The Øresund Traffic Forecast Model

Majken Vildrik Sørensen, mvi@ctt.dtu.dk[1]
Otto Anker Nielsen, oan@ctt.dtu.dk[2]
Jarn Schauby[3]

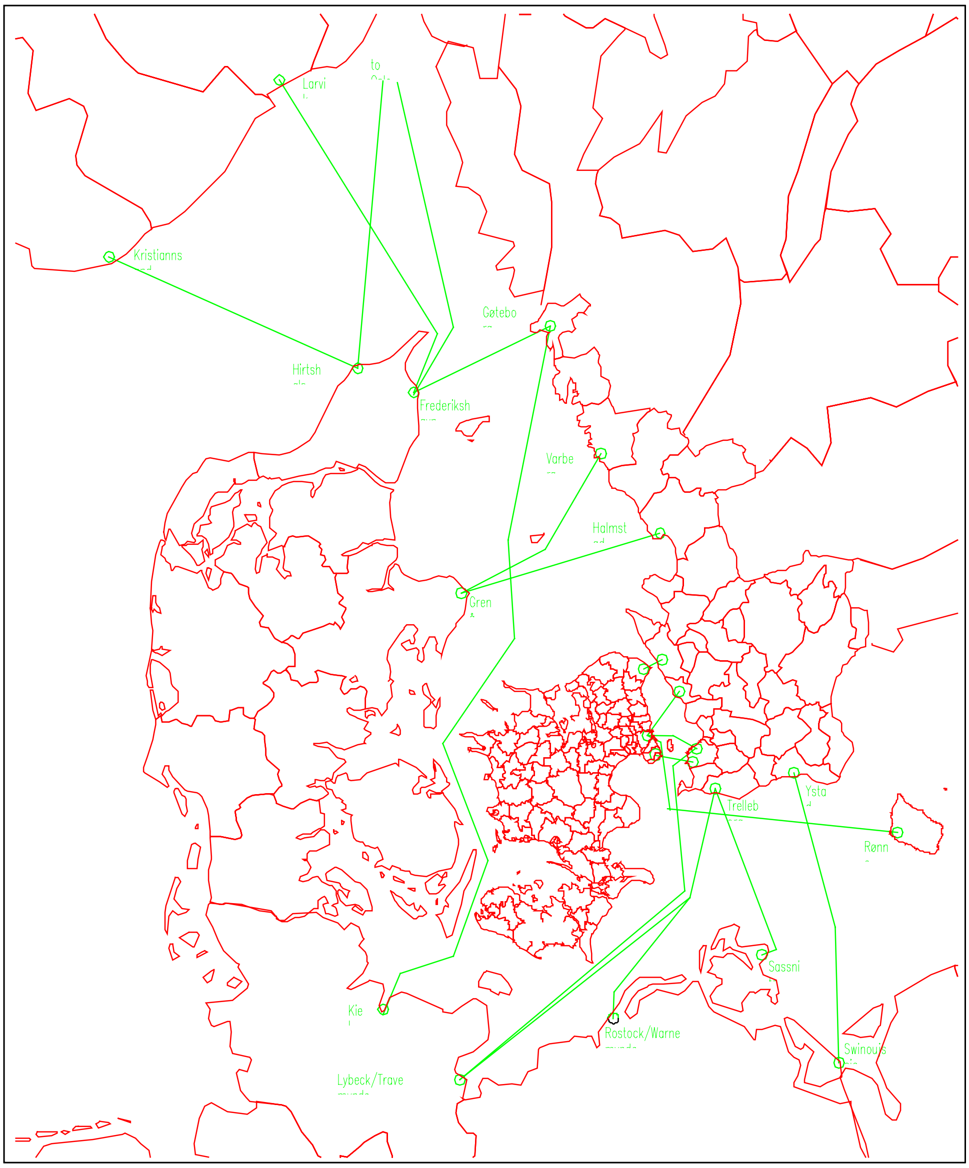

Figure 1 *Indication of Øresund crossings and zoning system (167 zones)*

---

[1] M.Sc., Ph.D.-student. Center for Traffic and Transport (CTT). Technical University of Denmark.
[2] Professor, PhD. Center for Traffic and Transport (CTT). Technical University of Denmark.
[3] Formerly at Øresundsbron. Vestersøgade 10. DK-1601 Copenhagen V.

Introduction

The Øresund Traffic Forecast Model was designed to forecast impacts of the fixed link crossing Øresund. The link connects the metropolitan area of Copenhagen with Malmö – the third largest city in Sweden. The link provides a direct rail and car connection between the two city-centres, previously only served by fast passenger boats. The latter was not connected to the rail network of Copenhagen, and parking facilities in Copenhagen near the boat was few and expensive. On the other hand several other ferry-lines connected other parts of the Copenhagen Region with the southern Sweden. Accordingly, the fixed link improved the connection between the two regions significantly – but to a varying extent for different parts of the region.

The fixed link also improved the infrastructure for longer trips between the Scandinavian Peninsula and the European Continent. However, in this respect the link compete with quite a number of ferry links.

From a modelling point of view, the task was difficult since the choice set included many overlapping routes and (sub-) mode-choices. For some trips, the link constitutes a significant improvement. But the link connects two countries, where culture, language and economy (e.g. labour legislation) provide a potential barrier for travel. Economic differences may also results in additional trips, e.g. shopping trips due to different VAT- and tax-levels on different goods.

To model this context, tenders were invited for a large-scale traffic model project. The resulting model was in many respects state-of-the-art: A utility-based model (nested logit) was estimated from joint RP- and SP-datasets collected specifically for the study. Several sub-models (segments) were estimated for short- and long trips, and for different trip purposes. In addition, both passenger and freight traffic were described by different (sub) models.

