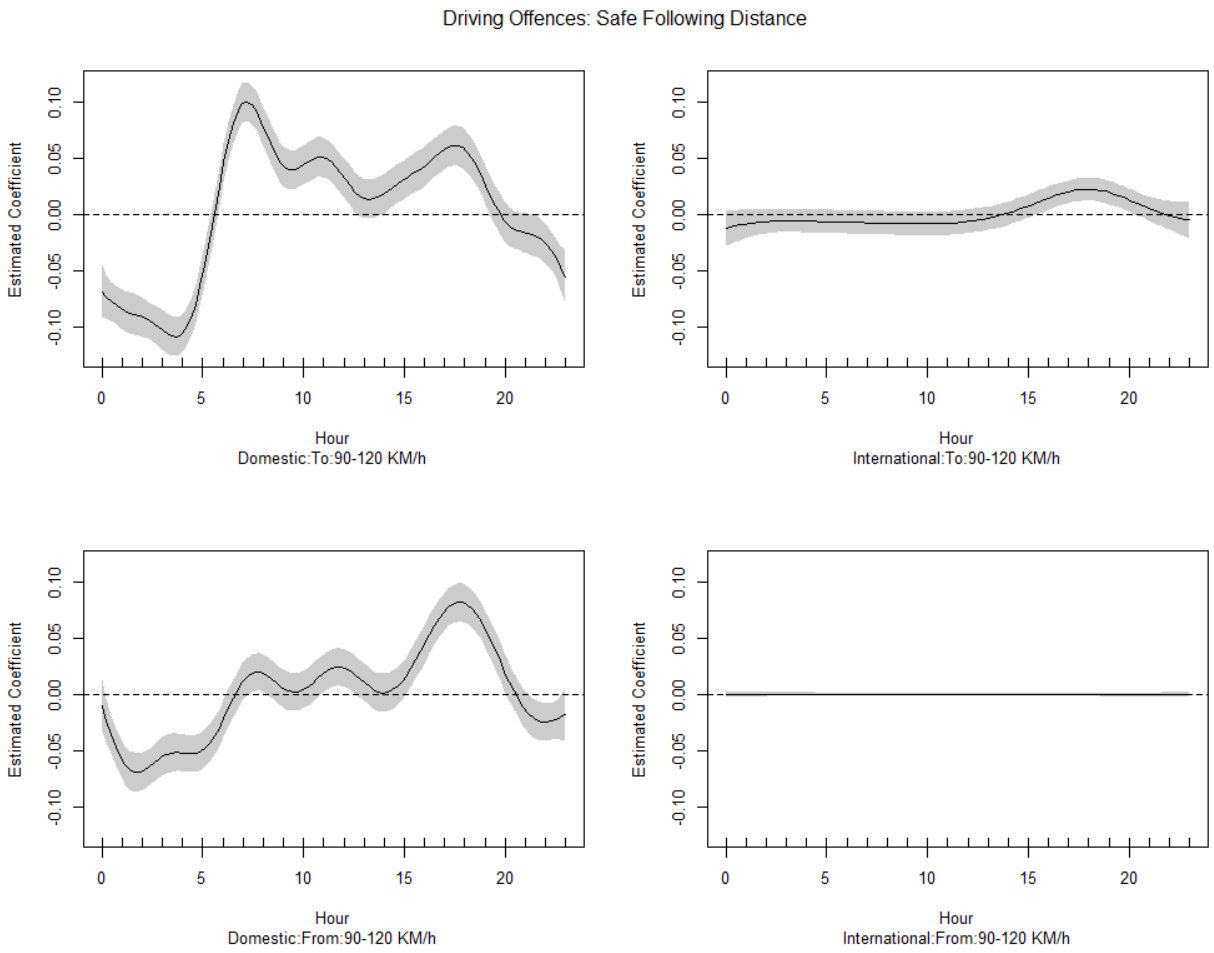


Figure 11: The estimated coefficients of Failing to follow a vehicle at a safe distance offences for 90-120 KM/h. Upper left: The domestic airport and "To" direction has positive effect working hours, Upper right: The international airport and "To" direction has only positive effect 15 to 20, bottom left: The domestic airport and "From" direction has positive effect between 10 to 20, bottom right: The International airport and "From" direction has zero effect.

