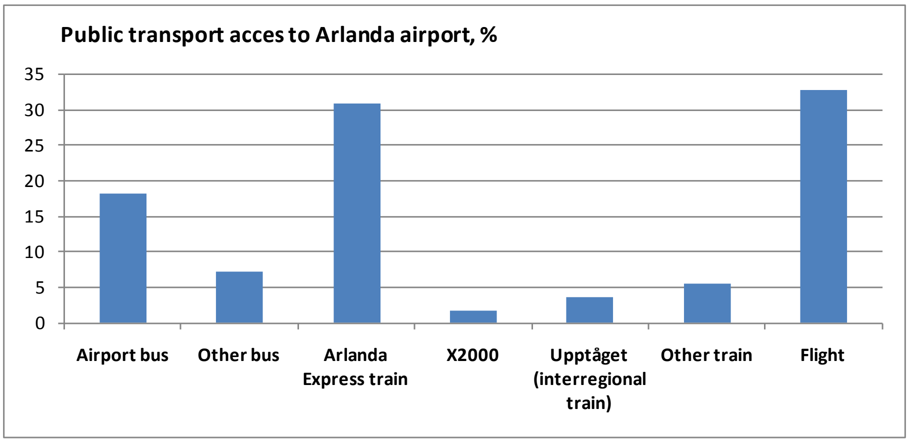
Towards a model for long distance passenger travel in the context of infrastructure and public transport planning



STAFFAN ALGERS
JOHN BATES
KJELL JANSSON
HARALD LANG
ODD LARSEN
HENRIK SWAHN

Report
Stockholm 2013

TRITA-TSC-RR 13-013
ISBN 978-91-87353-27-7
www.kth.se/abe

KTH Arkitektur och samhällsbyggnad
Avdelningen för trafik och logistik
KTH, SE-100 44 Stockholm

Royal Institute of Technology (KTH)
School of Architecture and the Built Environment
Division of Transport and Logistics

TRITA-TSC-RR 13-013

ISBN 978-91-87353-27-7

Towards a model for long distance passenger travel in the context of infrastructure and public transport planning

Consensus report by (in alphabetical order)
STAFFAN ALGERS
JOHN BATES
KJELL JANSSON
HARALD LANG
ODD LARSEN
HENRIK SWAHN

Kungliga Tekniska Högskolan (KTH)
Avdelningen för Trafik & Logistik
2013-12-20

Preface

It is likely that long distance travel will become more important from a transport policy point of view during the next few decades. Regional economic development and increasing focus on the interaction between regions, environment and climate policy, as well as technical development in passenger transport and infrastructure are important driving forces. Therefore, it is increasingly important that the tools for analysis and forecasting of long distance travel that are used by public agencies and other actors are of high standard and that they are developed to keep up with new developments and policy issues.

Models of long distance travel demand have a central role among the tools for analyses, primarily because they give important input to decisions about capacity as well as for societal Benefit Cost Analyses. The models that are now in practical use for transport planning and policy analyses by public agencies and for other purposes by other actors have all been around for many years. Despite the fact that both models and data for the models have been developed over the years, both users and researchers in the field have come to the conclusion that the current models have certain deficiencies in relation to current requirements. A particular and somewhat troublesome problem is that the outcome of an analysis may well differ considerably between different models.

In this report we outline possible and promising ways to improve the current model systems for long distance travel, which have the potential to give better traffic forecasts as well as more correct societal Benefit Cost Analyses. We propose that model development within the field of long distance travel is carried out in two time perspectives – the short to medium term, and the longer term. Our recommendations are based on the analyses of the central issues in long distance travel modelling which we have conducted within the project. In this process we have benefited from our joint experience as well as our contacts with Swedish and international research within the field. We have also benefited from the comparative analyses carried out in this project between the model implementation Sampers, which clearly dominates Swedish transport policy analyses, and one implementation of the model system Vips.

We gratefully acknowledge the financial support to this project from Vinnova, Banverket, SIKA and Storstockholms lokaltrafik as well as the active personal support from Fredric Almqvist, Banverket, Gunilla Wikström, SIKA. We are also grateful for the expertise regarding Sampers and Visum contributed respectively by Peter Roming and Josef Andersson. We are indebted to Andrew Daly for constructive criticism and suggestions. We would also like to thank the participants of a seminar that was arranged by the project team for many valuable comments and suggestions.