However, in reality the model failed to describe several impacts of the fixed link (as it turned out after the link opened). Especially, the model failed to describe the competition between fast boats and the rail link, and the magnitude of the social and economic barriers for short-distance traffic.

Since the model itself is described in Øresundskonsortiet (2000) its structure is only briefly described in section 2 of the present paper. Accordingly, the main focus is on lessons learned from the model project, starting with a short description of the forecast results versus actual traffic in section 3. Section 4 describes different modelling issues that are considered as the main problem with conventional models. Finally, section 5 sums up the conclusions and provide some guidelines for future models of projects that involve a mixture of regional and international (long-distance) traffic.

The model

Background and objectives

The Øresund Traffic Model has been designed for forecasting the future amount of traffic on the tolled Fixed Link between Copenhagen, Denmark and Malmö, Sweden and on around 25 competing ferry crossings between Denmark, Sweden, Norway, Germany and Poland.

The Fixed link has a total length of about 17 km and connects two major cities in Northern Europe. The Link opened in July 2000 for both road and trails traffic; access connections in form of new motorway sections and new rail lines have also been opened. The model is developed for Øresundsbro Konsortiet by COMVIN J/V (COWI Consulting Engineers and Planners, Denmark & MVA Consultancy, UK & InRegia, Sweden).

The aim of the model is to establish a comprehensive strategic planning tool, which will facilitate forecasting of future traffic levels with a wide range of planning and policy assumptions such as: different fare policies on the Fixed Link and competing ferry crossings, the transport service offered for both public and private transport users, macro-economic development, changed restrictions on land use, different transport policy options e.g. fuel taxes.

The model considers all relevant modes of travel including private vehicles, buses, trains, air passengers, various truck types and freight trains.

Scope of model

The model is based on an extensive set of Revealed Preference (RP) and Stated Preference (SP) surveys collected for this study. Statistical analysis of the survey data gives rise to model parameters, which determine much of the behavioral modeling in the Traffic Model. Survey data describing the observed pattern of trips in the study area forms an input to the forecasting process and is thus also part of the model. Inventory data is used to provide a description of the highway and public transport networks, and the services running on the public transport network.

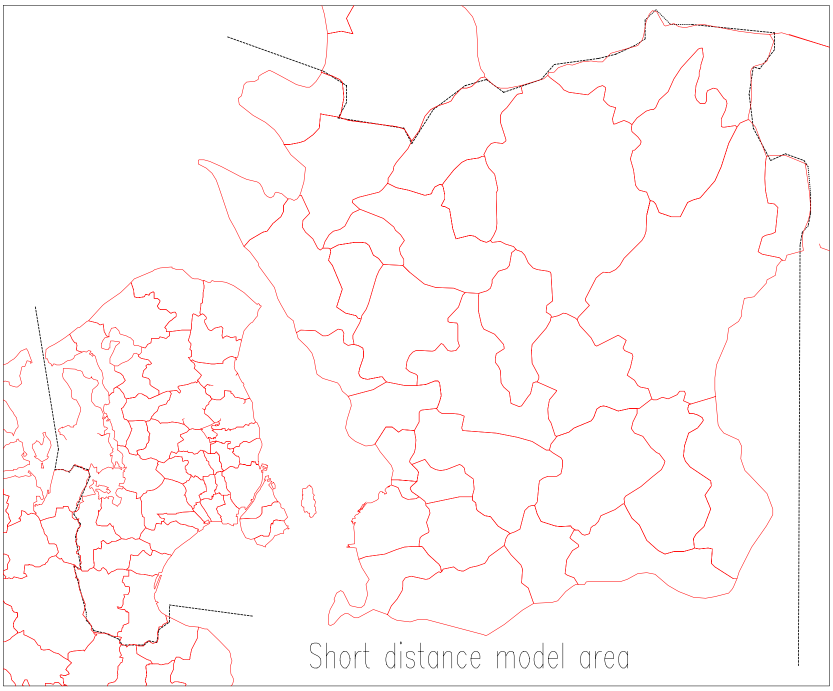

Figure 2 ***Area and zones of short distance mode***

The model contains distinct modeling for long-distance and short distance passenger trips and freight transport. Modeling of short distance trips is based on more extensive information and is therefore more precise than the corresponding modeling of long distance trips. Figure 2 shows the area and zoning associated with the short distance modeling, centered on the region of the Fixed Link.

Model overview

The Øresund Traffic Model contains a set of sub-models for the different elements of the forecasting, which are outlined in this section.

Short distance passenger model

Short distance trips are defined as trips whose origin and destination zones are both in the Øresund region, defined as the Hovedstad region and Skåne.

The short-distance model consists of a gravity model for generation of the growth of total short distance passenger trips in the Øresund region.

For people living in the region the traffic model includes in addition destination choice, mode choice and choice of crossing. For trips made by people not living in the region, the traffic model includes only mode choice and choice of crossing.

Choice of mode and crossing considers combinations of origins and destinations for Swedish and Danish sectors and for Swedish and Danish travelers separately across Øresund using absolute hierarchic logit modeling.
The choice of mode and crossing is modeled as a nested logit model, by choice of mode and under it choice of crossing. The choice is dependent on car availability.

The numbers of short distance trips, which are generated and attracted, to each zone are dependent on the forecasting of population, employment and economic activity data.

Long distance passenger model

Long distance trips are defined as the trips, which **do not** have both origin and destination in the Øresund region, which in turn is defined as the Hovedstad region and Skåne.

Unlike the short-distance model, no distinction is made about respondents’ residence or nationality.

The model generates the rate of change in long distance passenger trips between zones. This is done with a model describing tourist departures and arrivals by countries. First a calibrated matrix for the base year (1995) showing flows between countries is calculated. Then a corresponding matrix for a forecast year (e.g. 2010) is produced. The ratio between the two matrices gives the rates of change in flows between countries. These results are disaggregated into zones.

