Bar-Ilan University
Tradition of Excellence

Risk factors, driver behaviour and accident probability. The case of distracted driving.

Panagiotis Papantoniou
PhD, Research Associate

National Technical University of Athens

Tel Aviv, July 2016

Objective

The analysis of the effect of road, traffic and driver risk factors on driver behaviour and accident probability at unexpected incidents, with particular focus on distracted driving

The development of risky driver profiles and road characteristics leading to increased possibility of driver error

Methodological steps

- Literature review

- Methodological review

- Research questions

- Driving simulator experiment

- Statistical analysis

- Methodological contributions

- Key research findings

- Further research

Literature review

Several literature reviews were implemented in the following research topics:

Driver behaviour and road safety

Driver distraction

Types of assessing driver distraction

Driving simulator characteristics

Driving simulator studies on driver distraction

Driver behaviour and road safety

- **Human factors** are the basic causes in 65-95% of road accidents (Salmon et al., 2011).

- Human factors involve a large number of specific factors that may be considered as accident causes, including (Yannis et al., 2013):

  - **Driver injudicious action** (speeding, traffic violations etc.)

  - **Driver error or reaction** (loss of control, failure to keep safe distances, sudden braking etc.)

  - **Behaviour or inexperience** (aggressive driving, nervousness, uncertainty etc.)

  - **Driver distraction** (cell phone use, conversation with passenger etc.)

  - **Driver impairment** (alcohol, fatigue etc.)

Driver Distraction

- **Driver distraction** is generally defined as "a diversion of attention from driving, because the driver is temporarily focusing on an object, person, task or event not related to driving, which reduces the driver's awareness, decision making ability and/or performance, leading to an increased risk of corrective actions, near-crashes, or crashes"

- Driver distraction may include four different types: **physical** distraction, **visual** distraction, auditory **distraction** and cognitive **distraction** (Ranney et al., 2000)

- Driver distraction factors can be subdivided into those that occur outside the vehicle (**external**) and those that occur inside the vehicle (**in-vehicle**).

Types of assessing driver distraction

The following **experiment types** of assessing driving behaviour and driver distraction exist (Papantoniou et al., 2015):

- Field tests

- Naturalistic driving experiments

- Driving simulator experiments

- In-depth accident investigation

- Stated preference surveys

Driving simulator characteristics

Driving simulators allow for the examination of a range of driving performance measures in a controlled, relatively realistic and safe driving environment

Advantages
- safe environment
- greater experimental control
- large range of test conditions (e.g., night and day, weather conditions, road environments)

Disadvantages
- data generally include the effect of learning
- feeling of safety
- simulator sickness

Driving simulator studies on driver distraction

- Most experiments are based on **small samples**, limited to rural road environment and no explicit (if at all) simulation of ambient traffic

- Participants in almost all driving simulator experiments implemented a practice scenario, but no specific performance measures were used to assess the driver's **familiarization**

- No pattern could be identified as regards the selection of **number and duration of trials**

- In 30% of studies no **counterbalancing** in the different trials was reported

Methodological review

Two targeted literature reviews took place in order to investigate:

- key **driving performance parameters**
  - A lot of different methods and **measures** exist for evaluating driving performance

- **statistical analyses** implemented
  - In more than half of the examined studies the main statistical analysis is **repeated measures Anova**

  - Latent model analysis and especially **structural equation models** have never been implemented in the field of driver distraction.

Research questions

1. **Design and implementation** of a simulator experiment aiming to deal with the basic limitations found in the literature:
   - Large and representative sample
   - Randomisation of trials
   - Adequate practice drive
   - Investigation of an optimum number of driving factors

2. Need to demonstrate a composite driving performance measure in order to examine driver distraction as a **multidimensional phenomenon**

3. Development and application of an **innovative statistical analysis methodology**

4. Estimation of the **combined effect** of distraction sources, driver as well as road and traffic environment characteristics directly on driving performance.

Driving simulator experiment (1/2)

A common simulator experiment in the framework of two research projects:

- **Distract** - Analysis of causes and impacts of driver distraction

- **DriverBrain** - Analysis of the performance of drivers with cerebral diseases

An **interdisciplinary** research team:

- Dpt. of Transportation Planning and Engineering NTUA
- Dpt. of Neurology of the University of Athens Medical School, UoA
- Dpt. of Psychology, School of Philosophy, Pedagogy and Psychology, UoA

Driving simulator characteristics

- Foerst Driving Simulator FPF
- 3 LCD wide screens 40″
- total angle view 170°
- driving position and support base

Driving criteria

- Have a valid driving license
- Had driven for more than 3 years
- Had driven more than 2500km during the last year
- Had driven at least once a week during the last year
- Had driven at least 10km/week during the last year

The design of the driving scenarios is a central component of the present PhD thesis and includes:

Area type

- Rural area
  2,1 km long, single carriageway and the lane width was 3m, with zero gradient and mild horizontal curves
- Urban area
  1,7km long, lane width 3,5m, separated by guardrails

Distraction conditions

- No distraction
- Cell phone use
- Conversation with the passenger

Rural Road 2,1 km

Urban Road 1,7 km
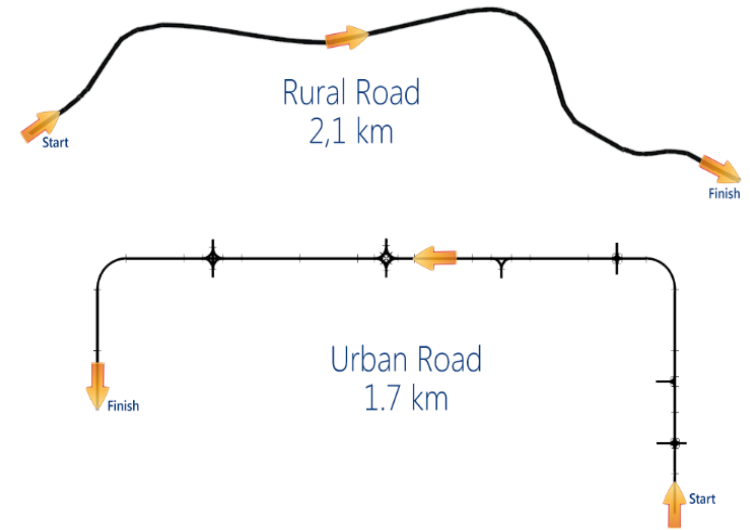
	Urban area		Rural area	
	Q_L	Q_H	Q_L	Q_H
No distraction	√	√	√	√
Cell phone use	√	√	√	√
Conversation with the passenger	√	√	√	√

Traffic scenarios

- Low traffic
  Q_L=300 vehicles/hour
- High traffic
  Q_H=600 vehicles/hour

Unexpected incidents

- Child crossing the road
- Sudden appearance of an animal

Randomisation

- The purpose of randomisation is to remove bias and other sources of extraneous variation, which are not controllable

Familiarisation

During the **familiarization** with the simulator, the participant practiced in:
- handling the simulator (starting, gears, wheel handling etc.)
- keeping the lateral position of the vehicle
- keeping stable speed, appropriate for the road environment
- Braking and immobilization of the vehicle

During this practice drive, two **unexpected incidents** took place.

