CHAPTER 12

Risk Evaluation and Road Safety

Sylvain Lassarre

GRETTIA, in Components and Systems at IFSTTAR, the French institute of science and technology for transport, spatial planning, development and networks, France

CONTENTS

ABSTRACT

Risk can be defined, following the reliability and epidemiology literature, as a probability of sustaining an injury more or less severe in a road accident. The main risk indicators used in public health and transportation, such as the mortality and fatality rates, are described with their theoretical background from the survival theory, and their common estimates based on accident, injury, and mobility data. The factors affecting the risk indicators: demographic, geographical, mobility, and economic, are reviewed with their main effects, especially on the mortality rate with the relevant statistical risk models, including some advanced statistical models useful to tackle some deficiencies of the classical models such as Smeed's law or Kopits World Bank model.

12.1 INTRODUCTION

We live in a risky society. What does it imply? It implies that as a road user we are subject to a potential adversity or threat, the vectors of which are the moving vehicles on the road. The accident can occur as an unwanted and unexpected event, such as a collision or a crash, which has consequences such as injuries because of the release of mechanical energy.

Risk has different meanings; here we treat risk as a probability of an adverse event leading to consequences. In this chapter, we are more interested in the objective rather than the subjective

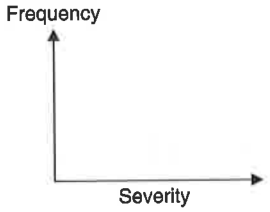

Figure 12.1 The two dimensions of risk.

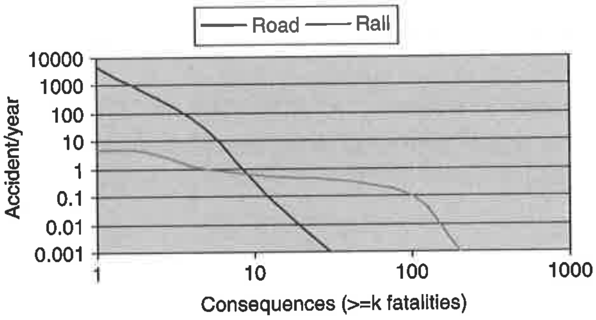

Figure 12.2 Farmer's curves for road and rail (source: Evans 1994). The scales are logarithmic. The y axis gives the frequency in number of accident/year and the x axis indicates the number of fatalities, and the curve indicates the number of accidents per year with more than a certain number of fatalities.

measure of the risk. According to this definition coming from epidemiology and reliability, risk has two dimensions: frequency/severity forming a plane in which accidents take place, some being very severe and less frequent, others rather frequent but slight.

On this risk surface, we can draw the Farmer's curve which gives the pattern of risk in the plane frequency/consequences. If the consequences are expressed as the number of fatalities per accident from 1 fatality to more than 100 fatalities per accident, in the case of a catastrophic accident, we can compare the risk curve of two systems of transport: road and rail. In the UK according to Evans (1994), the road risk is diffuse with a lot of fatal accidents with few deaths per accident. The rail risk is more concentrated on accidents with a high number of fatalities but with much less frequency. In the developing world, the road risk can be catastrophic, with accidents involving buses having more than 20 to 30 deaths in a single accident.

12.2 RISK INDICATORS IN PUBLIC HEALTH

We define the Individual Risk as the probability of occurrence of an adverse outcome during a stated period of time which leads to consequences such as death or injuries. This probability is related to the instantaneous failure rate, for a human an instantaneous death rate, for example by means of the hazard function

$$\text{Hazard function} = h(t) = \lim_{dt \to 0} \frac{P(\text{death between}[t,\ t + dt]/\text{live before } t)}{dt}$$

For a human, the death rate depends on age, according to a bath-tub curve. There is a strong risk at birth, then the risk decreases, then increases again at old age due to uncertainty. In the developed world, there is a peak around 20 years old due to road accidents, especially for men.

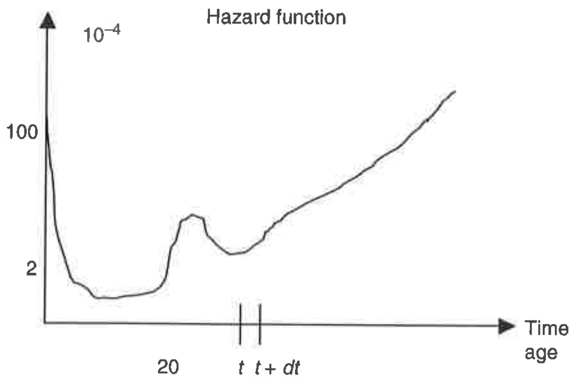

Figure 12.3 Death rate for a human in the western world (unit 1/10000 on the y axis).