The long-distance model contains an incremental hierarchical model for choice of mode and crossing based on the logit choice formulation of trips across the Øresund.

Two trip types are used relating to journey purpose:

- Employer business trips;
- Other trips (including leisure, vacation, shopping, commuting, etc)

For both trip types the modes for crossing the Øresund are the same as for the short-distance model.

Freight transport model

The freight growth model distinguishes between long and short distance transport. Both models are gravity based models but with different parameters and explanatory variables.

The freight model for choice of mode and crossing, forecasts the volume of freight in tonnes and the number of vehicles on each of the crossing links. The model is implemented as a hierarchical logit model in absolute form.

The freight model for choice of mode and crossing interacts with the freight growth model. Changes in supply induce increased traffic through the growth model and mode shifts in the model for mode choice for loaded vehicles. These changes are reflected in the model for empty vehicles.

The freight model is applied separately for two different commodity groups, bulk and general cargo. In addition, there is a separate model for empty trucks.

Definition of trip sectors and trip types the short distance passenger model

Short distance trips are defined as trips whose origin and destination zones are both in the Øresund region, defined as the Hovedstad region and Skåne.

The following *trip sectors* are considered in the model:

A Residents in Skåne traveling within Skåne	**B** Residents in Skåne traveling to or from the Hovedstad region
C Residents in the Hovedstad region traveling to or from Skåne	**D** Residents in the Hovedstad region Traveling within the Hovedstad region

The Swedish segments (A and B) and the Danish segments (C and D) are treated separately in the estimation.

The following *trip types* are defined:

Short distance trips made by residents in the Øresund region.

- Work trips

- Shopping trips
- Business trips
- Other trips (vacation, pleasure etc.)

Short distance trips made by people not resident in the Øresund region
- Business trips
- Other trips

For the first four trip types made by residents in the Øresund region (all considered home based, such that the home zone is the origin zone) the traffic models include traffic generation, destination choice, mode choice and choice of crossing. This part of the traffic model is estimated on the data for individuals living in Skåne or Hovedstadsregionen.

For trips made by people not living in the region, the traffic model includes only traffic generation, mode choice and choice of crossing. The models are estimated on individuals not living in the region.

Furthermore a number of special-purpose trips are included in the model:
- Airport access/egress trips
- Sailing trips = pleasure trips with no embarkation.

Forecast experiences

In general the model is describing the traffic situation around the link fairly well. Looking from a users point of view, the model is set up in order to enable the user to change almost every parameter in the model. Still the model seems to have a serious problem with the level of traffic, which is simply estimated too high.

The model is not user-friendly, even though huge attempts have been made to construct an interface including easy access to the change of vital parameters. Changing parameters require an in-depth knowledge of the whole model and an understanding of the interactions between the single parts of the model. The model is constructed in a modular manner, but the different parts of the model is defined with respect to the theory of traffic forecasting, not with respect to the problems and scenario for which the user needs answers.

The question here is off-cause, what the model is constructed for. The model is not a decision tool for building the fixed link across the Øresund. When the work of the model took place construction of the fixed link had already been decided. The model is more like a medium term forecasting model, which should describe development in traffic depending on regional development and choice of mode under different price regimes and different assumptions of the speed and strength of the regional integration between the too areas, greater Copenhagen and Skåne. This was important for the evaluation of the funding and of the repayment of the fixed link, which is solely financed by the users. Of equally importance was the determination of optimal prices both with respect to securing the repayment of the fixed link, and at the same time optimising the integrational process between the two previously separated regions.

It seems that the scope of the model is too broad, especially when it is added, that the user in several occasions has been tempted to use the model as a pure operative model evaluating short-term price scenarios. Looking from a user point of view it is obvious, that traffic models are still lacking precision and satisfactory utility and demand functions in order to be used as short term operative models.

The traffic forecast model for the Øresund fixed link has been used very intensive during the last three years. First, as a forecasting tool before the opening of the fixed link and later, primarily for testing different price scenarios. Experiences from using the model have in general been good. The model is probably one of the most open and well-documented models of its size, and the model generates a huge amount of detailed information. This should been kept in mind reading the following sections.

As previously described the model consists of a growth module for general growth and a module calculation trip generation on the local level.

The general growth is determined by GDP on a regional level, and it is possible to add special assumptions for the growth in individual sectors. Population, employment and income for a given sector determine the trip generation from the local sectors. Together the two modules determine the induced traffic and the traffic growth over the forecasting period. For local traffic different equations for different segments are estimated, e.g. commuting is depending on population and employment, while shopping is dependent on population and income.

In the short term the model does not seem to catch variations in traffic. GDP are too rude to catch short-term variations, and the specifications on the local segments are estimated on cross sectional data, probably resulting in a bias. We find, that time series models describing traffic by real disposable income, rate of savings and split on Danish respectively Swedish travellers give better results.

Another problem for the growth module is that costs, e.g. petrol prices, are determined in the cost module. Travelling costs enter the utility functions, and thereby eliminated the effect of changes in the relative prices. As the fixed link is an international connection, changes in relative prices are more than just changes in petrol prices. It is also changes in the rate of exchange between the Danish and the Swedish currency.

Using the model it become obvious that the level of traffic was estimated too high. One reason is probably that the separation of the two regions has been lasting for over 400 years. It is not possible to find a new equilibrium over night, even though this is what the model assumes. The model is a general equilibrium model assuming full information and immediate reaction. Introducing a S-shaped penetration curve for each segment solved the problem. Hereby the general result of the model was sustained and only the assumption of immediate reaction was moderated.

Later in this paper other reasons for wrongly specified estimations of the traffic level are discussed, but it should be mentioned here, that another solution to the problem would have been a change in the size of the economic or the cultural barriers. It is possible to adjust the barriers sufficiently to fit the realised traffic levels. The barrier for commuting should be 3 times higher and the barrier for shopping should be 1,5 times higher. Only minor changes are necessary occur on the remaining segments.