The following criteria must be verified (there is **no time restriction**) before the participant moves on to the next phase of the experiment:

Questionnaires

Driving behaviour questionnaire
- Driving experience - car use
- Self -assessment of the older driver
- Distraction-related driving habits
- Emotions and behaviour of the driver
- Anger expression inventory during driving
- History of accidents, near misses, and traffic violations

Self-Assessment and memory questionnaire
- Memory
- Self assessment
- Driving skills

Sample characteristics

The sample of the analysis consists of **95 participants**

- 28 young drivers aged 18-34 years old

- 31 middle aged drivers aged 35-54 years old

- 36 older driver aged 55-75 years old
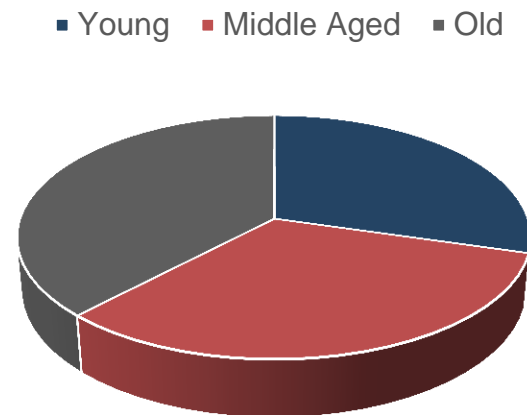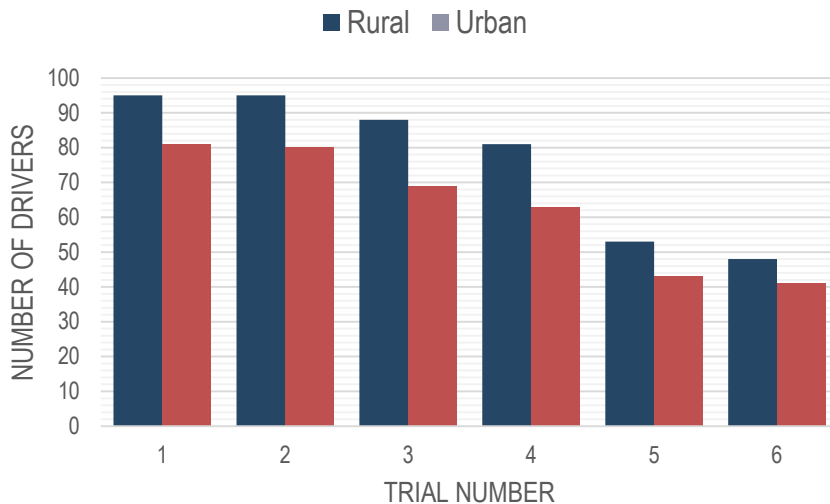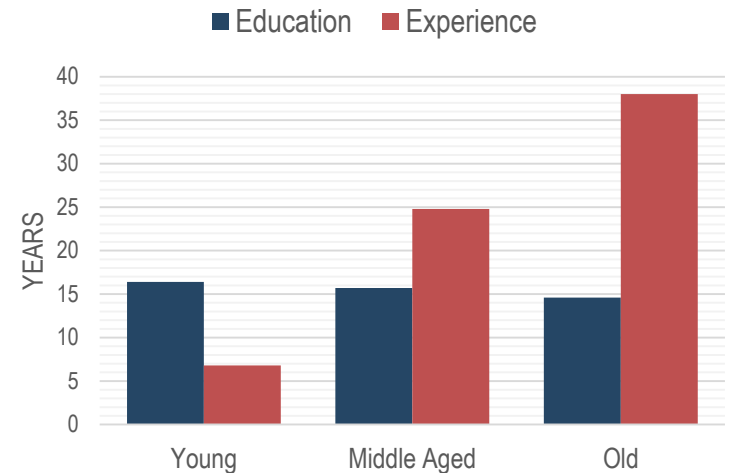
Statistical analysis methodology

Data collected from the driving simulator experiment and the respective questionnaires are analysed by means of a **dedicated statistical analysis method**:

1. **Descriptive analysis**
   (correlation table, boxplots)

2. **Regression analysis**
   (6 general linear mixed models)

3. **Factor Analysis**
   (2 factor analysis)

4. **Latent analysis**
   (4 structural equation models)

|	Variable	Explanation
1	Time	current real-time in milliseconds since start of the drive.
2	x-pos	x-position of the vehicle in m.
3	y-pos	y-position of the vehicle in m.
4	z-pos	z-position of the vehicle in m.
5	road	road number of the vehicle in [int].
6	richt	direction of the vehicle on the road in [BOOL] (0/1).
7	rdist	distance of the vehicle from the beginning of the drive in m.
8	rspur	track of the vehicle from the middle of the road in m.
9	ralpha	direction of the vehicle compared to the road direction in degrees.
10	Dist	driven course in meters since begin of the drive.
11	Speed	actual speed in km/h.
12	Brk	brake pedal position in percent.
13	Acc	gas pedal position in percent.
14	Clutch	clutch pedal position in percent.
15	Gear	chosen gear (0 = idle, 6 = reverse).
16	RPM	motor revolution in 1/min.
17	HWay	headway, distance to the ahead driving vehicle in m.
18	DLeft	distance to the left road board in meter.
19	DRight	distance to the right road board in meter.
20	Wheel	steering wheel position in degrees.
21	THead	time to headway, i. e. to collision with the ahead driving vehicle, in seconds.
22	TTL	time to line crossing, time until the road border line is exceeded, in seconds.
23	TTC	time to collision (all obstacles), in seconds.
24	AccLat	acceleration lateral, in m/s^2
25	AccLon	acceleration longitudinal, in m/s^2
26	EvVis	event-visible-flag/event-indication, 0 = no event, 1 = event.
27	EvDist	event-distance in m.
28	ErrINo	number of the most important driving failure since the last data set
29	ErrIVal	state date belonging to the failure, content varies according to type of failure.
30	Err2No	number of the next driving failure (maybe empty).
31	Err2Val	additional date to failure 2.
32	Err3No	number of a further driving failure (maybe empty).
33	Err3Val	additional date to failure 3.

Descriptive analysis

- **Database** development
  - Type of variable
  - Min, max, average value

- Several **boxplots** were developed in order to explain the effect of specific driver, road and traffic parameters as well as the examined distraction sources on selected driving performance measures

- A **correlation table** is investigating any of a broad class of statistical relationships between driving simulator variables
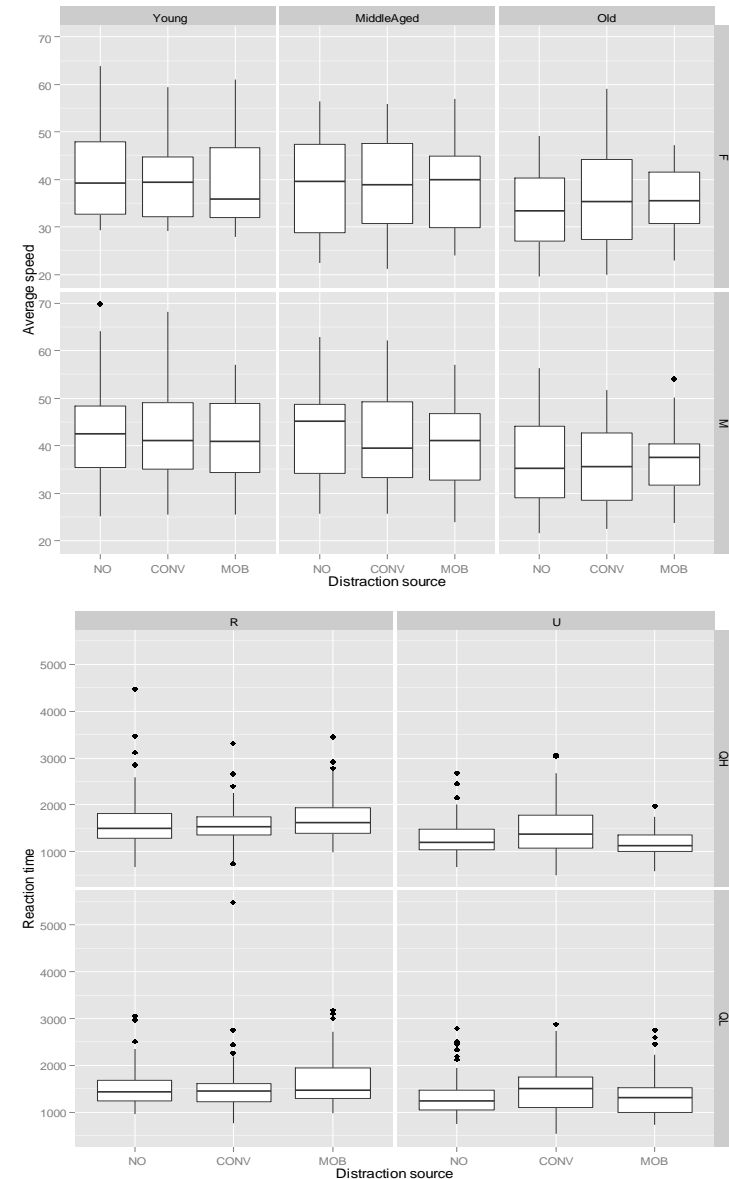
Regression analysis

- Within the framework of regression analysis, **6 general linear mixed models** are developed in order to identify several sets of explanatory variables that covary with specific driving performance measures of the driving simulator dataset.