From the cumulative failure (death) rate, which is equal to the integral of the instantaneous rate

$$\int_0^t h(u)du$$

we can calculate the risk of death during a certain duration or up to a certain age.

$$\text{Risk} = P(Y < t) = 1 - P(Y \geq t) = 1 - \exp(-\int_0^t h(u)du) \approx \int_0^t h(u)du$$

For example, the risk of dying between 20 and 21 for a young French male in the 1980s was equal to 2/1000, that is to say, two chances over one thousand to die.

$$P(\text{death}[20,\ 21]) = P(20 < Y \leq 21) \approx \int_{20}^{21} h(u)du = (21 - 20) \times \frac{2}{1000}$$

It is also possible to calculate by integration the life expectancy from this curve at different ages: at birth, at 20, 40, 60, ...

12.2.1 Mortality rate and number of years of life lost

When we select traffic accidents as the cause of death, we get the mortality rate due to road accidents, which has a peak around 20 years old with young males, and starts to increase with age due to the physical vulnerability of old persons in the case of a crash.

We used to characterize the risk in public health by the **Mortality rate** expressed as a number of fatalities/person $*$ year. It is estimated by the ratio

$$\frac{\text{Number of fatalities in a year}}{\text{Number of inhabitants exposed during a year}}$$

This mortality rate due to traffic accident is very low. In France in $2000 = 13.5 * 10^{-5}$, in $2003 = 10 * 10^{-5}$

These figures are based on an actuarial estimation relying on population statistics by gender and age and on statistics on fatalities due to traffic accidents. Such estimates are subject to the migration effect, as some foreigners may die on the road and some residents may die by accident abroad.

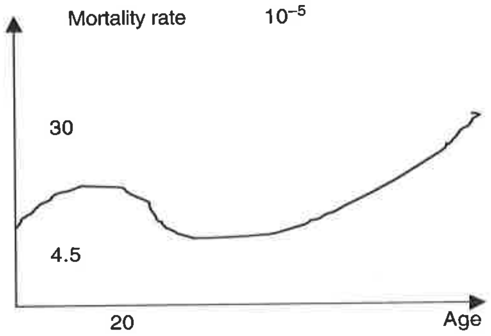

Figure 12.4 Mortality rate pattern due to traffic accidents in western world (unit 1/100000 on the y axis).

By aggregating individual risks, we obtain the burden as number of deaths or injuries that results from the exposure of the population to the risk of traffic accident during a period of one time unit, one year for example, as the summation over the whole population of n individuals of the indicator of death, which is a random variable taking the value 1 for death with probability $\lambda \times 1$ equal to the mortality rate over one year, and 0 with the probability $1 - \lambda \times 1$

$$N_{t,t+1} = \sum_{i=1}^{n} D_i\left([t, t+1]\right) \quad \begin{array}{l} P(D_i = 1 \text{ (death)}) = \lambda \times 1 \\ P(D_i = 0 \text{ (life)}) = 1 - \lambda \times 1 \end{array}$$

The collective risk, which is similar to the average burden, is equal to the product of the individual risk by the total exposure equal to the total number of individuals exposed during one year to risk

$$E\left(\sum_{i=1}^{n} D_i\right) = \sum_{i=1}^{n} E(D_i) = \lambda \times n \times 1$$

The burden can be expressed as a number of fatalities by using the mortality rate or by the number of injuries (severe and light) by using the morbidity rate.

When we compute the number of fatalities, each death has the same value. We can choose to put a value on each death according to some criteria. In public health, the value corresponds to the number of years of life lost because of premature death due to a traffic accident. Their summation leads to another risk indicator used in public health, the total **Number of years of life lost**. It is equal to the product of the life expectancy taken at the average age at death for traffic accident by the total burden. If the age at death for 910 killed pedestrians is 24 years and the life expectancy at 24 is 40 years, then the number of years of life lost is equal to $40 * 910 = 36400$ years. As young people are much more involved in fatal accident than adult or old people, the ranking of road accident as a cause of disease increases from rank 10 to 3 by using the number of years of life lost instead of the mortality rate.