The authors share the responsibility for the contents of the report.

Staffan Algers, John Bates, Kjell Jansson, Harald Lang, Odd Larsen, Henrik Swahn.

Contents

Executive summary

Objectives

The objective of this study is to set out conditions for how demand for long distance travel should ideally be modelled, to compare existing models to these ideal conditions and to suggest how further research could contribute to improving the capability of applied models.

The transport policy relevance of long distance travel is growing

Although Long distance travel (here defined as journeys of more than 100 km one way) contributes a relatively small part of total travel in terms of the number of journeys, it is considerably more important for total transport and traffic work (passenger kilometres viz. vehicle kilometres).

Environmental policies as well as climate policy now influence the development of societies within all sectors. The increasing integration of distant labour markets, as well as the growing specialisation of production of goods and services, seems to require more mobility over longer distances, while there is considerable political pressure to reduce impacts on environment and climate from the transport sector. Major changes in public transport supply as well as for the use of the private car and air for long distance travel will be important political considerations during the coming decades. Therefore it is important to have adequate and efficient analytical tools for analysing and forecasting total long distance travel and its distribution over origin-destination relations, modes and travel paths, as well as to provide the basis for societal Benefit Cost Analyses.

Current models for long distance travel demand need improvement

The model system Sampers dominates Swedish transport policy analysis related to long distance travel. The model systems Vips and Visum are also used, but to a considerable lesser extent – predominantly for the analyses of specific issues relating to major infrastructure and traffic projects – often within the railway sector. These different models have in some cases turned out to give divergent estimates of various aspects of demand – and sometimes the differences have been quite significant. Naturally this provokes questions about the general reliability of the models as well as whether one or the other of the models is more correct.

The criticism of the demand models is, however, not confined to cases where estimates from different models diverge. There has also over the years been criticism originating from the regular users of the models. Demand estimates have been questioned on a priori or face validity grounds. While not all of this criticism may be justified or even relevant, much of the criticism has been taken seriously and has led to various development actions in order to improve the performance of the models or the data that the models rely on. However, within such a continuous development process it is both costly and difficult to implement major changes of the fundamental properties of the models, and this has limited the range of possible improvements.

Demand and supply of long distance travel are complex – prioritizing and choice between alternative approximations are necessary in modelling

In this study we have tried to develop a general understanding of the problem of modelling demand for long distance travel by pointing at the multi-faceted as well as complex nature of the modelling problem. In doing so, we have set out conditions for how long distance travel should be modelled, as guidance for further development of practical applications. Approximations of the true reality always have to be made in modelling, and these approximations will always incur some bias and error. In practice models will make use of a range of such approximations. The ways to approximate used in different model systems will in the end also lead to differences in the estimates generated by the models. The choice of particular approximations to use are inter alia influenced by how the model builders and their commissioners have assessed and prioritized different aspects of the reality that the model aims to reflect.

The fundamentals of current models were developed more than 20 years ago.

Two major model systems for long distance travel demand currently in use in Sweden differ in scope and basic principles

Between the dominating model in Sweden, Sampers, and the model system Vips there is a fundamental difference, namely that the Vips system is not a comprehensive demand model since it only deals with the problem of distributing a **given** demand for movements between relations on modes and travel paths.

In order to enhance our understanding of how the model results are influenced by the different approximations they use we have studied these approximations theoretically in some detail and compared the outcomes from the two model systems based on "experimental" practical model runs where the "experiments" have been designed to highlight the effects of specific algorithms and approaches. Because of the limited scope of the Vips demand model these comparisons have been confined only to the part that is common ground for the two model systems, namely the distribution of demand over modes, lines, and routes, and the estimation of the effects on the travellers' level of service (consumer surplus change) of assumed transport measures.

In addition, the model systems Sampers and Vips differ with respect to the general model structure (regarding the approach to modal split and route choice) as will be briefly explained below.

The Sampers model comprises two distinct parts, namely a network model, Emme/2 that uses an optimisation algorithm called OS (optimal strategy) and a discrete choice model (logit type). The latter model is used to estimate total demand and its distribution over relations and modes. The network model is used for three purposes. Firstly it is applied to generate the estimates of travellers' travel time components that are used for the initial estimation of the logit model together with behavioural survey data. Secondly the network model is used to provide input data about travel time components in the model application phase, when travel demand and its distribution are calculated by the logit model. Thirdly it is used to assign travel demand per mode and relation to different travel paths represented in the network.