- Average speed

- Reaction time

- Lateral position

- Average headway

- Speed variability

- Lateral position variability
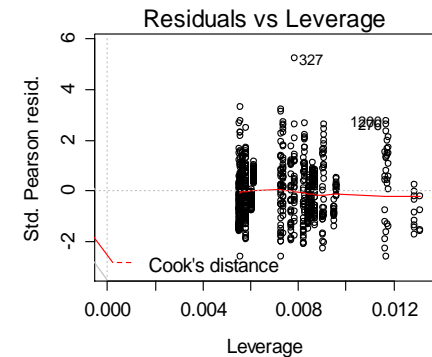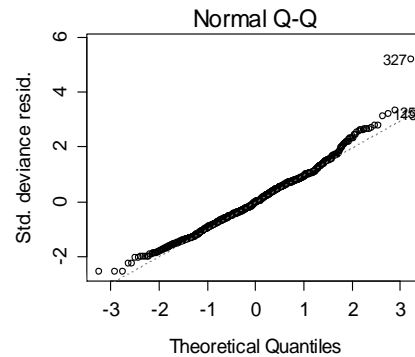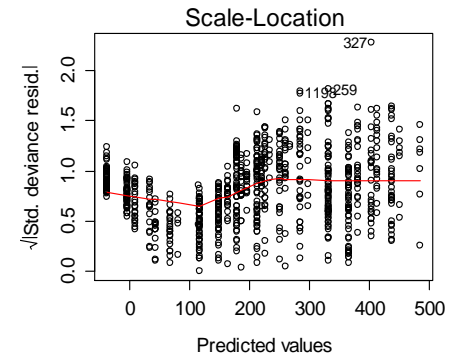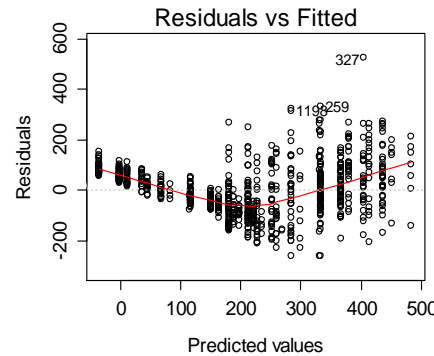
Factor analysis

- Two factor analysis are developed in order to investigate which observed variables are most highly correlated with the common factors of **driving performance** and **driver error** and how many common factors are needed to give an adequate description of the data

- Regarding **driving performance**, 5 factors are best fitted in the specific database. The interpretation of the results revealed that the five factors are: lateral measures, speed measures, vehicle direction measures, headway as well as vehicle revolation

- The variables that tend to explain better the "**Driver Error**" factor are: numbers of Outside Road Lines, Sudden Brakes and High Rounds per Minute

Structural Equation Models (1/2)

- Structural Equation Modeling is a very general, powerful **multivariate analysis** technique that includes several analysis methods

- SEM involves the evaluation of two models:

- **Measurement Model**
  - The part of the model that relates indicators to latent factors
  - The measurement model is the factor analytic part of SEM

- **Path model**
  - This is the part of the model that relates variable or factors to one another (prediction)
  - If no factors are in the model then only path model exists

Structural Equation Models (2/2)

Path diagram

- Latent variables

- Observed variables

- Single-headed arrow $\rightarrow$
  Regression Coefficient or factor loading

- Double headed arrow $\leftrightarrow$
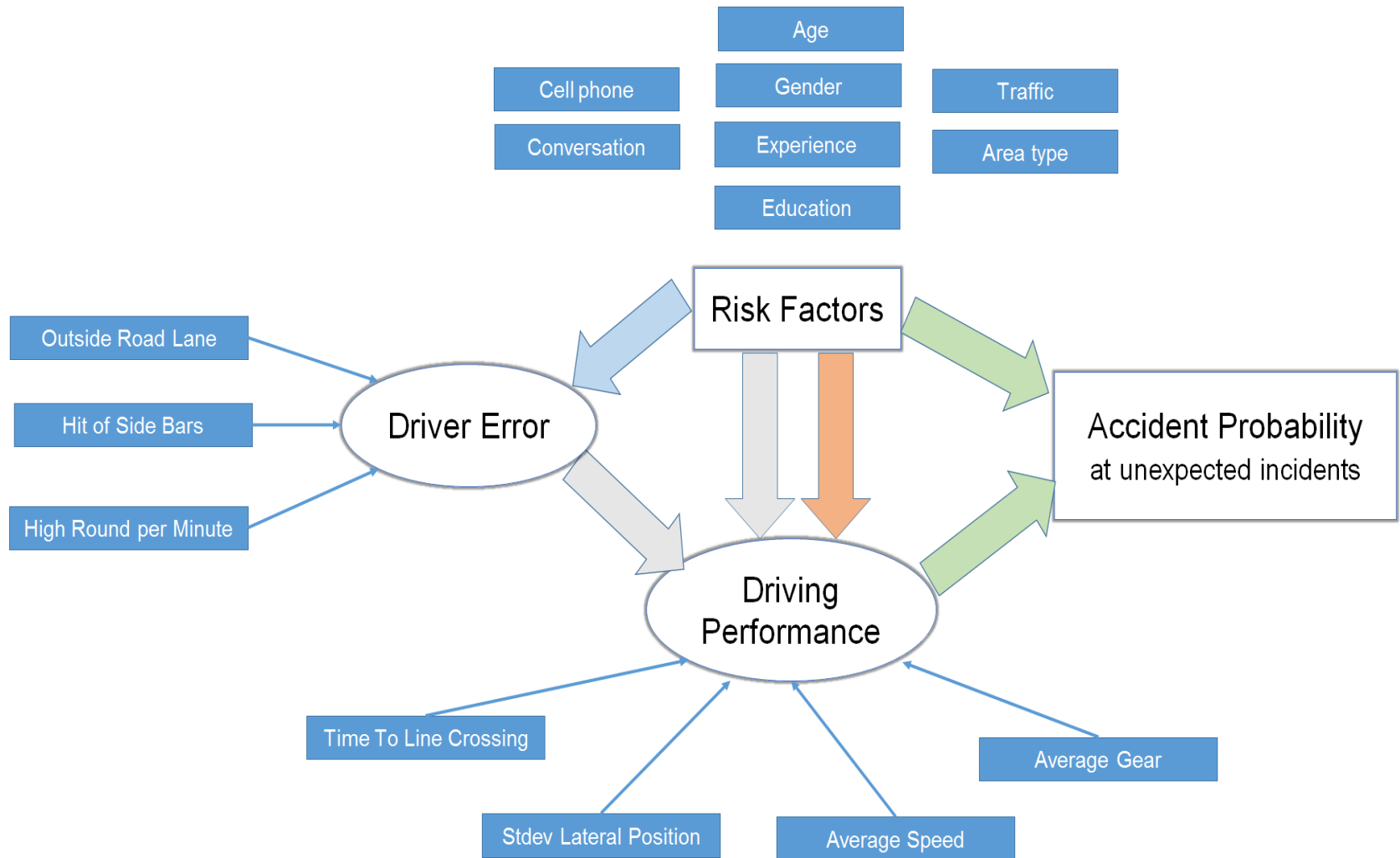 Correlations
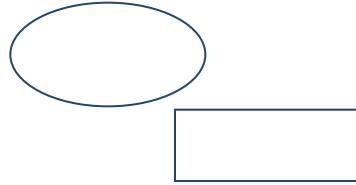
Goodness-of-fit measures