This risk indicator is completed with the Number of years lived with disability which is estimated from the burden of injured people. When we add both indicators, we end with the Number of disability adjusted life years (= YLL + LLD). 1 DALY = Loss of 1 healthy year is the unit used in The Global burden of diseases by WHO.

12.2.2 Factors influencing the mortality rate

In epidemiology, we are interested in the variations of the accident risk according to age and sex to explore the differences between male and female and between young and old, as well as the spatial and temporal distribution of the accident risk among different spatial units.

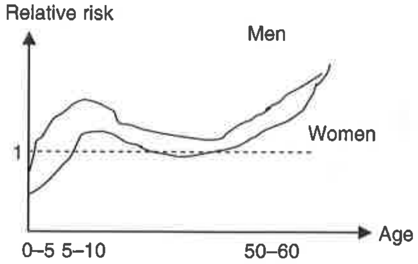

Figure 12.5 Mortality ratio (reference male 0–5 years old) of pedestrian according to age and sex.

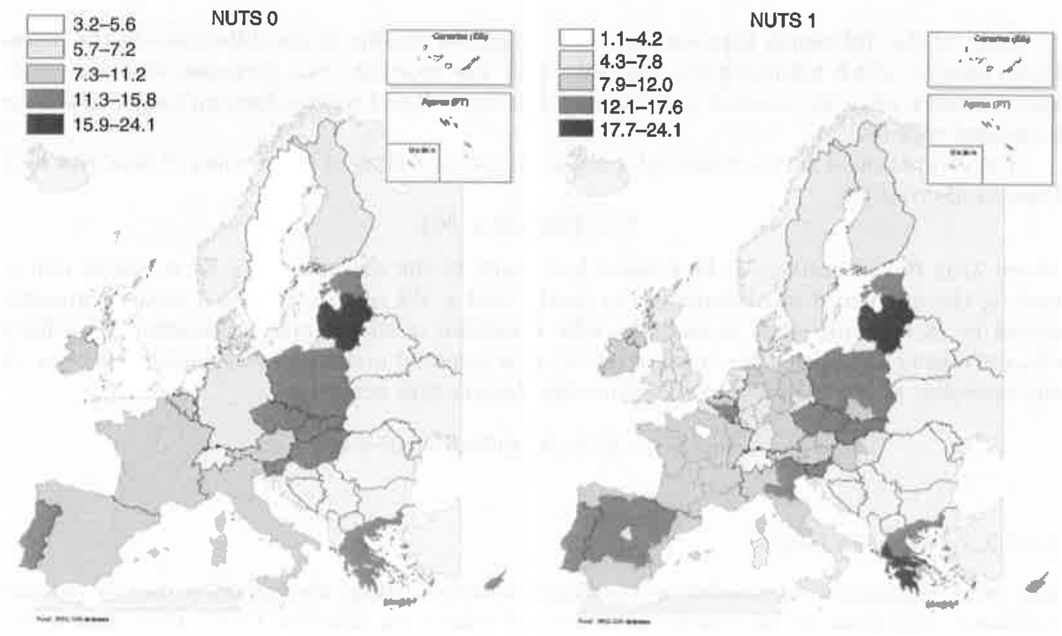

Figure 12.6 Mortality rate in Europe (source: Eksler et al 2008) (unit 1/100000).

12.2.2.1 Demographic factors

The mortality rate varies according to gender and age. The patterns of variation are not similar among road users. The peak of risk is around 10 years old for pedestrians and around 20 years old for car drivers or occupants in the western world. A comparison of the relative mortality rate according to age and sex for walking reveals that there are two main vulnerable groups: young boys (5–10) and old women (>65).

12.2.2.2 Geographical factors

The risk can be estimated on different geographical areas. Most of the time we follow some administrative division of the territory, as in regions. The spatial distribution of the mortality rate is useful to compare countries or regions as in Europe. The risk is smaller in northern Europe than in southern Europe (Eksler et al 2008).

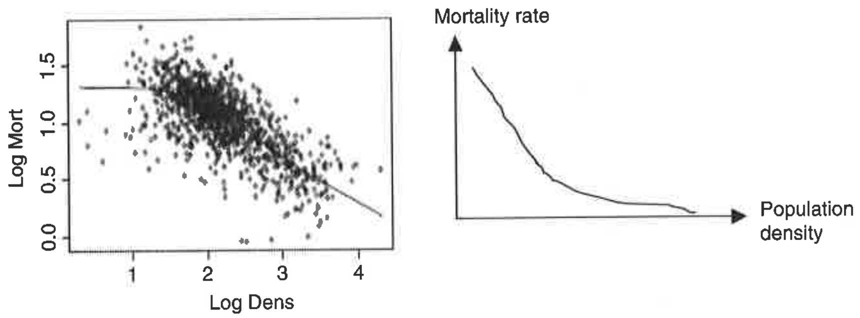

Figure 12.7 Mortality rate and population density (source: Eksler et al 2008) (unit 1/100000).