The Vips model, on the other hand, uses only a network model, for distribution of demand over both modes and travel paths. The algorithm of this network model is called RDT – Random Departure Time, which is different from the algorithm that is used by the Sampers' network model.

The differences between the two model systems are thus both in the basic model structure and in the algorithm that is applied in the network model parts of the two model systems. These two approaches to the modelling problem represent differences in the emphasis and assessment of various aspects that are expected to influence travellers' choice of mode and travel route. One key difference relates to the assumed significance of travellers' individual (or travel group) preferences for modes and consequently how such differences are treated. Another key difference relates to the treatment of timetables. The approach taken in the Sampers system gives more consideration to individual preference differences, whereas the approach taken in Vips is designed to give more consideration to timetable effects, i.e. average schedule delay.

In transport modelling, the concepts generalised cost (GC) or generalised price (GP) are often used. These concepts aim to include the cost of all inputs carried by the traveller for production of the travel service, part of which may be captured by the fare paid, and including the input of time, discomfort etc by the traveller. While the two models are generally similar in the assumptions that they make, a particular difference will be seen in the treatment of waiting time/service frequency. It should also be noted that while the concepts are straightforward when they apply to identifiable services, there are further theoretical issues which apply when they are used to relate to the average or "composite" service over a number of different alternatives.

As already noted, in Vips the travellers' choice of travel path simultaneously covers both the mode and line choice, based on the generalised cost for different travel alternatives. This generalised cost also factors in mode-specific aspects. In Sampers, on the other hand, the possibilities that travellers' varying preferences for different modes will influence their choice of mode can be allowed for by means of a mode constant in the generalised cost/utility function of the logit model. Such a constant reflects some of the inertia of mode choice, so that the effect of changes of generalised cost components on mode choice is thereby mitigated. There is much empirical evidence that individual preferences for modes, in addition to variations of travel conditions for individuals or travel parties, influence mode choice.

Among the properties that an ideal model of long distance travel demand should have is a proper account of individual preference differences in many dimensions, as well as an explicit account of individual preferences with respect to departure/arrival time and how these departure/arrival time preferences interact with time table information. We have therefore come to the conclusion that it would be useful to try to combine both properties within one demand model. Of course, there may be different ways to do this, and we suggest some research ideas in this direction.

In the study we illustrate by means of examples and analytical discussion how the different model principles, as well as the approximations chosen in relation to an ideal model, will influence the estimates of change of generalised cost, and hence the demand changes.

Calculation of change of consumer surplus due to different measures

Change of consumer surplus (CS; i.e. the change of the travellers' utility) is often the most important component in a societal Benefit Cost Analysis for various measures relating to traffic or infrastructure supply. Therefore it is important that CS change is calculated as correctly as possible. For this the demand model plays an important role, since it is used to calculate the main components of CS change, namely changes of travel volumes for modes and travel paths, as well as providing estimates of the generalised cost changes induced by the measure.

We have established that there are sometimes differences between the outcomes of the two model systems with respect to the consumer surplus change due to some policy measure. Such differences can sometimes be fully or partly explained by differences in the set of parameters used by the practical system implementations, e.g. in relation to values of time. However, there are also other differences that are not readily explained by different parameter values, in particular for cases where headways are changed and/or for actions that influence the properties of travel alternatives with more than one mode. In such cases differences in the way the two systems calculate CS change, as well as the way by which journeys with mode combinations are treated, influence the outcomes.

Differences in the calculated consumer surplus change between the two model systems could, in real cases, be quite large, as is illustrated by the High Speed Train example presented in the report. Clearly, as we say in the report, the way by which consumer surplus change is calculated should be consistent with the demand model, as well as theoretically correct. The differences that we have observed based on detailed analyses show that these requirements may not be fulfilled today. Therefore we think it is necessary to carry out a more in-depth study of the calculation of consumer surplus change related to different demand models and in particular for the current model implementations.