- Standardized Root Average Square Residual (SRMR) < 0,08

- Root Average Square Error of Approximation (RMSEA) < 0,08

- Comparative Fit Index (CFI) > 0,90

- Tucker Lewis Index (TLI) > 0,90

Latent analysis overview

The latent variable reflects the underlying **driving performance** and the objective is the quantification of the impact of distraction, driver characteristics as well as road and traffic environment on driving performance
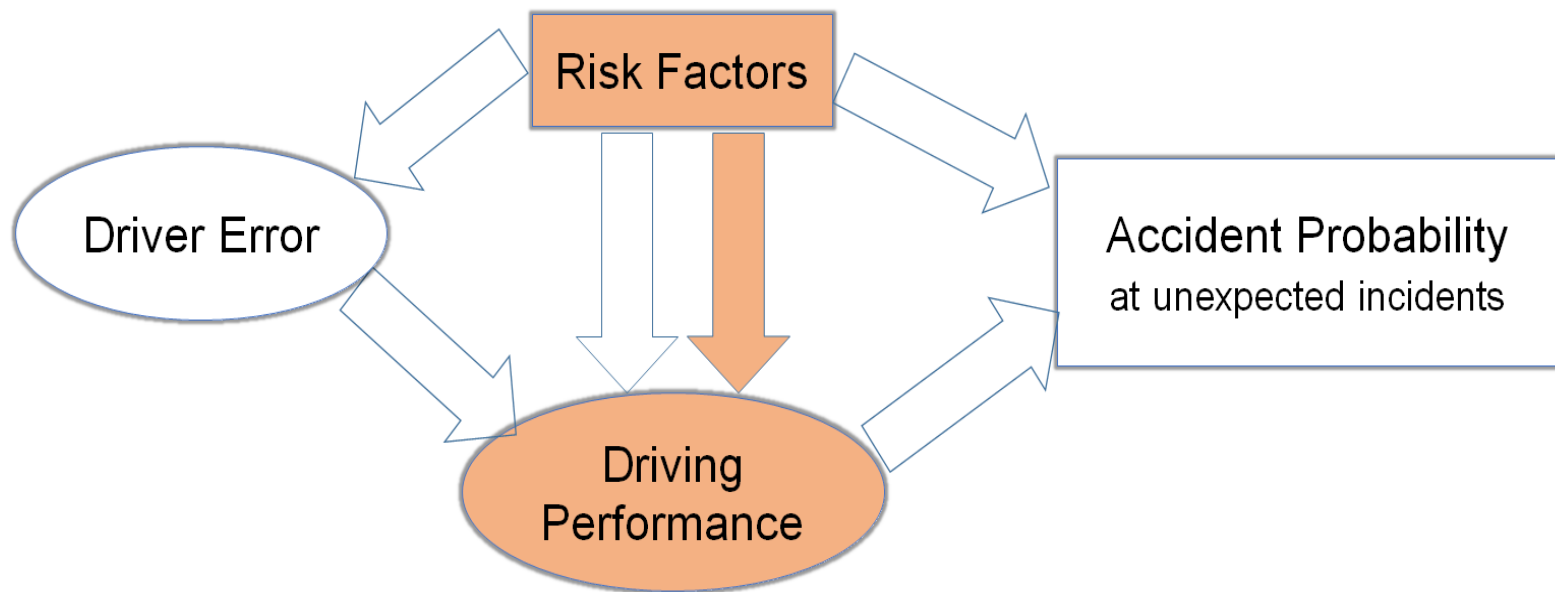
	Est.	Std.err	t value.	P(>|z|)
Latent Variable				
Driving Performance				
Average Speed	1,000	-	-	-
Stdev Lateral Position	-0,085	0,004	-23,909	0.000
Average Gear	0,048	0,002	21,887	0.000
Time to Line Crossing	-0,109	0,005	-19,972	0.000
Regressions				
Driving Performance				
Distraction – Cell phone	-1,099	0,342	-3,213	0.001
Area - Urban	-15,596	0,467	-33,410	0.000
Traffic - Low	1,123	0,285	3,943	0.000
Gender - Female	-1,154	0,303	-3,802	0.000
Age	-0,155	0,027	-5,755	0.000
Experience	0,083	0,032	2,630	0.009

Summary statistics

Minimum Function Test	305,74
Degrees of freedom	20

Goodness of fit

| SRMR	0,061
RMSEA	0,136
CFI	0,867
TLI	0,809
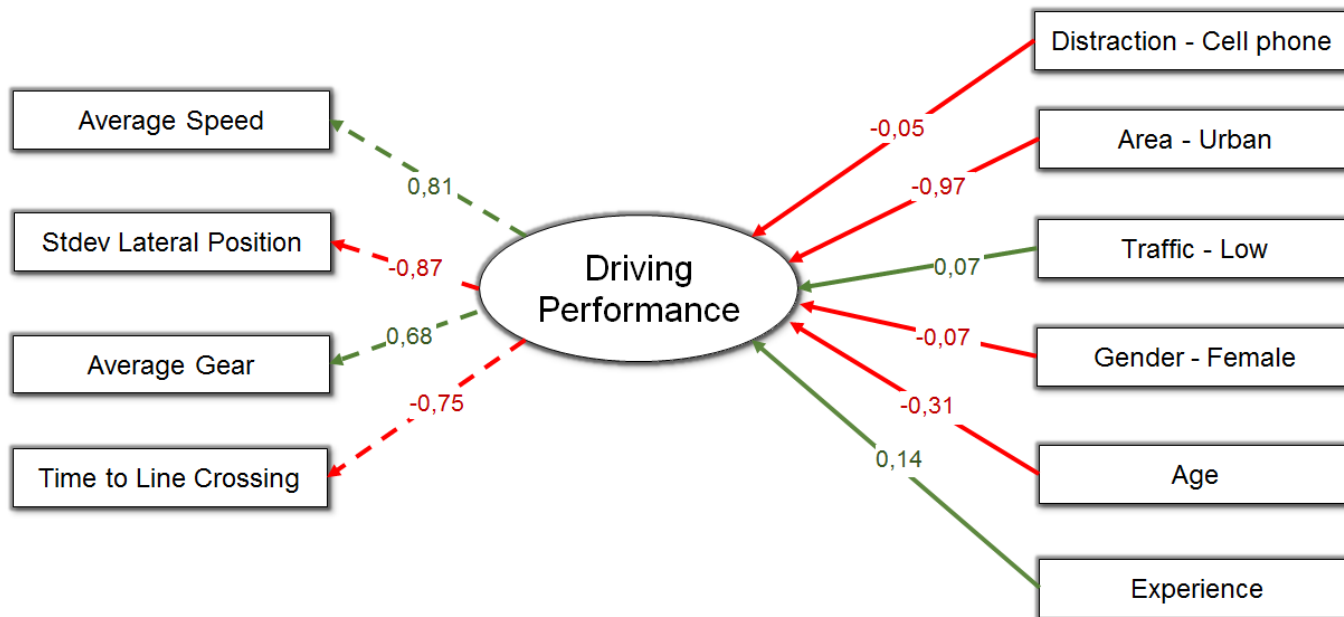
- The effect of **cell phone** on driving performance is definitely negative
- **Conversation with the passenger** does not has a statistically significant effect
- Risk factors that affect driving performance include driver characteristics (**age**, **gender**, driving **experience**), **area** type and **traffic conditions**