Some of the differences between regions or countries are due to the differences in the population density, which influences the mortality rate. The mortality rate decreases with the population density when we consider regions as spatial units. Rural regions have an higher rate than urbanized regions.

The estimation of such a relationship relies on the modelling of the number of fatalities by a Poisson distribution

$$Y_i \sim \text{Poisson}(\lambda_\text{i}\ N_i)$$

where λ_i is the parameter of the Poisson law, equal to the mortality rate for a spatial unit i, and N_i the number of inhabitants in the spatial unit i. We use a generalized mixed regression model by introducing some co-variables (the logarithm of the population density and a fixed effect "country" β_c with c the country index) plus a spatial non structural random effect ν_i on the mortality rate, which follows a normal distribution with zero mean:

$$\log(\lambda_i^\nu) = \beta_\text{c} + \beta(\log(\text{PopDens})) + \nu_i$$

12.2.2.3 *Mobility factors*

The most well-known relationship is the Smeed's curve relating the mortality rate to the motorization rate, equal to the number of motorized vehicle per inhabitant (car, lorry, bus, moto, and moped). The motorization rate increases with the economic development and tends to stabilize in developed countries. The curve was fitted originally with data for G-B from 1907–1947 (Smeed 1949) and for a panel of 68 countries between 1960–1967 (Smeed, 1968). From this curve, by regression, Smeed's law has obtained in which states the mortality rate increases with the motorization rate, with an elasticity of 1/3.

$$\frac{\text{fatality}}{\text{inhabitant}} = \text{c}\left(\frac{\text{motorised vehicle}}{\text{inhabitant}}\right)^{\frac{1}{3}}$$

$$\frac{\text{fatality}}{\text{motorised vehicle}} = \text{c}\left(\frac{\text{motorised vehicle}}{\text{inhabitant}}\right)^{-\frac{2}{3}}$$

$$\text{fatality} = \text{c}\ (\text{motorised vehicle})^{\frac{1}{3}}\ (\text{inhabitant})^{\frac{2}{3}}$$

It is in fact a Cobb-Douglas production function of the number of fatalities with two factors: the population and the fleet of motorized vehicles.

This law has some defects; first of all, it has been estimated by a regression over two highly correlated variables: population and vehicle fleet; secondly, it does not integrate the socio-technical progress of the road transportation system due to the improvement of vehicles and of drivers. Due to this failure, the law is not able to reproduce the peak in the trend of the number of fatalities in the western countries around the seventies.

Another class of road risk models developed by Koornstra and Oppe (1989, 1991) is able to predict this change from positive to negative of the number of fatalities. The model is based on a time series analysis of the number of fatalities on an annual basis and the number of vehicle * kilometers (equal to the number of motorized vehicles multiplied by the average of the number of kilometers driven in one year by a motorized vehicle). They model separately the fatality risk equal to the ratio of these two numbers by a decreasing exponential function over time and the number of vehicle * kilometers by a logistic function because there is a phenomenon of saturation about the motorized mobility in the developed world.

The evolution of mortality rate is nearly proportional to the number of vehicles (or vehicle * kilometers) mediated by a socio-technical factor, which is related to the safety of the road transportation system, and decreases the rate of the risk over time because of the socio-technical improvement of the system.

$$\frac{\text{fatality}_t}{\text{inhabitant}_t} = c\, e^{-bt} (\text{motorised vehicle})^{[0.8-1.2]}$$

There is a competition among the three components of the road transportation system: safety/mobility/demography. The rate of variation in the number of fatalities is the sum of the rate of variations of these three components:

$$-\mathrm{b} + \frac{\mathrm{dmotorisation}}{\mathrm{dt}} + \frac{\mathrm{dpopulation}}{\mathrm{dt}}$$

It is positive, and the number of fatalities increases over time, if the rate of socio-technical progress is lower than the sum of the increase in motorisation or mobility plus the increase in population. It is negative, and the number of fatalities decreases, if the safety progress is greater than the increase in mobility and population. In that case, safety outweighs mobility, in the former case mobility outweighs safety. What happens is that when the motorisation starts growing in a country, the growth rate is very strong, around 8%. With an increase in population of around 2%, it means that the safety rate has to be greater than 10% to counterbalance the effect of mobility. Usually the safety rate is around 5% per year and we have to wait for a slowing down of the mobility due to saturation to get a decrease in the number of fatalities.