The cultural and economic barriers to traffic between two countries are conceptually very satisfactory. They enter the utility functions as a cost element and thereby take part in the shaping of the implicit demand curve. As an element in forecasting they are very rude. In the model they are set like dummies, and there is no systematic way of changing the barriers over time. More work must be done in order to decompose the effects, and the barriers must be included in a more consistent theoretical framework.

When the model is used for different price scenarios the implicit price elasticises and thereby the utility functions take an important part. As will be described later the model has some problems in the model split between public transport and the existing high speed ferries. For cars the split seems to work satisfactory and the division of traffic between ferries and the fixed link seem reasonable compared to the actual traffic.

The price elasticises for local traffic, induced traffic, are very high, for the segments shopping and visit/vacation more than –3. Taking the size of the region into consideration, 3 million inhabitants within a radius of 100 km, this is not unrealistic in principle. The local traffic before the opening of the fixed link accounted for no more than 1,000 – 1,500 cars per day. After the opening of the bridge we must realise that the elasticities probably are estimated to be too high.

In general, our experiences with the model are positive concerning the international traffic, and unsteady concerning local traffic. During the process we have relaxed different assumptions in the model and tested different theories in trip generation. Basically all results have shown even higher traffic levels than the levels reported in the model. Today we are inclined to believe that the implicit demand functions (utility functions) for local traffic are less adequate and could be improved.

Modelling issues

The previous chapter gave a brief introduction to the model while the following section will put special emphasis on some issues that may be critical for the model as well as of general interest for modellers. The paper is focused towards the regional passenger models, since the methodological findings for these are believed to be of most general interest.

The long distance model followed traditional approaches (nested logit models), which turned out to give good results for both passenger and freight traffic. The reader is accordingly referred to the technical documentation (Øresundskonsortiet, 1999). No regional freight model was commissioned, as no satisfactory tender was received for this. This pinpoints a need for research as well as practical experiences for this type of model.

The regional models were constructed within a relatively limited timeframe though some issues that had not been dealt with before should be incorporated into the model. Some shortcuts have been made and the consequences of these on the model outcome have not in all places been evaluated. These evaluations are included here as well as general comments on the model structure.

Trip generation

A growth model performs the generation of trips. Rates of growth are estimated at zonal level and are subsequently applied to the collected data from 1995/96. The rate of growth is subdivided into four levels; firstly, traditional growth between zones for trips performed by residents in the Øresund region (Greater Copenhagen Area and Skåne). Then growth at national level (non-residents trips) where growth between (groups of) countries (π_{IJ})[4] is measured which applies to all zones within the country. Thirdly, a growth factor for land based trips to/from the airport in Copenhagen (Ψ) due to an increase of the number of flights departing from this airport rather than the airport Sturup just across the Sound in Sweden.

The model design with a growth model retains the overall trip pattern observed in the base data, which is often referred to as a pivot-point method, i.e.

$$T_i^{u(1)} = T_i^{u(0)} \cdot \frac{T_i^{u(1)est}}{T_i^{u(0)est}} \qquad (1)$$

where $T_i^{u(0)}$ is the base (survey) value for trips in the base year, and *est* are the corresponding estimates for the base *(0)* and forecast year *(1)* respectively.

The model was estimated on a cross section in time (only one year) and space (only trips crossing – or potentially crossing – the Sound, not internal trips in the two regions).

The individual's choice of whether to perform a trip includes (with some exceptions) the knowledge of the generalised costs. These are contingent on the network as there are a quite limited number of crossings of the Sound. In the year 2000 as in the forecast situation, a new link is added to the network at a new location[5], implying altered generalised cost for numbers of OD relations. Since the 1995/96-generation of trips were based on the generalised costs of trips within the 1995/96 network and the 2000 is a extrapolation of these patterns, the 2000 trip patterns will inevitably be based on the 1995/96 costs as well.

Since the trips crossing the Sound only constitute a small number of trips in the Øresund region[6], it was difficult to estimate the base matrices. Furthermore, an option of where to cross the Sound was added by the opening of the bridge, which could be expected to alter the destination choice pattern significantly for trips crossing the Sound – which also turned out in reality. These effects are both in conflict with the underlying assumption of equation (1); that is, the base matrix has a high quality and that changes to this caused by the project under evaluation is (fairly) marginal.

---

[4] The notation follows the official documentation of the model in Øresundskonsortiet (1999).

[5] Before the crossing, only a hydrofoil boat connected Kastrup/Copenhagen with Malmö, the two largest cities in the region. Car ferries operated south of Copenhagen and approx. 40 km to the north at Helsinore/Helsingborg. The only rail connection was between Helsinore/Helsingborg. The fixed link contains a motorway as well as a rail line running between the central stations of Copenhagen and Malmö.

[6] The daily number of trips in the greater Copenhagen region is about 10 millions, in Sweden the number of trips is about 3 millions (both dependent on the geographical limitation of the region). About 30.000 trips cross the Sound per day.

The conclusion must be, that pivot-point methods are not suitable in this type of model context. This error in the number of trips generated is inherited throughout the model.

Barriers

As the link between Sweden and Denmark did not replaces an existing link but merely created a new corridor – which indeed changed the accessibility to regions on the other side of the Sound, the model needed to include the concept of two less integrated regions. The proposed mean to include this has been two measures of barriers where one is devoted to the economic aspect while the other is a measure of cultural integration.