The latent variable reflects the underlying **driver error** and the objective is the quantification of the impact of distraction, driver characteristics as well as road and traffic environment on driving error
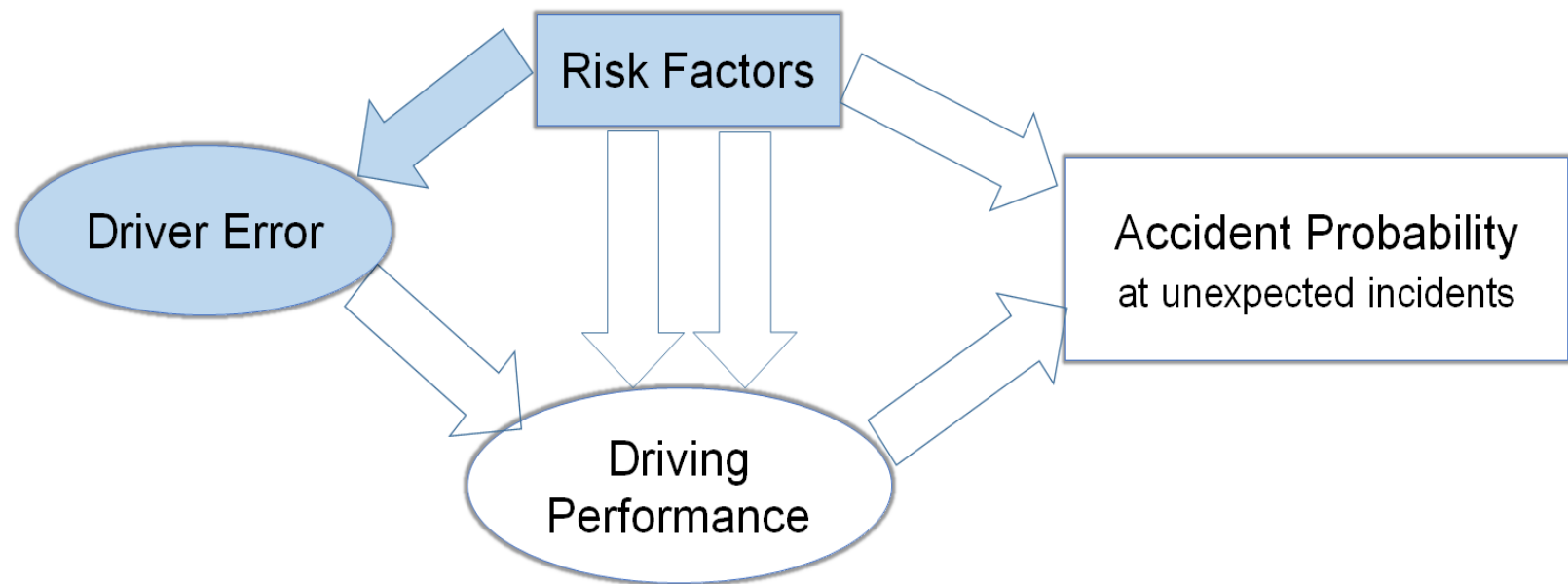
SEM regarding driver error (2/3)

	Est.	Std.err	t value.	P(>|z|)
Latent Variable				
Driving Error				
Hit Of Side Bars	1,000	-	-	-
Outside Road Lanes	0,741	0,257	2,887	0,004
High Rounds Per Minute	0,680	0,243	2,803	0,005
Regressions				
Driver Errors				
Gender - Female	0,359	0,076	4,739	0.000
Age	0,031	0,009	3,393	0.001
Area - Urban	-0,393	0,062	-6,383	0.000
Experience	-0,030	0,010	-3,050	0.002
Education	-0,021	0,010	-2,167	0.030
Summary statistics				
Minimum Function Test	62,19			
Degrees of freedom	10			
Goodness of fit				
SRMR	0,032			
RMSEA	0,096			
CFI	0,823			
TLI	0,682			

Risk Factors

Driver Error

Accident Probability
at unexpected incidents

Driving
Performance
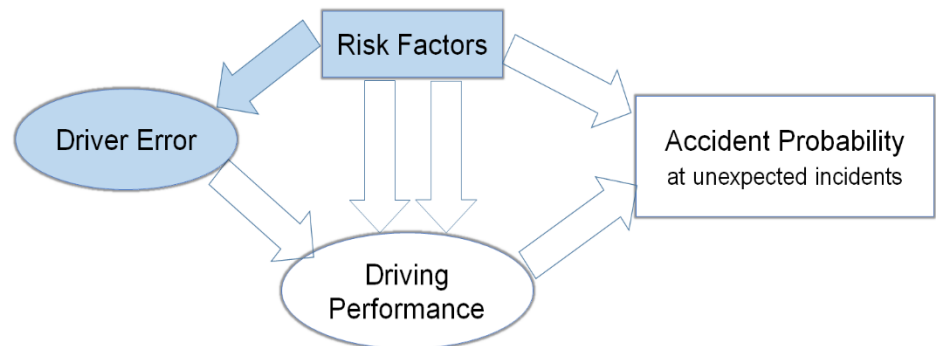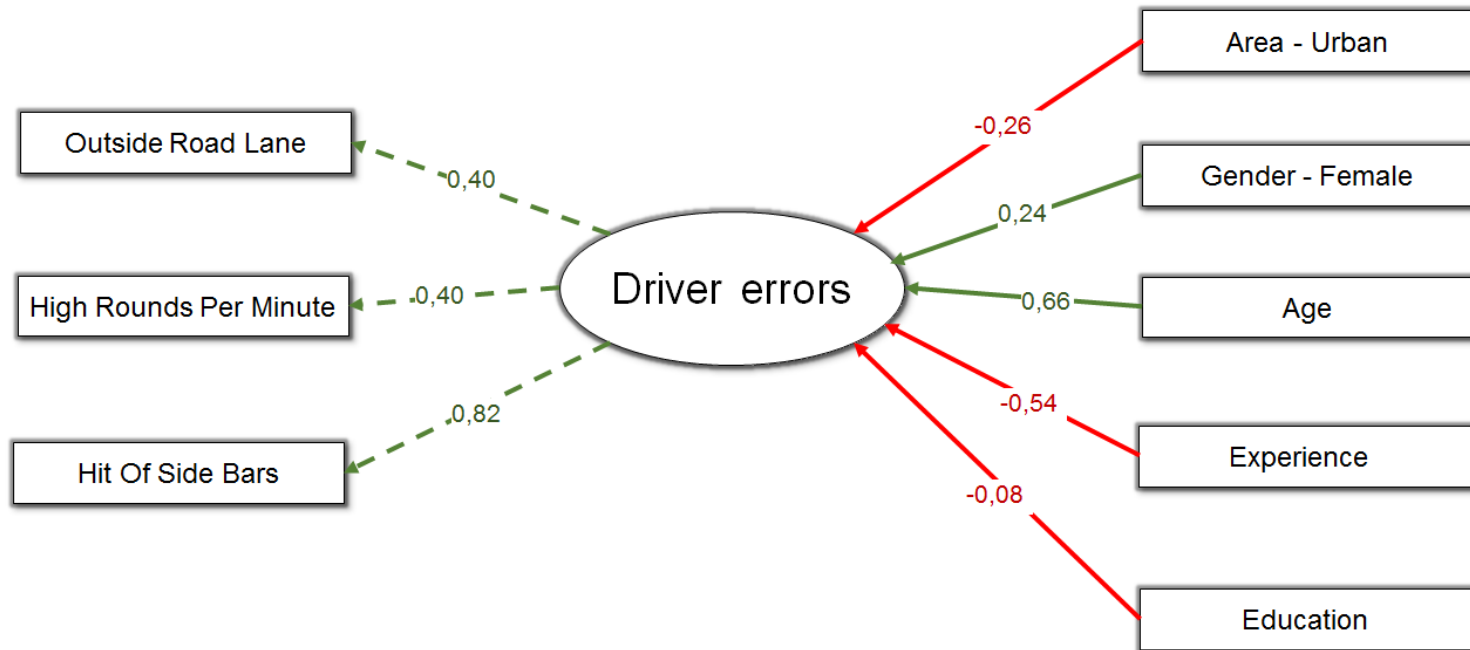
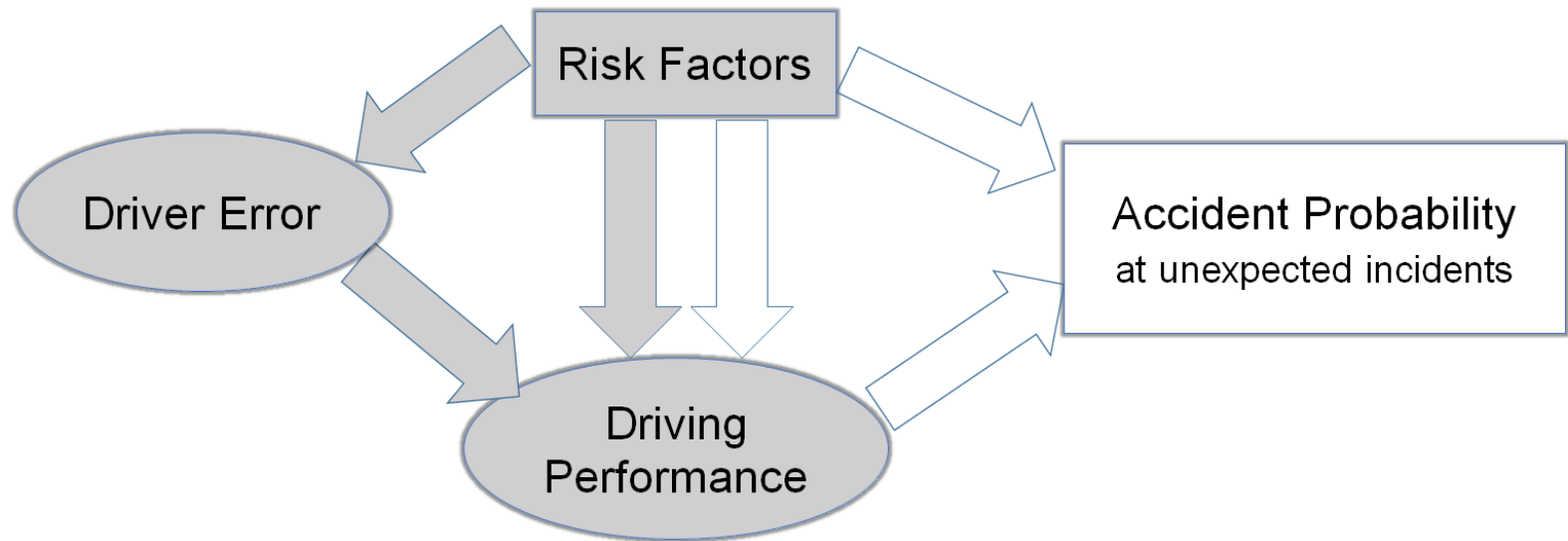- Neither conversing with a passenger nor talking on the cell phone has a statistical significant impact on driver error
- Risk factors that affect driver error include **gender**, **age**, **experience**, **education** and **area type**