In France (Lassarre and Hoyau 2009), the number of vehicle * kilometers increased at a rate of 7.5% up to 1973, the year of the first world oil crisis, which reduced the increase to 2.5% . The safety index measured by the number of fatalities per vehicle * kilometers decreases according to an exponential function at a rate of 5.4% per year. It follows that the number of fatalities rises steadily at a rate of 7.5–5.4 = 3.1% per year up to 1973, then decreases at a rate of 2.5–5.4 = 2.9% per year after 1973. On the fatality risk curves we can observe the impact of important national safety measures taken in 1973 (speed limits on national roads and motorways plus seat belt use) and 2002 (Speed cameras) which immediately reduced the number of fatalities by 14% and 17% respectively.

12.2.2.4 Economic factors

Many attempts have been made at an aggregated level (country level) to relate an economic indicator such as GDP or unemployment rate to the mortality rate. On the basis of an analysis of a macro-panel of countries from all over the world over a twenty year period, Kopits and Cropper (2005) end with a Kuznets curve relating GDP to mortality rate. The mortality rate

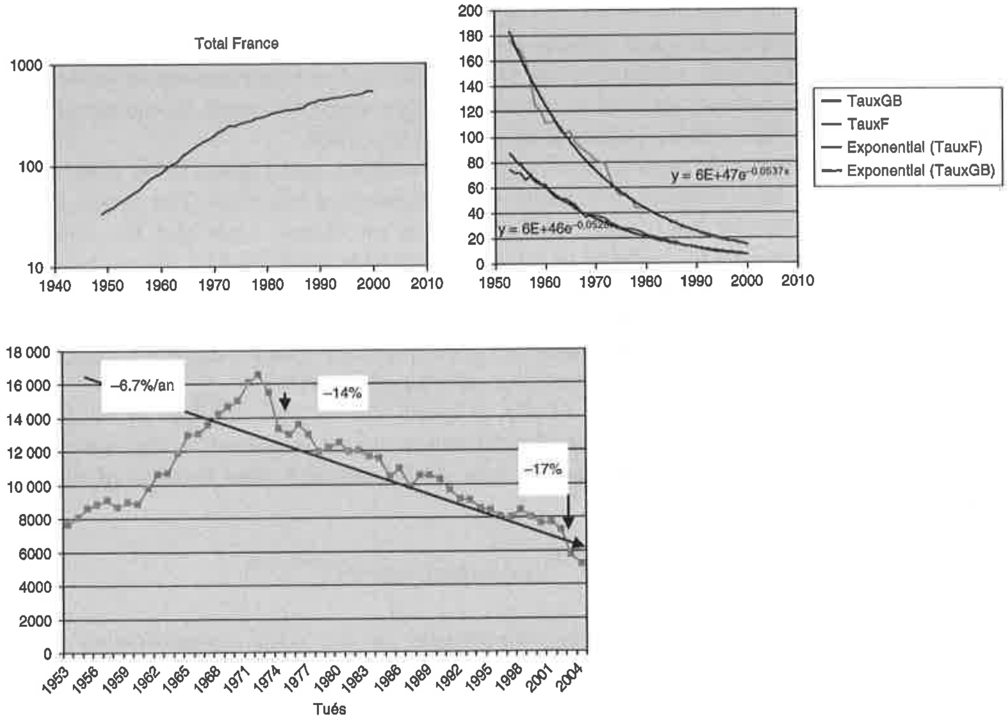

Figure 12.8 Evolution in France of the number of vehicle ∗ kilometers (logarithmic scale, unit 10^8), fatality risk (unit 10^{-8}) and the number of fatalities (source: Lassarre and Hoyau 2009).

increases with the economic development up to a turning point at a value of 8500 US$ then starts decreasing. At the country level, one needs to attain a certain degree of development in order to invest in safety for reducing the burden of road accidents, in the same way as for reducing environmental pollution.

Such models have to be examined carefully, as they rely first of all on the quality of the data collected (with uncertainty about the counting of fatalities in developing countries) and on the validity of the model and of its estimation techniques.