Five main issues for further research have been identified

We have identified five main areas that are of crucial importance in the process to develop better models for long distance travel demand, namely:

1. How important for long distance travel decisions is heterogeneity with respect to travellers' preferences as well as situation- or traveller-specific travel conditions, and what feasible alternatives are there for approximating such heterogeneity? The choice of the most suitable modelling approach depends on the answers to these questions.

2. Which existing (or newly developed) network algorithm is most suitable in the context of alternative basic modelling approaches to long distance travel to project travel lines, and for calculating travellers' generalised cost?

3. How important is it to model correctly journeys with combinations of modes in the light of current and potential travel patterns?

4. How could consistency between the demand model and the module for calculation of consumer surplus change be attained in practice for existing models as well as for new/modified models?

5. How could full consistency be attained between the different levels of a comprehensive demand model?

In the report there are certain hypotheses as to the answers to these questions. However, it is clear that the answers – so far – are uncertain and that considerably more efforts are required to reach more definite answers.

The future development of long distance demand models should have two time perspectives and study four alternative strategies

We propose that the future development of long distance demand models be planned for the short to medium term (0-4 years) as well as for the longer term (4-10 years).

In order to establish the potential for improvements as well as to make concrete suggestions for development, we think that the following four strategies should be further investigated in the short to medium term time perspective:

1) Study the possibility for further developing and improving the traditional approach with a combination of a logit model and a network model, primarily focussing on the development of the logit model with a specification and structure that is more suitable than existing specifications for long distance travel

2) Continue the investigation – already started in this study – of the possibilities for improving the models that are based on the RDT algorithm in three ways. The first would be to further develop methods to include traveller heterogeneity in the RDT-algorithm. The second would be to develop methods to modify the RDT algorithm through using distributions of demand and supply other than the uniform distributions that are implemented today. The third development area is to further study and to develop ways to more stringently calibrate a model implementation based on RDT.

3) Study the possibilities and potential benefits of combining RDT and logit by using estimates of generalised cost from RDT for the estimation and application of a logit model for mode choice and/or for the modelling of travel frequency as well as the travel distribution between relations.

4) Carry out further studies of the option of using a timetable-based model. A first step might be to make a pre-study that in more detail illuminates the properties of two existing timetable-based model systems, namely Sampers' and Visums timetable versions.

Within all four development lines above it is necessary to constantly bear in mind how the calculation of change of consumer surplus can be dealt with consistently.

We have pointed out the complexity as well as the multi-faceted nature of long distance travel itself and the factors that determine this type of travel. In the long run development, it is possible to take more of these general aspects on board, and to find ways that are less tied in with earlier modelling traditions and sunk cost considerations relating to existing model systems. Therefore, we propose the development of a long term research program for modelling long distance travel demand. Hopefully, the findings of this project could to some extent prove useful and be taken on board for such a research program.

Förord

Det finns mycket som talar för att det långväga resandet under de kommande decennierna kommer att få en allt större transportpolitisk betydelse. Regional utveckling och samspelet mellan olika regioner, miljö- och klimatpolitiken liksom den tekniska utvecklingen inom transportområdet är viktiga drivkrafter. Det blir därför allt viktigare att de analysverktyg som myndigheter och andra aktörer inom transportområdet använder för att analysera och prognosera utvecklingen av det långväga resandet håller en hög klass och utvecklas i takt med områdets utveckling och de nya frågeställningar som aktualiseras.

Modeller för efterfrågan på långväga resande intar en central position i den nödvändiga arsenalen av analysverktyg, främst därför att de ger ett viktigt underlag för kapacitetsdimensionering och för samhällsekonomiska bedömningar. De modeller som nu används i den transportpolitiska planeringen och av andra aktörer inom transportsektorn har alla ett antal år på nacken. Trots att de modeller som används idag och de data som utnyttjas i modellerna successivt har utvecklats har både användare av modellerna och forskare bedömt att modellerna har vissa brister i relation till dagens krav. En svårighet är också att beräkningsutfallen för en viss frågeställning kan skilja sig avsevärt mellan olika modeller.