Two latent variables are created regarding driving performance and driver error while the objective of this analysis is the quantification of the impact of driving errors, distraction, driver characteristics as well as road and traffic environment on driving performance

	Est.	Std.err	t value.	P(>|z|)
Latent Variable 1				
Driver Errors				
Hit Of Side Bars	1,000	-	-	-
Outside Road Lanes	0,547	0,214	2,559	0,010
High Rounds Per Minute	0,950	0,276	3,436	0,001
Latent Variable 2				
Driving Performance				
Average Speed	1,000	-	-	-
Stdev Lateral Position	-0,085	0,004	-23,117	0,000
Average Gear	0,049	0,002	22,043	0,000
Average TTL	-0,108	0,005	-20,114	0,000
Regression 1				
Driving Performance				
Driver Errors	-51,016	11,417	4,468	0.000
Gender – Female	-16,739	3,799	-4,407	0.000
Age	-2,244	0,681	-3,297	0.001
Experience	2,103	0,694	3,031	0.002
Regression 2				
Driver Errors				
Gender - Female	0,311	0,076	4,068	0.000
Age	0,042	0,010	4,125	0.000
Area - Urban	-0,300	0,068	-4,395	0.000
Experience	-0,040	0,011	-3,815	0.000
Education	0,004	0,001	3,174	0.002

Summary statistics

Minimum Function Test	608,01
Degrees of freedom	40

Goodness of fit

|	
SRMR	0,088
RMSEA	0,158
CFI	0,793
TLI	0,711
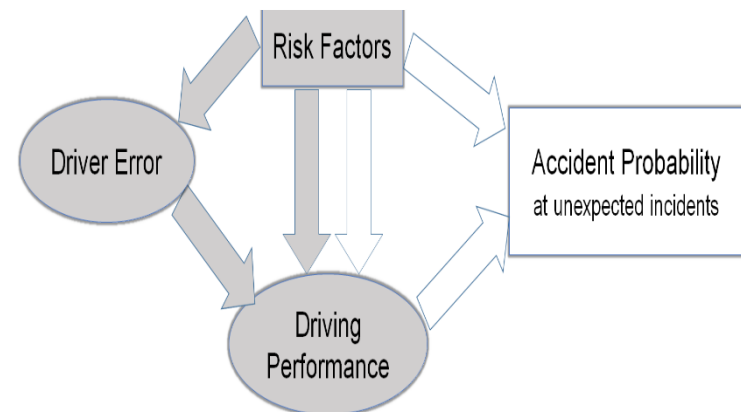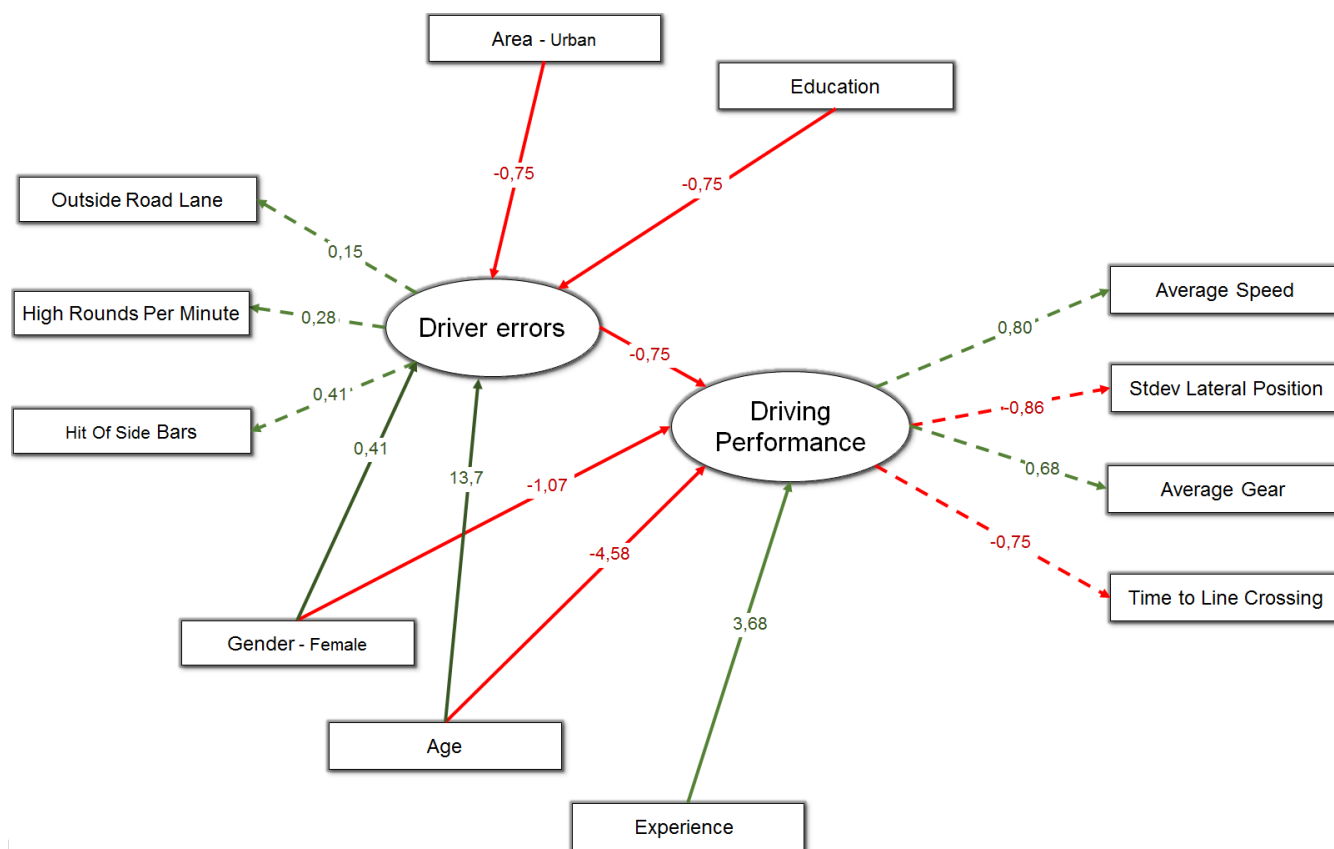
- Driver error is a crucial factor that negatively affects driving performance
- Neither **road characteristics** (area type, traffic conditions) nor the **distraction sources** examined (cell phone use, conversation with a passenger) have a significant impact on this model