To describe the estimation of the barriers a basic knowledge of the utility function is required. The function reads as follows:

$$\begin{aligned} U_{ijmw}^{uk} &= \alpha^u + \gamma d_{ij}^u + \gamma\beta t^v + \gamma\beta\rho \cdot f\left(t^{hp}\right) + \gamma\beta\mu\left(t^a + t^e\right) + \gamma\beta\sigma \cdot ch \\ &\quad + \gamma\delta t^f + \gamma\varepsilon t^h + \gamma\theta t^c + \gamma\eta + \zeta \cdot CAV^k \end{aligned}$$

where:

α	ASC for mode and crossing
γ	Coefficient for generalised cost
d_{ij}^u	Monetary cost from zone *i* to *j* for purpose *u*
β	VOT inv time
t^v, t^a, t^e, t^f, t^h, t^c	time component for resp. inv time (*v*), access (*a*), egress (*e*), ferry (*f*), headway (*h*), check-in (*c*)
ρ	coefficient transforming waiting time to inv time equivalents
$f(t^{hp})$	first waiting time transformation; min. of half the headway and 10 minutes
μ	coefficient transforming access/egress time to inv time equivalents
σ	time per interchange
ch	no of interchanges
δ, ε, θ	VOT ferry time, headway, check-in time
η	link cost
ζ	coefficient for car availability
CAV^k	Car availability

In the forecast model the following values estimated from SP data are employed[7]

$\beta = \delta = \theta = 0.57$ DKR/min
$\eta = -1.04$ DKR for Fixed Link, 0 otherwise
$\varepsilon = 0.33$ for ferry headway, *0.15* for public transport on Øresund
$\rho = \mu = 2$
$\sigma = 5$ min/interchange,
Hence, α, γ and ζ are left to be estimated.

[7] It should be noted that the SP data consists of 728 interviews, though the models are based on 442 (short distance) and 328 (long distance) SP interviews.

For each sector the levels of utility are estimated. The barriers are estimated from the difference in utility levels between sectors, where the utility of the destination choice is formulated as:

$$\xi_n \cdot U_{ij*} + \lambda_n \cdot \Gamma_{ij} ,$$

where ξ_n is a structural parameter (estimated in sectors B and C which are transferred to A and D respectively), * denotes the logsum, λ_n the barrier effect for sector *n* (defined zero for sector B, C) and d_{ij} indicator for trip between *i* and *j.*

Finally the barriers are deduced and the division into economic and cultural barriers is based on the following assumptions.

- There is no cultural barrier for Swedes going *to* Sweden as well as for Danes going *to* Denmark.

- The economic barriers are specific for sectors (B and C) and do not vary by nationality of the traveller.

The barriers are estimated separately for each direction on the Sound. As the cultural barrier is estimated by assuming no cultural barrier for individuals going to their original country (while living in the other) there needs to be a group of respondents present in the sample who changed their residence from one side of the Sound to the other. In sector B (Swedes travelling to Denmark) the fraction of non-Danes applied for the deduction of the cultural barrier is quite low for three of four segments (4, 4, 9%). Similarly for sector C (Danes travelling to Denmark) with 2, 6, 10 % for the same three segments. This appears to be a (very) small reference groups – though no figure was present in the documentation from the consultants.

As mentioned above, the majority of the coefficients are estimated jointly from SP data, while the remaining are calibrated following the estimation. This does not ensure consistency between the estimates (possible scale effects from the logit model are ruled out by the use of VOT in the utility function rather than the estimated coefficients). The documentation does not clearly state whether the remaining coefficients are estimated by means of a discrete choice model, or how. If they are estimated from a discrete choice model (NL) then a scale is introduced to the coefficients again, as the logit model can only estimate scaled coefficients[8]. This would significantly affect the model results.

The estimation of the coefficients β are restricted by the not overwhelming number of interviews performed. This implied that no separate models (coefficients) are estimated for each sector. This in fact corresponds to assuming that Swedes and Danes are homogeneous – an assumption that is not empirically based. This can lead to poorly estimated coefficients (low t-values), wrong size as well as a less good fit of the model (likelihood) causing low quality forecasts.

It should be noted that the barriers are not estimated jointly with the β-s, and the 'estimate' of the barrier effect was found as the difference between the deterministic

[8] See e.g. Ben-Akiva & Lerman (1985).

parts of the utility – without taking into consideration the random error term in the logit model[9].

When applied in the forecast model the change (reduction) in the barriers are given as percentage change compared to the modelled barrier. The reduction applied to economic as well as the cultural barriers without regard to the sign of the barrier.

All the problems mentioned above relates to problems of estimating the model. Basically, it was difficult to carry out sufficient interviews (only a very small fraction of the total traffic in the region crosses the Sound).

Alternatives

The type of statistical model applied in this analysis is a Nested Logit model (NL), which partitions the overall choice situation (choice of one alternative from all alternatives) into a number of smaller choice situations (bundles of alternatives to choose from). Alternatives in each bundle (nest) share some characteristics, which differ from other nests' intra-nest shared characteristics. Alternatives within the same nest are assumed statistically independent, in all but the characteristics just mentioned.

The hierarchical structure applied for this analysis is choice of mode before choice of route.

Since crossing of the Sound used to be limited to ferry routes – regardless of the mode of transport, it is necessary to describe the trip by mode of transport before the Sound, on the crossing as well as after. The relevant modes are bus (B), car (C), train (T) and (unspecified) public transport (P) on either side and 8 ferry links where **one** allowed for trains and 5 were restricted to only passenger transport. Disembarkation is labelled D. The ferries are grouped by between which harbours they run (H, L, C, D), it should be noted that parallel services were available. This adds up to that the chains of modes described below were available at the time of data collection. The total amount of alternatives is 21.