I denna rapport skisserar vi tänkbara vägar för förbättring av nuvarande modellsystem för långväga resande beträffande efterfrågeprognoser och beräkning av konsumentöverskott för samhällsekonomiska bedömningar. Vi föreslår utveckling såväl på kort och mellanlång sikt som på längre sikt. Rekommendationerna bygger på analyser av centrala frågeställningar inom området som gjorts inom ramen för projektet, där vi utnyttjat författarnas samlade erfarenhet och kontakter med svensk och internationell forskning inom området. Vi har också dragit nytta av jämförande analyser mellan det i svensk trafikplanering dominerande modellsystemet Sampers och en implementering av modellsystemet Vips.

Vi vill framföra ett varmt tack till VINNOVA, Banverket, SIKA och AB Storstockholms lokaltrafik som finansierat arbetet i projektet samt till Banverkets företrädare Fredric Almkvist och SIKAs företrädare Gunilla Wikström som aktivt medverkat i projektets referensgrupp och där tillfört många värdefulla synpunkter. Vi tackar också Peter Roming och Josef Andersson som medverkat i arbetet som experter på Sampers respektive Visum, Andrew Daly som bidragit med konstruktiv kritik och värdefulla förslag samt till alla dem som under arbetets gång tillfört värdefulla synpunkter bland annat vid det seminarium som arrangerades inom ramen för projektet.

Författarna står gemensamt bakom innehållet i denna rapport.

Staffan Algers, John Bates, Kjell Jansson, Harald Lang, Odd Larsen, Henrik Swahn

Sammanfattning

Studiens syfte

Syftet med denna studie är att analysera och skissera vilka villkor och krav som bör uppfyllas då man modellerar efterfrågan på långväga resande samt att jämföra några existerande modeller med dessa villkor och krav. Med utgångspunkt i dessa jämförelser föreslås ytterligare forskning som konkret skulle kunna bidra med kunskap som ger förutsättningar för utveckling av bättre tillämpade modeller.

Det långväga resandet (inom landet) får en växande betydelse i trafikpolitiken

I Sverige bidrar det långväga resandet med en mindre del av antalet resor men det är mycket betydelsefullt för det total transport- och trafikarbetet.
Miljö- och klimatpolitiken påverkar samhällsutvecklingen inom alla områden. Integration av arbetsmarknader och specialisering av varu- och tjänsteproduktion ställer krav på utveckling av mobiliteten samtidigt som höga krav ställs på att minska trafikens miljö- och klimatpåverkan. Stora förändringar kan komma att behöva övervägas under de kommande decennierna såväl när det gäller utbudet av kollektivtrafik för långväga resande som när det gäller den individuella biltrafiken. Därför är det angeläget att de verktyg som används fungerar så väl som möjligt för att analysera och bedöma det långväga resandets omfattning och fördelning på olika relationer samt dess fördelning på olika färdmedel och resvägar och för att ge underlag för samhällsekonomiska bedömningar.

Dagens modeller för efterfrågan på långväga resande har brister

Idag dominerar modellsystemet Sampers i den reguljära svenska transportpolitiska analysen som berör långväga resande. Modellsystemen Vips och Visum används också, men i betydligt mindre omfattning och oftast för analyser av speciella frågeställningar i anslutning till specifika projekt, oftast järnvägsprojekt. Dessa olika modeller har i praktiska tillämpningar i många fall visat sig ge skilda resultat – i flera fall har dessa skillnader varit betydande, vilket naturligtvis väcker frågor om modellernas tillförlitlighet generellt och i vad mån och/eller i vilka avseenden den ena eller andra modellen är mera rättvisande och tillförlitlig.

Även i den löpande tillämpningen har modellerna kritiserats av olika skäl och från olika utgångspunkter. Trots att det över tiden har skett en viss löpande utveckling och förbättring av dem är det svårt och kostnadskrävande att förändra respektive modells egenskaper på ett mera grundläggande sätt.

Efterfrågan och utbud för långväga resande är komplexa företeelser – modellering kräver prioritering och val mellan olika approximationer

I denna studie har vi försökt utveckla en generell förståelse för problemet att modellera efterfrågan på långväga resande genom att visa på problemets mångfacetterade och komplexa natur. I anslutning till detta har vi definierat vissa villkor och krav som bör uppfyllas av modeller för långväga resande och som kan tjäna som vägledning för utveckling och tillämpning av dessa modeller. Approximationer av den ”sanna verkligheten” måste göras i modellerna, men dessa approximationer innebär alltid en avvikelse ifrån verkligheten. Modellerna innehåller i praktiken en mängd sådana approximationer. Skillnader mellan olika modellsystem uppkommer bland annat på grund av det sätt man valt att approximera inom olika delområden. Valet av approximationer som en gång gjorts beror på hur de som byggt modellerna (och deras uppdragsgivare) värderat och prioriterat olika aspekter.