4 Comparisons between other countries airports

4.1 The plot

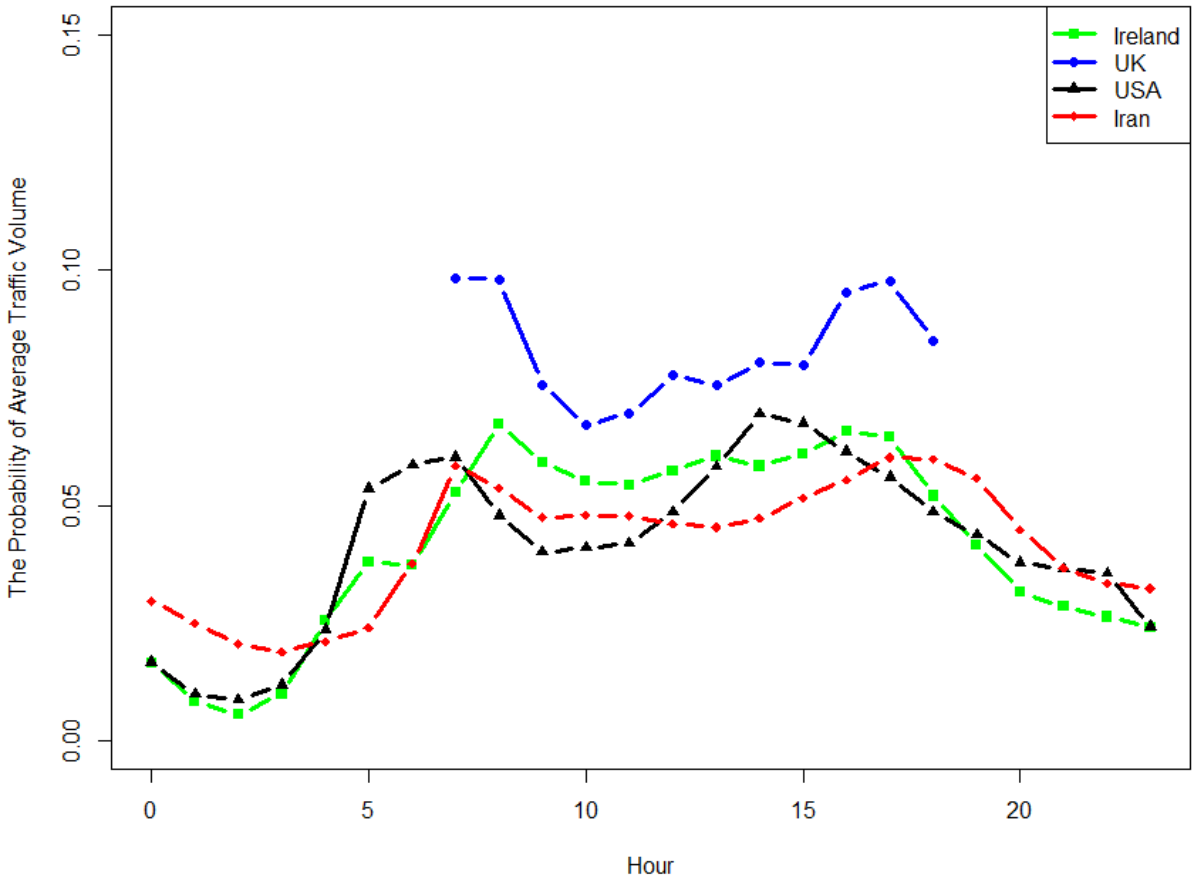


Figure 12: The Probability of Average Traffic Volume in different Airport Countries.

4.2 The Data resources

4.2.1 USA

There are one airport:

Table 4- The list of traffic count sites near the Hartsfield–Jackson Atlanta International Airport in the US

Row	Airport name	Traffic Count Name	Description	Time-Interval
1	Hartsfield–Jackson Atlanta International Airport	0000063_1407 - 063-1407 0000063_R160 - 063-r160 0000063_R651 - 063-r651 0000063_8264 - 063-8264	SR 000300 BEG AT RPY 403225 L SR 040300 BEG AT Airport Blvd	2018

We use only Hartsfield–Jackson Atlanta International Airport dataset because they are available online.

Data Profile

The Georgia Department of Transportation's Traffic Analysis and Data Application (TADA!)

Description

The Georgia Department of Transportation's Traffic Analysis and Data Application (TADA!) website presents data collected from the Georgia Traffic Monitoring Program located on the public roads in Georgia. The Website uses a dynamic mapping interface to allow the User to access data from the map as well as in a variety of report, graph, and data export formats.**Website**

https://gdottrafficdata.drakewell.com/publicmultinodemap.asp

4.2.2 UK

There are two airports:

Table 5- The list of traffic count sites near the Heathrow Airport and Gatwick Airport in the UK

Row	Airport name	Count Point IDs	Time-Interval
1	Heathrow Airport	Left (38599,47625,75148,78401) Up (26118,56114,942674,16112,3612136013) Right (58197,6123,73633,36309) Bottom (953083,942678,73734)	2018
2	Gatwick Airport	46035,18231,57660,36274,90299,946199	2018

The data is not available for all hours. The average of all count point were used.

Data Profile

Road traffic statistics

Data Description

Summary and street-level traffic data for road-links on the motorway, 'A' road and minor road network in Great Britain. Use this website to find information about the road traffic statistics collected over the last 18 years, across 34,416 manual count points. The latest information available covers 2018.

Website

https://roadtraffic.dft.gov.uk/

4.2.3 Ireland

There are two airports:

Table 6- The list of traffic count sites near Dublin, Cork and Shanon airport in Ireland

Row	Airport name	Traffic Count Name	Name	Description	Time-Interval
1	Dublin	NRA 000000001011	TMU N01 000.0 N	M01 Airport Link Road Between R132 Swords Road and Jn2 Dublin Airport	2018-10-23 to 2019-10-22
2	Cork	NRA 000000001271	TMU N27 000.0 N	N27 Between Cork Airport and Cork, Ballycurreen, Co. Cork	2018-10-23 to 2019-10-22
3	Shanon	NRA 000000001191	TMU N19 001.5 S	N19 Between R472 and Drumgeely roundabout, Ballymurtagh Co Clare.	2018-10-23 to 2019-10-22

We use only Dublin and Cork dataset because they are complete for that period.

Data Profile

TII Traffic Data Site

Data Description

The TII Traffic Data website presents data collected from the TII traffic counters located on the National Road Network. The Website uses a dynamic mapping interface to allow the User to access data in a variety of report formats.

Website

https://www.nratrafficdata.ie/c2/gmapbasic.asp?sgid=ZvyVmXU8jBt9PJE$c7UXt6