Table 1. Alternatives and their correlation

Transport chain	Crossing of Øresund	No. of interchanges
BBB, CCC	H, D	0
TTT	H	0
CDC, CDP, PDC, PDP	H. L, C, D	2

The alternatives depicted are clearly not uncorrelated as they pair wise share transport modes on part of the trip or use the same link on the Øresund crossing. Had these common features been limited to either mode or route it might have proven sufficient with a nested logit model which is in fact used in this model, where

[9] The utility of an alternative in a (variant of) logit model is defined as $U = \beta X + \varepsilon$ where βX is referred to as the deterministic part of the utility and is in fact the only part that is estimated. The random part ε is assumed to follow a Gumbel distribution and to be IID across the individuals in the model. When differences between stochastic estimates are to be found not only the deterministic but also the variance of the random part should be included.

the upper level nesting split the choice between the common aspect (whether that would be mode or choice[10]).

In this case with two levels of common aspects it is not possible to construct a nesting structure, in agreement with assumptions of the NL model. Alternative models that can incorporate these features are Error Component Models (Ben-Akiva *et al*, 1993)[11] or the Cross-Nested Logit model (see Vovsha, 1997), which are described in further detail below. However, it is stressed that the use of these techniques are given as recommendations for future model projects. At the time when the model was constructed, both EC and CNL models were very new proposals in the research community, and very little practical experiences existed.

Error component models

The error components model (EC) employ that individuals have heterogeneous preferences towards various characteristics of a trip; that is they weight differently the relative importance of pairs of attributes of a trip. This variation in preferences applies between individuals as well as between periods of time (from trip to trip) for the same individual. Hence, this type of models can incorporate aspects, such as when a link is part of two different routes and is valued differently for each of these routes. In other words, this corresponds to mixing the choice of mode and route into a combined step where the choice is between alternatives that consists of both a mode and a crossing of the Sound. In the following, a brief introduction is given to error-component models.

In utility maximising models, travellers are represented as choosing alternative i from the set of n alternatives when i has the highest utility. A random utility function for i consists of a discrete part V_i and a random part ε_i, as illustrated in (2):

$$U_i = V_i + \varepsilon_i \qquad (2)$$

If the deterministic utility V_i can be measured and the distribution of ε_i can be stated, the probability of making each choice can be calculated. The logit model can be derived by assuming ε_i independent and identically distributed, *iid,* across alternatives and respondents, and to follow the Gumbel distribution. The probability P_i of choosing alternative i from the set of n alternatives is then given by the conventional logit formula (3):

$$P_i = \frac{e^{V_i}}{\sum_j e^{V_j}} \qquad (3)$$

The deterministic component V_i of the utility U_i is often assumed to be a linear function of travel time T_i and travel cost C_i:

$$V_i = \beta_c C_i + \beta_t T_i \, , \qquad (4)$$

[10] The order of mode and crossing is determined in the model estimation (by the logsum parameter).

[11] Other names for these models are Random Coefficients Logit, Random Parameters Logit and Mixed Logit.

where β_c is the cost coefficient and β_t the time coefficient, both are to be estimated from data.

The implicit Value Of Time (VOT)[12] can be shown to be given by the ratio $v_t = \beta_t / \beta_c$. Accordingly, the VOT is constant for multinomial logit model. This is obviously a problem if VOT differs between travellers due to their different preferences or tastes. The assumption of constant VOT can be relaxed by segmentation into different trip purposes, e.g. that business travellers have higher VOT's *as a group* than commuters. The same applies for differences in VOT's for respondents travelling by different modes where VOT differences can be related to different comfort levels offered by the different modes. The problem of the constant VOT can also be relaxed by segmentation into income groups or by other socio-economic variables. However, all these modifications still assume the same VOT for all persons *within each group*, and the model is simply designed as a number of 'parallel' logit models.

Accordingly, logit models can handle random *iid* variation of the utility between individuals while they cannot explain differences between preferences (coefficients) within one group. That is, the utility variation cannot be related to known data, it must just be constant over individuals and alternatives. However, in practice there are important variations that do have effect over observations and alternatives.

For example, suppose that the VOT (coefficient) v_i is the only systematic variation. Consider the simple choice situation with independent alternatives and Gumbel distributed random variation, where we for simplicity assume no other elements in the utility function than time and cost (refer to Ben-Akiva *et al*, 1993, for a more general case). With fixed VOT, v_t, the utility function is:

$$U_i = \beta_c (c_i + v_t t_i) + \varepsilon_i \qquad (5)$$

With this design, one can assume a distribution of v_t[13]. In a multi-modal network, such as the Øresund crossing, the utility function will include many more elements. However, it is an important feature, that correlation between alternatives can be described by the ε_i[14] following techniques usually used for assignment models, i.e. stochastic user equilibria based on probity models. While taste variations can be described by the distribution of v_t.

[12] This is a measure of the relative valuation (trade off) of time compared to cost, where the coefficients are measures of the preferences.

[13] This model incorporates a variation over individuals with a variance related to the amount of time spent, since v_t is multiplied by the amount of time spent in each alternative, which is therefore no longer *iid*. Accordingly, such models are theoretical complicated, but fully feasible today. See e.g. Ben-Akiva & Bradley (1993) , Sørensen & Nielsen (2001) and Sørensen (2002).

[14] $\varepsilon_i = \sum_{a \in R(i)} \varepsilon_a$ represent random variations, as the sum of variations along the *arc (a)* used along the route *(R)*. If the sums of ε_a at *route-level* followed a Gumbel distribution, this specification for constant ξ would result in the conventional logit model, where all alternatives are assumed independent. However, link variations that add to give Gumbel distributions at route level are not known and probably do not exist. Moreover, in many choice situations independence is a very poor assumption, which has resulted in the use of nested logit models, that allow correlation between some specific alternatives. However, in route – and joint route and sub-mode - choice models the number of alternatives is so large that this approach is completely unsuitable. The problem can be partially avoided by using a multinomial probit model, where ε_a is assumed to follow the normal distribution, and where correlation between different routes is allowed (Sheffi, 1985). Although the number of routes still presents serious problems for model estimation. In model application, the normal distribution has the disadvantage that negative link-costs must be avoided. Therefore Gamma-distributed ε_a's are recommended, since the Gamma distribution is reproductive in mean and variance -just like the Normal distribution - given that the variance is linked to the mean by a fixed coefficient of variation (Nielsen and Jovicic, 1999), and it has the advantage of being non-negative.