Grunderna i de modeller som används idag utvecklades för mer än 20 år sedan.

De två dominerande modellsystemen i Sverige har olika omfattning och bygger på olika grundprinciper

När det gäller de modeller som används i praktiken idag i Sverige, Sampers och Vips, finns en fundamental skillnad nämligen den att Vips inte är en fullständig efterfrågemodell utan endast behandlar problemet med att fördela en given efterfrågan på långväga resor i olika relationer på färdsätt och resvägar.

För att studera och fördjupa förståelsen av hur de olika approximationer som används då man modellerar efterfrågan på långväga transporter påverkar resultaten, har vi mera i detalj studerat vilka approximationer som utnyttjas i de nu praktiskt använda modellsystemen Sampers och Vips och om och hur dessa approximationer påverkar efterfrågeberäkningarna och vilka skillnader som finns mellan de två modellsystemen. På grund av Vips-modellens begränsade omfattning har dessa jämförelser endast omfattat den del som båda systemen hanterar, nämligen fördelningen av efterfrågan på färdmedel och resvägar samt beräkning av effekter på resenärernas standard (förändring av konsumentöverskott) av antagna transportåtgärder. Vi har försökt belysa vilka grundläggande egenskaper hos två av dessa modeller, nämligen Sampers och Vips, det är som förorsakar skillnader i beräkningsresultat

En grundläggande skillnad mellan modellsystemen Sampers och Vips (när det gäller fördelning av ett givet resande mellan färdmedel och resvägar) finns när det gäller modellens generella struktur.

Sampers-modellen innehåller två distinkta delar nämligen en nätverksmodell, Emme/2, med en optimeringsalgoritm (kallad OS – Optimal Strategy) och en diskret valmodell (logit). Nätverksmodellen används dels för att generera restidskomponenter för estimering av logitmodellen dels för att beräkna restidskomponenter för resenärerna för olika representationer av reella nätverk i samband med logitmodellens tillämpning samt därefter för att fördela resandet i dessa nätverk. Nätverksmodellen är uppdelad på färdmedel och fördelningen mellan färdmedel sker i logitmodellen.

Vips-modellen består däremot endast av en nätverksmodell med en optimeringsalgoritm (RDT –Random Departure Time) som skiljer sig från den som används i Sampers-modellens nätverksmodell.

De grundläggande skillnaderna mellan modellsystemen ligger således dels i den övergripande modellstrukturen (för det här aktuella delproblemet av långväga resande) dels i den algoritm som tillämpas i respektive nätverksmodell. Dessa två lösningar speglar skilda betoningar och värderingar av olika aspekter av hur ett givet resande per relation fördelas på färdmedel och resvägar. En grundläggande skillnad finns när det gäller vilken betydelse man tillmäter individuella variationer i resenärernas preferenser för olika färdmedel och hur man behandlar sådana skillnader. En andra grundläggande skillnad ligger i det sätt som modellerna beaktar trafikanternas förhållningssätt till tidtabeller. Den ansats som valts i Sampers lägger större vikt vid skillnader i individuella preferenser i andra avseenden än individuella önskemål om tidpunkt för avresa/ankomst medan Vips lägger större vikt vid den senare aspekten och därmed till trafikanternas förhållningssätt till tidtabeller.

Bland de villkor som vi identifierat som centrala att beakta i modeller för långväga resande finns såväl skillnader i individuella preferenser som tidtabellsanpassning från trafikanternas sida. Vi har mot denna bakgrund dragit slutsatsen att en möjlig väg att förbättra dagens modeller vore att kombinera de två nämnda modellsystemens ansatser i syfte att utveckla en modell som uppfyller båda villkoren. I rapporten skisserar vi några forskningsförslag med just detta syfte.

I studien belyser vi med olika exempel och analytisk diskussion hur de skilda modellprinciperna och approximationerna påverkar beräkning av förändring av generaliserade kostnader och efterfrågeförändringar.