The long-term model of the mortality rate and the GDP can take a linear form after a logarithmic transformation; the coefficient β_0 is the elasticity which is constant but could differ among the countries

$$\log FAT_{it} = \alpha_i + f_i(t) + \beta_{0i} \log GDP_{it} + u_{it}$$

It includes a fixed effect of a country or a region, and a trend translating the socio-technical progress, depending on the country or region if necessary, which could be

- a linear trend $\beta_i \mathrm{t}$,
- a linear trend plus interventions (national road safety measures taken) $\beta_i \mathrm{t} + \omega_i \mathrm{I}_{it}$,
- a parabolic trend $\beta_i \mathrm{t} + \eta_i \mathrm{t}^2$.

In the Kopits model (2005), the elasticity is no more constant with the introduction of the square of the logarithm of GDP

$$\log FAT_{it} = \alpha_i + f_i(t) + \beta_{0i} \log GDP_{it} + \beta_{1i} \log GDP_{it}^2 + u_{it}$$

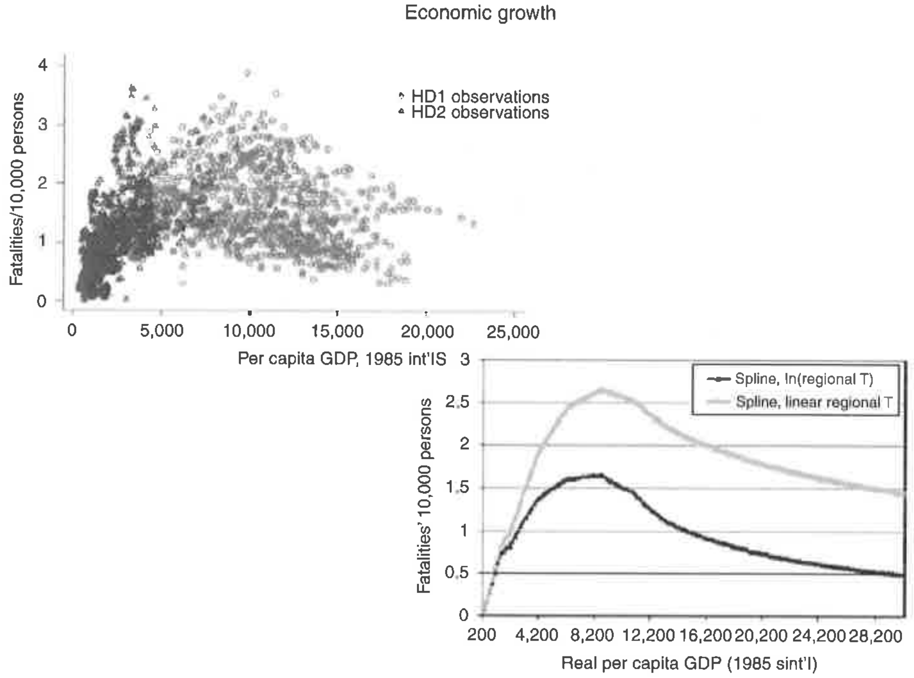

Figure 12.9 Data points and curves fitting per capita GDP and mortality rates according to Kopits (source: Kopits and Cropper 2005).

We get the Kuznets curve (an inverted U curve) with a concave parabola if $\beta_{0i} > 0$, $\beta_{1i} < 0$. In this case the elasticity is a linear decreasing function of the GDP

$$\frac{\text{dlog}\, FAT_{it}}{\text{dlog}\, GDP_{it}} = \beta_{0i} + 2\beta_{1i}\text{log}GDP_{it}$$

Some precautions have to be taken in the regression by introducing some proxy variables playing the role of common factors among the macro-panel and some structural models of the variances (autoregression and heteroscedasticity). Furthermore, the hypothesis of homogeneity of the elasticity (same value for all countries) has to be tested before applying such a hypothesis as in the Kopits model. The same applies for the test of a hypothesis of the same type and value of trend. And finally one has to consider the possibility of cointegration between these time series as we know that over more than twenty years the number of fatalities is not a stationary time series but could be integrated of order 1 or 2 (combining a deterministic and a stochastic linear trend) and that the deflated GDP is integrated of order 1 (Dupont et al 2014).