The latent variable reflects again the underlying driving performance of the participants and the objective is the quantification of the impact of driving performance, distraction, driver characteristics as well as road and traffic environment directly on **accident probability at unexpected incidents**
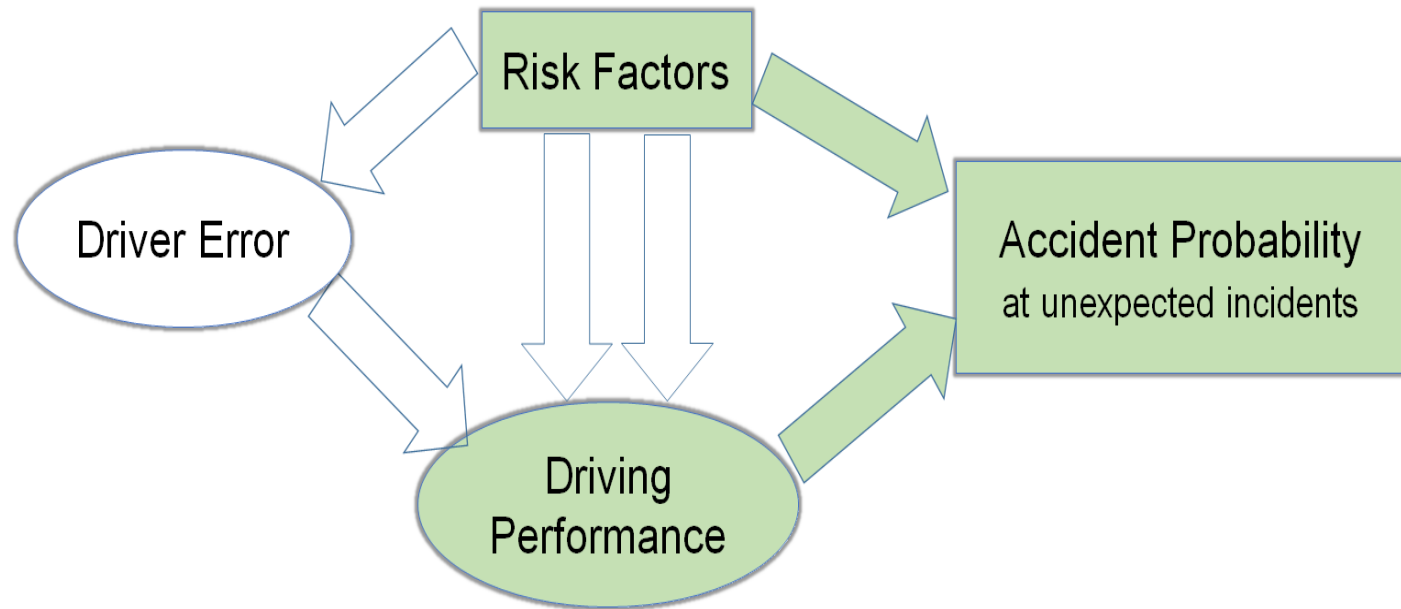
	Est.	Std.err	t value.	P(>|z|)
Latent Variable				
Driving Performance				
Average Speed	1,000	-	-	-
Stdev Lateral Position	-0,085	0,004	-23,803	0.000
Average Gear	0,048	0,002	21,836	0.000
Average TTL	-0,109	0,005	-20,046	0.000
Regression				
Accident				
Driving Performance	-0,007	0,002	-3,119	0.002
Gender - Female	0,074	0,034	2,198	0.028
Traffic – Low	0,104	0,033	3,142	0.002
Distraction – Cell phone	0,081	0,033	2,463	0.014
Regression				
Driving Performance				
Gender - Female	-1,147	0,307	-3,737	0.000
Area - Urban	-15,614	0,468	-33,386	0.000
Distraction – Cell phone	-1,099	0,343	-3,208	0.001
Traffic - Low	1,131	0,286	3,956	0.000
Age	-0,156	0,028	-5,593	0.000
Experience	0,083	0,032	2,557	0.011

Summary statistics

Minimum Function Test	352,62
Degrees of freedom	31

Goodness of fit

SRMR	0,061
RMSEA	0,136
CFI	0,867
TLI	0,807
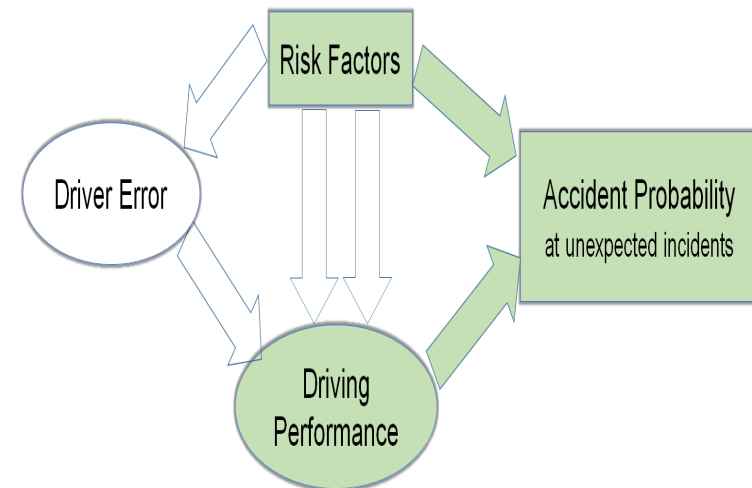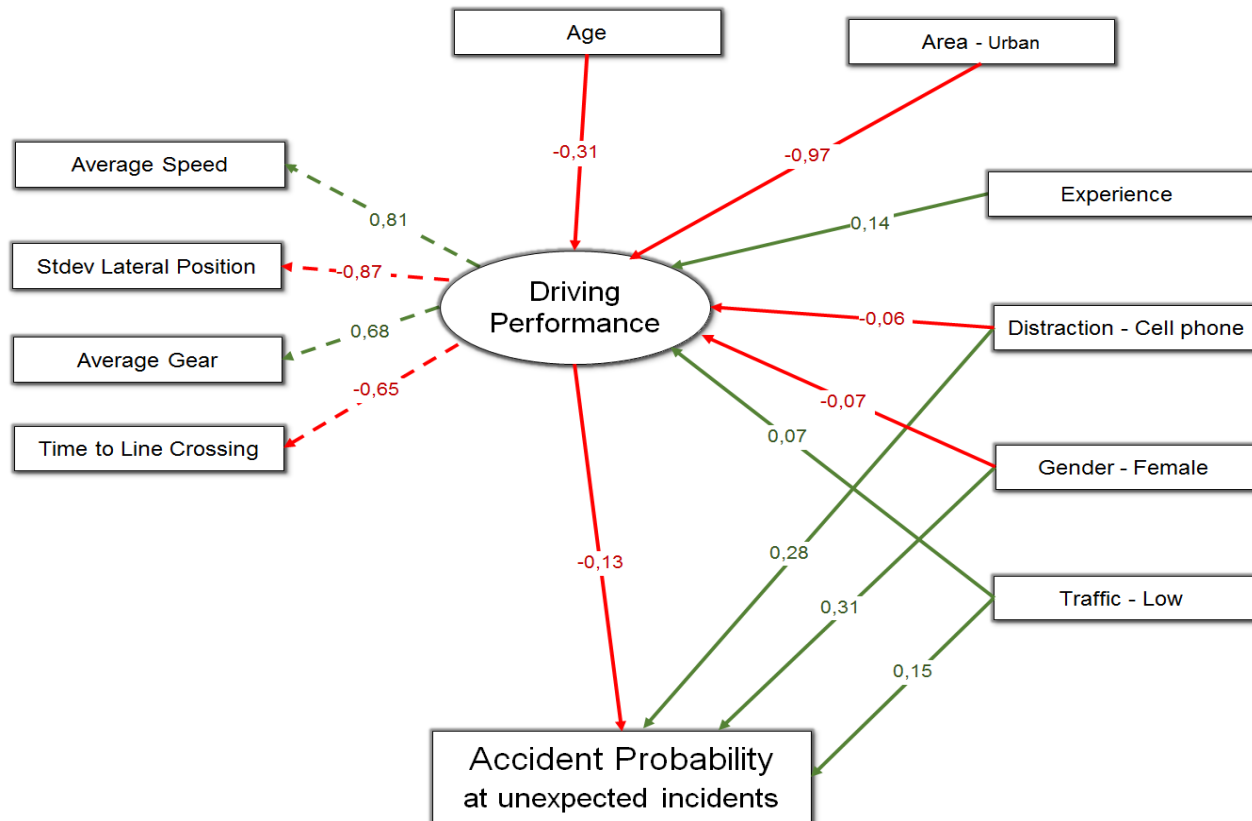
- **Cell phone** use has a negative effect on accident probability
- Drivers self-regulate their driving performance better while conversing with a passenger
- **Female** drivers at **low traffic** are more prone to accidents at unexpected incidents