Using error-components in complex utility functions combined with techniques from assignment models have been used to solve complicated choice situations in several Danish models.

In the Harbour Tunnel model (Paag & Daly, 2000), such an assignment model (Nielsen *et al* 2000c) was used to describe the choice of a tolled tunnel facility in a dense urban road network suffering congestion (Copenhagen). The results were compared with a more traditional approach, describing the choice between the tunnel and alternative routes by a hierarchical logit model. This comparison showed that it is absolutely feasible to describe the tunnel choice as part of the route choice problem.

In the Copenhagen-Ringsted study (Nielsen *et al,* 2000b) a similar approach was used to describe the joint mode and route choice problem in a complicated public transport network (Nielsen & Jovicic (1999), following earlier work in Nielsen (2000a).

Cross-nested logit models

The Cross-Nested Logit model (CNL) is a fairly new type of model, which allows for alternatives in different nests to be correlated. The nests are, like in the traditional nested logit model, grouped into bundles where alternatives in each nest share some characteristics not shared by the other alternatives. An example of the CNL model is shown in figure 3, where in fact the same alternative is present in more than one nest, while the combination of characteristic and alternative is not present more than once. This allows for the covariance at the lower level.

Hence, this model allows for structures like those present in the Øresund model where choice of crossing is dependent on choice of mode, meanwhile choice of mode is dependent of choice of route.

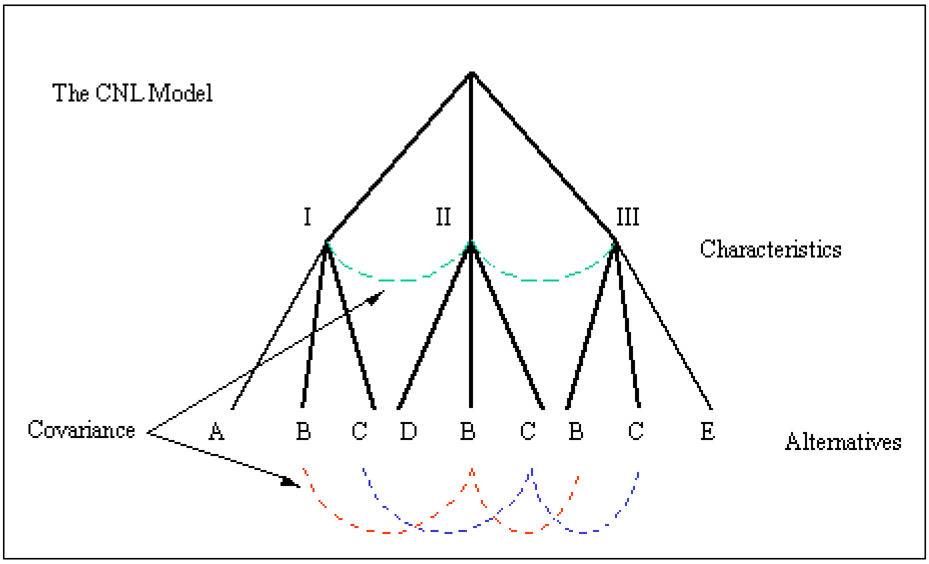

Figure 3 The CNL model

The CNL model is not the only one to incorporate similarities between alternatives – the Paired Combinatorial Logit model (Koppelman & Wen, 1998) incorporates joint aspects by means of similarity coefficients, while the Probit model (Ben-Akiva & Lerman, 1985) allows for a full covariance structure between the alternatives.

Other aspects

Several other issues may explain the differences between the forecasted traffic, and the measured traffic after opening the fixed link.

The number of interchanges was incorporated into the model (utility function) only as added travel time (fixed amount of minutes per interchange) valued as travel time. Recent results found in other studies (see e.g. Nielsen *et. al.*, 2000b) suggest that the disutility of interchange time is significantly higher than in-vehicle travel time. In addition to this, the interchange in itself may cause a disutility to the individual (referred to in the literature as a *interchange penalty*), which is not included here. These may bias the coefficient β for in-vehicle time.

Moreover, problems with the network for the Copenhagen area have been detected; some zones are not linked logically to the network, which implies the traveller in the model has to take a detour compared to the real world implying additional generalised costs.

Another issue is the concept of park and ride which has in fact been present in the after bridge situation. On the Swedish side of the bridge there is a link from the north-south motorway crossing the railway line only 2 km from the bridge. At this place there is a rail station with a free parking facility (compared to the parking charge of up to 25 DKR/ hour in Copenhagen and the distance is approx. 25 km by car); the train fare is 60 DKR per person and the travel time is 28 minutes to Copenhagen City, while the car fare is approximately 200 DKR and travel time 30 minutes.

The paper has put the main emphasis on the short distance effects. In the model for long term effects measures as location of residence are incorporated. However, the effect of changed residence does only seem to effect the trip to and from work (if workplace is maintained) and not all other trips undertaken by the individual/ household. It seems rather obvious that if an individual changes residence to the other side of the Sound then the trips to/ from family/ friends are influenced as well. The trips all by one have had one of the trip ends changed (though the frequency may be lowered). Hence, the generalised cost for change of location should make allowance for this. The influence of this on the model results is somewhat limited.

Organisational issues

Traffic models are often used as basis for decisions, after which the models are abandoned. In the case of the Øresund model, it was used as a decision tool before opening the fixed link, as well as it is being used as a tool for planning – e.g. prize strategies – after the link opened. Accordingly, before and after studies of the modeled versus measured traffic has been carried out which provided the findings referred to in the paper. This gave new insight in traffic modeling, compared to many model studies where such analyses are not carried out.