One has to be cautious about the results coming from such models on macro-panel data. Nevertheless from an analysis made by an IRTAD (International Road Traffic and Accident Database) working group (ITF 2014), we could admit a positive relationship between GDP and mortality rate and fatality risk. As we expect, a relationship between GDP and the mortality rate (because GDP influences the motorization rate or the mobility rate and acts as a proxy for this variable), the relationship between GDP and fatality risk is more informative in the sense that it shows a direct effect on the number of fatalities through the risk, plus an indirect effect through the mobility. The unemployment rate is negatively related to the mortality rate or the fatality risk, maybe because unemployment affects mainly the young people who are more at risk.

12.3 RISK INDICATORS IN ROAD TRANSPORT

People are not exposed permanently to the risk of traffic accident. In fact they are exposed to the risk when they are on the road or on the street as a pedestrian, cyclist, driver or passenger of a motorized vehicle. Note that some people, because of their jobs (policemen, street cleaners, or professional drivers), are more exposed to the risk.

The exposure is no more 24 hours long but limited to the duration of the presence on the road or street network. This exposure can be measured either by an amount of time, like a number of hours spent in the traffic, or by the distance driven on the network. The information on these exposure indicators comes from mobility surveys for persons (households) or goods (companies) or from counts on the road network by vehicle categories. In addition to the internal traffic of local people and companies, one has to consider the transit traffic from abroad coming or passing through the territory. We can also use the gas sales as a proxy for estimating the number of vehicle * kilometers driven by motorized vehicles in a country, knowing the consumption rate of liter per kilometer.

We end with a series of individual risk indicators:

- Rate of implication in an injury (fatal) traffic accident per hour or per kilometer,
- Rate of being injured or killed in an injury (fatal) traffic accident per hour or per kilometer.

The second one is called the fatality risk in terms of number of deaths per million of hours or kilometers. Such rates can be calculated for each road user type: driver of a car, motto, lorry, bicycle, passenger in a car, motorcycle, lorry, bicycle, public transport such as a bus, or pedestrian.

These rates can be related to the vehicle as done by the insurance companies. See by example the premium related to the number of kilometers driven. Usually a premium is calculated relative to the power and age of the vehicle, the residence of the owner, and the age and sex (beware of discrimination) of the main driver. We end with a new series of individual risk indicators which are:

- rate of implication in an injury (fatal) traffic accident per vehicle * kilometer
- rate of being injured or killed in an injury (fatal) traffic accident per vehicle * kilometer.

Sometimes these indicators can be downgraded by taking the number of vehicles, that is to say the size of the fleet of motorized vehicles in a country, as exposure. When calculating these risk indicators for each type of motorized vehicle, we could distinguish between an internal and

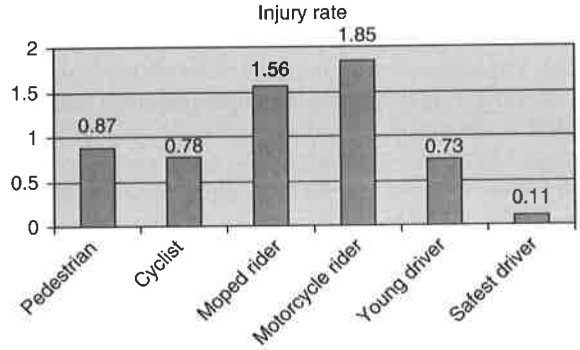

Figure 12.10 Injury rate in Norway (injuries per million person kilometres of travel) (source: Elvik 1999).

TABLE 12.1 Fatality rates per hour or kilometer in Great-Britain for transport modes and other activities (source: Evans 1994).

Great-Britain	**Fatality rate per 100 million hours**	**Fatality rate per 100 million kilometres**
Passenger travel by		
Bus	1.4	0.06
Rail	6	0.1
Car	12.4	0.4
Water	16	0.8
Air	20	0.04
Foot	27	7
Bicycle	64	4.6
Motorcycle	342	11.4
Employment		
All work	0.9	
At home		
All ages	2.6	
People over 75	22	

an external risk by taking into account either the individuals inside the vehicle, like the driver and the passengers of the car, or the individuals involved in the collision with that type of vehicle but situated on board another vehicle or as vulnerable road users such as cyclists or pedestrians. The fatality risk for the driver of a lorry is 15 to 20 times less than the fatality risk for the occupants or a car hit by a lorry. The internal risk for trucks is very low, but the external risk is rather high because of the mass of the vehicle, which causes a lot of damage.