Scientific contributions

Risky driving profiles

Estimation of driving performance and safety

Innovative statistical analysis methodology

Large and rigorous driving simulator experiment
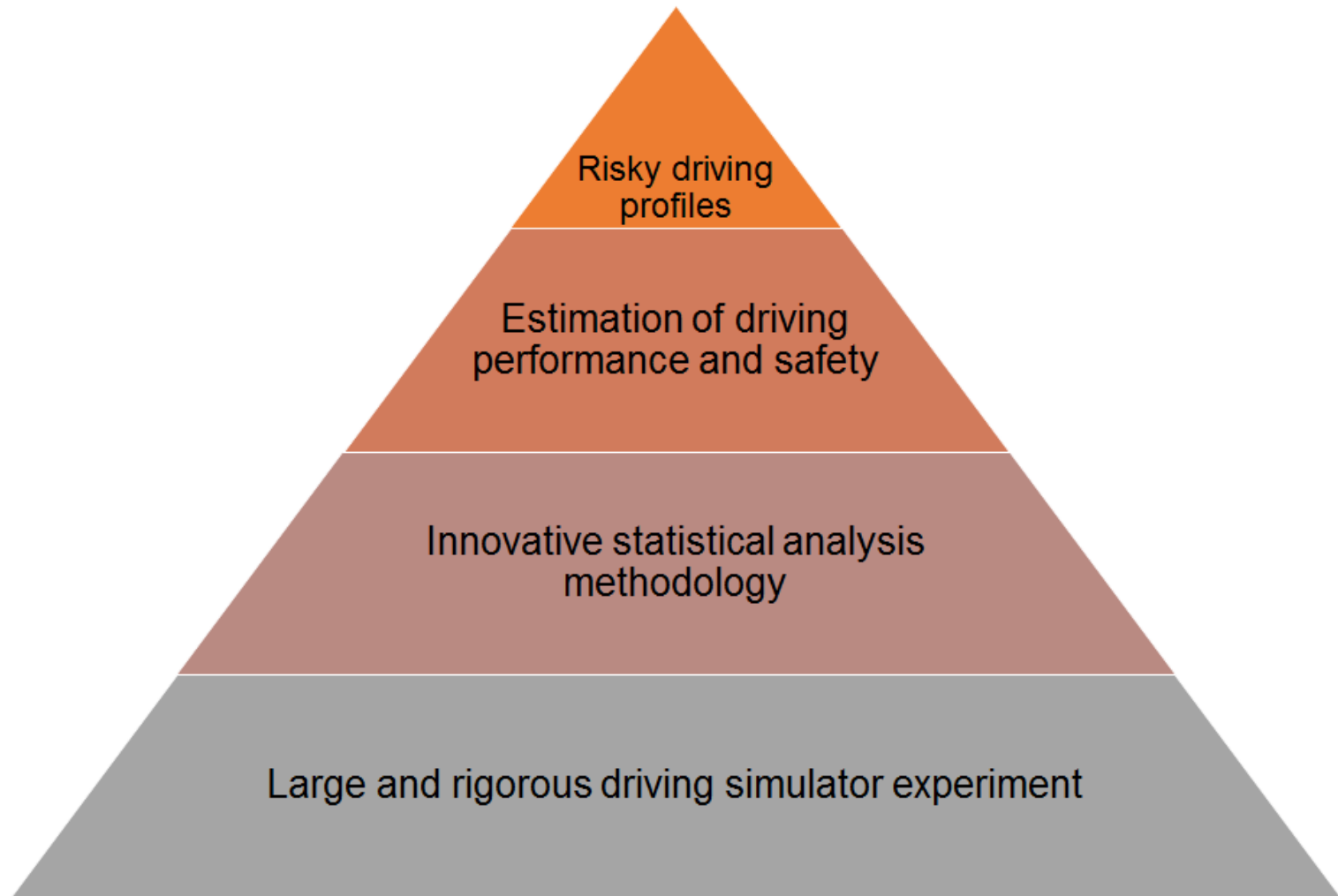
Methodological contributions (1/2)

Design and implementation of a large and rigorous driving simulator experiment

The basic limitations found in the literature that the present experiment tackled are the following:

- Large and representative sample
- Randomisation of trials
- Adequate practice drive
- Investigation of an optimum number of driving factors
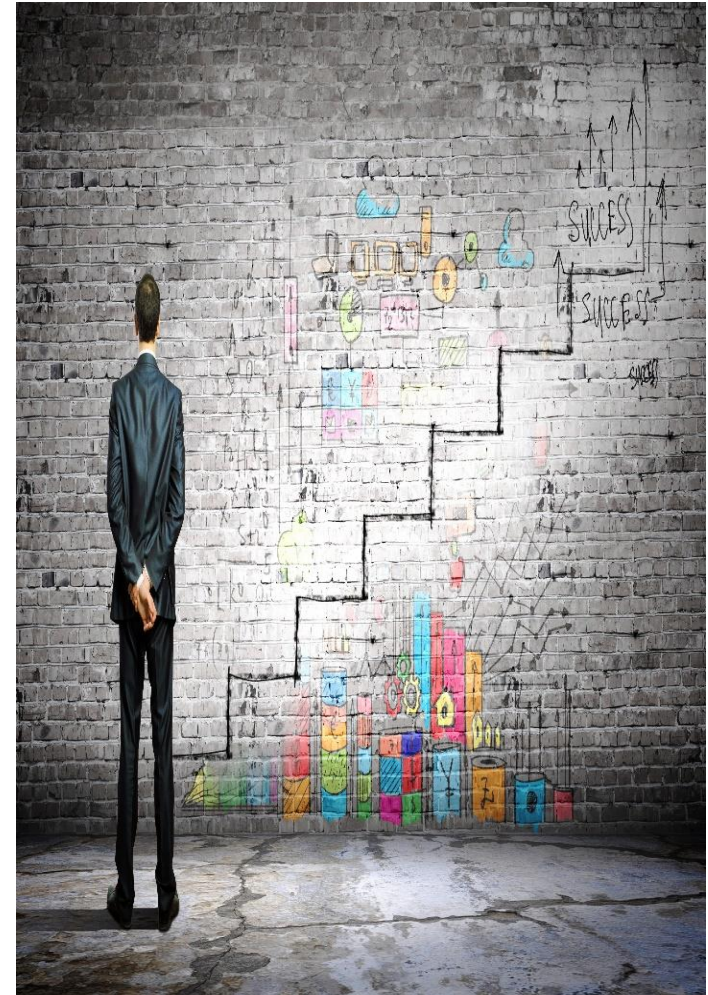
Methodological contributions (2/2)

- Development and application of an **innovative** statistical analysis methodology

- **Latent analysis** through Structural Equation models is implemented for the first time in the field of driving performance and traffic safety

- Estimation of the **combined effect** of distraction sources, driver as well as road and traffic environment characteristics directly on driving performance

Key research findings (1/2)

- Results regarding the effect of **driver distraction** indicate the different effect on driving performance between cell phone use and conversation with the passenger

- **Driver characteristics** play the most crucial role in driving performance (gender, age, experience)

- Driving performance is worst in urban areas and high traffic conditions probably due to the complex **driving environment**

Development of risky driver profiles regarding driver error and accident probability at unexpected incident. Results indicate that:

- more likely to commit driving errors are **young or old female drivers at urban areas**

- more likely to be involved in an accident at an unexpected incident are **female drivers in low traffic conditions while talking on the cell phone**

Further research

- Investigation of the effect of other parameters such as **alcohol**, **fatigue** etc. on driving performance through latent analysis

- Development of Structural Equation Model on different experimental methods (**Naturalistic experiments, field test etc.**)

- Further investigation of the parameters that affect the **compensatory** behaviour of the driver

- Investigation of different **types of cell phone** use such as a hands-free, Bluetooth, typing an sms etc.)

Risk factors, driver behaviour and accident probability. The case of distracted driving.

Bar-Ilan University
Tradition of Excellence

Panagiotis Papantoniou
PhD, Research Associate

National Technical University of Athens

Tel Aviv, July 2016