As in most model projects – at least in Denmark – tenders was invited for the actual work, and a consortia of consulting companies carried out the work. In many respects, this approach is recommended since the tender with the best price and quality can be selected. However, by outsourcing the work the documentation tends to be written at a late stage of the work – or after the model has been finished.

Accordingly, it is difficult to follow the work as a client, and external review of the work is also difficult. If – in addition – the work is delayed, little time is left for quality control before the model is being used for project evaluation.

Accordingly, much emphasis must be laid in formulating demand to the documentation and interaction between client and consultant.

Conclusions

The paper described the Øresund traffic forecast model. It is a utility-based model (nested logit) estimated from joint RP- and SP-datasets collected specifically for the study. Several sub-models (segments) were estimated for short- and long trips, and for different trip purposes. In addition, both passenger and freight traffic were described by different (sub) models.

To model this context, tenders were invited for a large-scale traffic model project. The resulting model was in many respects state-of-the-art. Nevertheless, some of the model segment turned out to predict traffic impacts that differed significantly from the actual (measured) impacts. This was especially the case for the regional passenger models, while the traditional approaches for long distance trips proved reasonable.

Concerning future models for situations where route and (sub) mode choices are correlated, and where the impacts concerning the cross-section under evaluation is not marginal, the following recommendations can be given:

- Traditional nested logit models proved insufficient. Nor can a mix of traditional assignment models and demand models be recommended. Two alternative approaches are suggested (1) To model the choice by a multi-modal assignment model, where the utility functions are estimated with random coefficients (Error Component Logit, also referred to as mixed logit). (2) To model the choice situation by a Cross Nested Logit (CNL) model (1) is recommended in cases with many alternatives, (2) in cases with a limited number of alternatives.

- When changes in supply are relatively large, it is dissuaded to use pivot-methods in the demand model (i.e. model relative changes in flows compared to the present situation, rather than using an absolute model). Or – at least – great care should be taken.

- Models of traffic across national boarders with cultural, linguistic and economic barriers (including tax policies, labour legislation, etc.) are especially complicated to estimate. Separate utility functions for each nationality as well as for each type of travellers (within versus between nations) must be estimated, and barrier effect must be part of the estimation – not calibrated afterwards. However, these request also demand a larger sample than collected in the Øresund model.

Concerning the organisation of modelling projects, we recommend contractual binding agreement concerning delivery of documentation if the proposed model design is changed during the estimation and calibration phase. Also a not too tight time-schedule is desirable, since this makes it possible for the client to order independent review/advises to secure model quality. Naturally, such recommendations conflicts with the often tight (or even unrealistic) time-schedule that policymakers set up for modelling projects.

References

Ben-Akiva, M., Bolduc, D. & Bradley, M. (1993). Estimation of Travel Choice Models with Randomly Distributed Values of Time**.** *Transportation Research Record.* 1413, 88-97.

Ben-Akiva, M. og Lerman, S.R. (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press Cambridge, Mass.

Koppelman, F.S. & Wen, C-H (1998). The paired combinatorial logit model: properties, estimation and application. *Transportation Research Part B Methodological.* Vol. 34 (2), p. 75-89. Elsevier Science Ltd.

Nielsen, O.A. & Jovicic, G. (1999b). A large-scale stochastic timetable based transit assignment model for route and sub-mode choices. *European Transport Forum (PTRC).* Proceedings of Seminar F, Transportation Planning Methods, Vol. P 434. 169-184. Cambridge.

Nielsen, O.A. (2000a). A Stochastic Transit Assignment Model Considering Differences in Passengers Utility Functions*. Transportation Research Part B Methodological*. Vol. 34B, No. 5, pp. 337-402. Elsevier Science Ltd.

Nielsen, O.A., Daly, A. & Hansen, C.O. (2000b) A Large-scale model system for the Copenhagen-Ringsted railway project. *9th International Conference on Travel Behaviour Research*. *Travel behaviour Research: The Leading Edge*. Book edited by David Hensher. Pergamon press. Chapter 35, pp 573-596

Nielsen, O. A., Daly. A. & Frederiksen, R. D. (2000c). A stochastic multi-class road assignment model with distributed time and cost coefficients. *9th International Conference on Travel Behaviour Research*. Proceedings, Vol. 12, Application Workshop 4: Large scale model systems. Gold Cost, Queensland, Australia, July.

Daly, Andrew; Nielsen, Otto Anker (2002). Large Scale Model Systems. Paper / workshop report in *Travel behaviour Research: The Leading Edge*. Book edited by David Hensher. Pergamon press. Chapter 18, pp 285-296.

Paag, H., Daly, A. & Rohr, C. (2000). Predicting Use of the Copenhagen Harbour Tunnel. To be published at. *9th International Conference on Travel Behaviour Research*. *Travel behaviour Research: The Leading Edge*. Book edited by David Hensher. Pergamon press. Chapter 35, pp 573-596

Sheffi, Y. (1985). *Urban Transportation Networks*. Prentice Hall, Inc, Englewood Cliffs, NJ.

Sørensen, M.V. & Nielsen, O.A. (2001) Assessing Distributions of Error Components in a Complex Traffic Model. To be presented at the *World Conference on Transportation Research (WCTR)*, Korea July, 2001

Sørensen, M.V.. (2002) Error components in demand estimation. To be presented at the *European Transport Comference (ETC),* Cambrigde, England, September.

Vovsha, P. Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area. *Transportation Research Record*. 1997 (1607), p. 6-15.

Øresundskonsortiet (1999). Traffic Forecast Model. The Fixed Link across Øresund. Øresund Konsortiet[15]

[15] Can be obtained by contacting Øresundsbron. Vester Søgade 10. DK-1601 Copenhagen V. E-mail: info@oeresundsbron.com, Internet: www.oeresundsbron.com .