These indicators are considered to measure the safety performance of the road transportation system. The fatality rate per billion of vehicle ∗ kilometre is a classic risk indicator which decreases regularly along an exponential function over time due to socio-technical learning, as has been shown previously. The fatality rate per vehicle ∗ kilometre decreases with GDP as shown by Kopits with a regression model on a macro-panel of OECD countries.

We can compare the performances of different modes of transport which are not by road, such air and rail, and also the transport risks with other risks generated by other activities, such as going to work or staying at home (Evans 1994).

12.4 MODELS OF ACCIDENT FREQUENCY AND SEVERITY

In the Insurance industry, the practice in risk evaluation is to separately model the frequency and the severity of crashes of the vehicles that are insured. The number of accidents is supposed to follow a Poisson distribution. The number of victims in an accident or the cost of an accident is a random variable Z, which could be a Pascal distribution for the number of injuries or a log-normal distribution for the cost of an accident.

Then, the number of victims or the total cost in a set of accidents is a random sum of random variables which follows a compound Poisson distribution

$$N_g(t) = \sum_{i=1}^{N_a(t)=n} Z_i = Z_1 + Z_2 + \cdots + Z_n \qquad \begin{array}{l} E(N_g(t)) = E(N_a(t))E(Z) \\ Var(N_g(t)) = E(N_a(t))E(Z^2) \end{array}$$

It means that the number of fatalities or injuries is an overdispersed Poisson distribution because the variance is not equal to the mean but greater.

For the number of accidents, especially for counts on road section or junction, a negative binomial distribution is used instead of the Poisson distribution. Some other specific distribution such as the zero inflated Poison distribution could also be used.

12.5 CONCLUSION

Risk of road traffic has been extensively described and analyzed through probabilistic and statistical risk models coming from epidemiology and reliability theories. Individual and collective risks are measured and estimated to assess the importance of the burden of accidents through classic risk indicators such as mortality and fatality rates, as well as morbidity and injury rates. The estimators are subject to the exhaustivity of the counts of accidents and victims and of the precision of the mobility indicators used as measures of exposure to the risk. The relationships between the risk indicators and the main risk factors are rather well documented by means of aggregated risk models, but the exactness of the form of the relationships requires advanced statistical technics and not just simple linear regression fits.

REFERENCES

Dupont, E., Commandeur, J. J., Lassarre, S., Bijleveld, F., Martensen, H., Antoniou, C., Papadimitriou, E., Yannis, G., Hermans, E., Perez, K., Santamarina-Rubio, E., Usami, D. S. and Giustiniani, G. (2014). Latent risk and trend models for the evolution of annual fatality numbers in 30 European countries. *Accident Analysis & Prevention*, 71, 327–336.

Eksler, V., Lassarre, S. and Thomas, I. (2008). "Regional analysis of road mortality in Europe," *Public Health*, vol. 122, no. 9, pp. 826–837.

Eksler, V., Lassarre, S. and Thomas, I. (2008). "Géographie de la mortalité routière en Europe," *Revue internationale de géomatique*, 20/4, 481–498.

Evans A. (1994). Evaluating public transport and safety measures. *Accident Analysis & Prevention*, 26, 4, 411–428.

International Transport Forum (2014). Road safety and economic development. Relationship between road safety performance and economic developments – International review. IRTAD research report, OECD, Paris.

Kopits, E. and Cropper, M. (2005). Traffic fatalities and economic growth. *Accident Analysis and Prevention*, 37 (1), 169–178.

Kopits, E. and Cropper, M. (2008). Why have traffic fatalities declined in industrialized countries? Implications for pedestrians and vehicle occupants. *Journal of Transport Economics and Policy*, 42 (1), 129–154.

Lassarre, S. and Hoyau, P.-A. (2009). Evolution récente du risque routier en France et Grande-Bretagne. In Delorme R., Lassarre S. Eds Les régimes français et britannique de régulation du risque routier: la vitesse d'abord. Synthèse INRETS n° 57.

Oppe, S. (1989). Macroscopic models for traffic and traffic safety. *Accident Analysis and Prevention*, 21, 225–232.

Oppe, S. (1991). Development of traffic and traffic safety: global trends and incidental fluctuations. *Accident Analysis and Prevention*, 23(5), 413–422.

Smeed, R. J. (1949). "Some statistical aspects of road safety research", *Journal of the Royal Statistical Society. Series A (General)*, 112 (1): 1–34.

Smeed, R.J. (1968). Variations in the pattern of accident rates in different countries and their causes, *Traffic Engineering & Control*, 10 (7), pp. 364–371.