

COST-329 Workshop on Models for Traffic and Safety  
Development and Interventions

**DATA REQUIREMENTS IN RELATION TO  
CARE AND TIME SERIES ANALYSIS**

*Dr. George Yannis*

Department of Transportation Planning and Engineering  
National Technical University of Athens

Brussels, 10 June 1999

# THE QUESTIONS

*Do we have the data we need?*

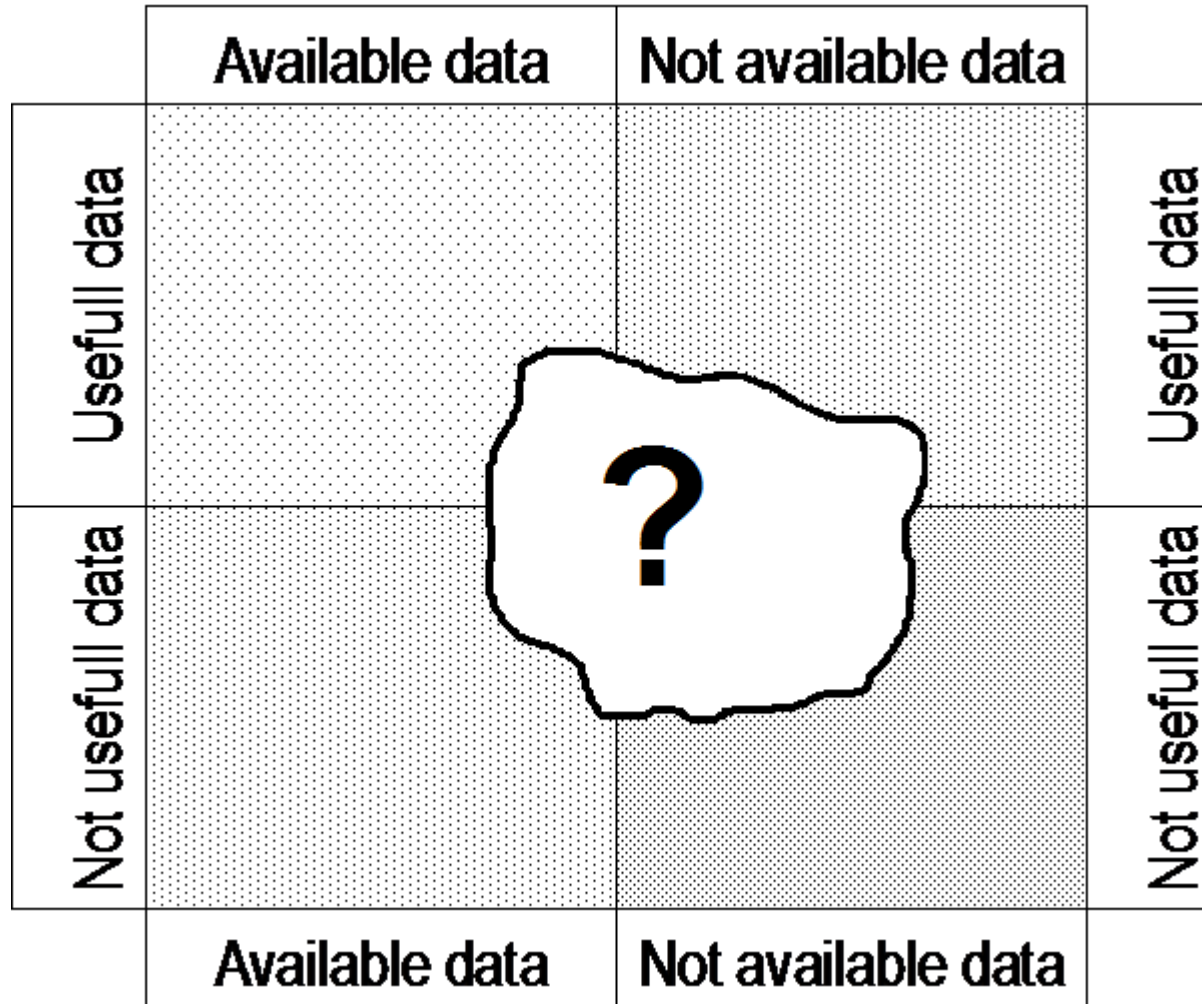
*Do we need the data we have?*

Road accident analysis at international level

Cross-country comparisons

Exchange of experience

# USEFULNESS & AVAILABILITY OF DATA



# **METHODOLOGY**

## **COST 329 Surveys**

accident analysis priorities and necessities  
data availability and usefulness

## **Experience from the development of international data files**

the CARE database and the CAREPLUS project  
other data files (IRTAD, ECMT, UN, EUROSTAT, IRF)

## **Experience of COST-329 experts**

analyses at national level  
international research

# PRIORITIES AND NECESSITIES FOR TIME SERIES ANALYSIS

- multiple aims: monitoring and prognosis of traffic safety, international comparisons
- basic criterion: fatalities
- type of variables: number of casualties, number of accidents
- disaggregation: by country
- types of model types: the descriptive/prognostic, causal
- time basis: yearly or monthly
- time horizon for prognosis: 1-3 years or 4-10 years
- basic additional variables: exposure (veh-kms)

# ABOUT ROAD ACCIDENT DATA

- Data concerning accident variables are in general available and sufficiently disaggregate for model-based road accident analysis
- Data availability is in general daily and the average starting year is in the early seventies
- Basic data quality problems refer to definitions (killed 30-days, hospitalised) and under-reporting
- Usefulness of accident variables is in general acceptable

# ABOUT EXPOSURE (TRAFFIC) DATA

- A combination of periodical and extraordinary traffic counts and surveys is used for the estimation of the number of vehicle-kilometers (and passenger-kilometers) and their characteristics (by road and vehicle type and by road user type).
- The basic data insufficiencies comprise
  - poor availability
  - poor reliability
  - incomparability
  - inappropriate disaggregation

# ABOUT ADDITIONAL DATA

- **additional exposure data**

vehicle fleet by type and age, population by age and sex, road network length by type, fuel sales, drivers, etc.

- **weather conditions**

temperature, rain, sunshine, snow/ice

- **economic factors' data**

fuel and consumer price index, wage index, GNP/GDP, unemployment, alcohol consumption per capita, road construction and maintenance budget, road safety budget, etc.

- **safety factors' data**

drinking and driving, seat belt and helmet use by road user type, average speed by road type, etc.



# DEALING WITH LACK OF EXPOSURE DATA

- **Absolute numbers and trends**  
(limited use)
- **Severity Indices**  
(nr. of persons killed per 100 persons injured, etc.)
- **Induced exposure**  
(the driver of the 2nd vehicle is presumed not responsible)
- **Percentages related to accident type**  
(accident and manoeuvre type)

# CONCLUSION

- Today, model-based road accident analysis at European level has a great potential
  - priorities and necessities are well defined and converging
  - a lot of useful data variables are available
- The more the data are useful, the more difficult is to find them
- Further work and research should focus on data compatibility and availability, especially for traffic data (pan-european data collections systems and surveys)