Beräkning av förändring av konsumentöverskott vid olika åtgärder

Den komponent som oftast spelar störst roll vid en samhällsekonomisk värdering av en förändring i trafiksystemet är hur den påverkar konsumentöverskottet (nyttan för resenärerna). Det är därför viktigt att beräkna denna förändring så korrekt som möjligt. Efterfrågemodellen har här en mycket stor roll att spela eftersom trafikåtgärdernas förändring av resandet på olika färdmedel och linjer tillsammans med beräkningen av förändringen av trafikanternas generaliserade kostnader ligger till grund för beräkningen av förändringen av konsumentöverskottet.

Vi har konstaterat att förändringen av konsumentöverskottet till följd av en viss trafikåtgärd i vissa fall skiljer sig mellan Sampers/Samkalk och Vips. Vissa sådana skillnader eller en del av skillnaden förklaras av skillnader i t ex de tidsvärden som används. Det finns dock även flera skillnader som inte låter sig förklaras av detta. Det gäller t ex vid åtgärder som innebär förändring av turtäthet och även i vissa fall åtgärder som till en inte obetydlig del får sin effekt genom att de påverkar de generaliserade kostnaderna för en del av en resa som genomförs med en kombination av flera färdmedel. I dessa fall kan dels beräkningssättet i Sampers-systemets Samkalk-modul , som är en tillämpning av den så kallade ”rule of one half” dels förekomsten av och sättet att behandla resor som innehåller kombinationer av flera färdmedel spela in.

Dessa skillnader i beräkning av konsumentöverskottsberäkning kan i reella fall vara stora. Vi konstaterar i rapporten att sättet att beräkna konsumentöverskottet bör vara konsistent med och teoretiskt korrekt givet efterfrågemodellen. De skillnader som vi observerat och de mera detaljerade analyser vi genomfört visar att det kan vara så att dessa krav inte är uppfyllda idag. Vi bedömer därför att det är nödvändigt att genomföra en fördjupad studie av hur väl dagens modellsystem beräknar förändrat konsumentöverskott och om kompletterande ansatser kan medföra förbättringar.

Fem huvudfrågor har identifierats som kräver fortsatta insatser för forskning och utveckling

Vi har identifierat fem huvudområden som centrala för en fortsatt utveckling som syftar till bättre modeller för efterfrågan på långväga resande. Dessa områden är:

1) Vilken betydelse har heterogenitet när det gäller resenärernas preferenser, situationsspecifika förhållanden etc. för resebesluten? Valet av den i någon mening ”bästa” modellansatsen beror i hög grad på svaret på denna fråga.

2) Vilken existerande eller nyutvecklad algoritm för allokering av det långväga resandet i nätverket av trafik respektive för beräkning av trafikanternas reseuppoffring är bäst lämpad vid modellering av långväga resande?

3) Hur viktigt är det att på ett korrekt sätt modellera resor som använder kombinationer av färdmedel?

4) Hur säkerställer man i praktiken konsistens mellan efterfrågemodell och beräkningar av förändringar av konsumentöverskott – i dagens modeller och eventuellt nyutvecklade modeller

5) Hur kan man så långt som möjligt uppnå konsistens mellan den övergripande efterfrågemodellen och modellen för val av färdmedel och resväg?

I rapporten konkretiseras vissa hypoteser när det gäller svaret på dessa frågor. Vi kan dock konstatera att dessa svar tills vidare är osäkra och att det krävs väsentligt större insatser för att nå fram till säkrare svar.

Fortsatt utveckling bör ske med två olika tidsperspektiv och studera fyra alternativa strategier

Vi föreslår att den fortsatta utvecklingen planeras dels på kort till medellång sikt (0-4 år) dels på längre sikt (4-10 år).

De fyra strategier, som enligt vår uppfattning bör studeras närmare på kort och mellanlång sikt när det gäller att identifiera potential för förbättringar och att ta fram konkreta förslag till åtgärder är följande:

1) Möjlighet att förbättra ansatsen med kombination av logitmodell och nätverksmodell i första hand genom utveckling av logimodeller som är specificerade på ett bättre sätt för långväga resande

2) Fortsätt den undersökning som redan påbörjats i studien av möjligheterna att förbättra modellansatser som bygger på RDT-algoritmen dels genom att på lämpligt sätt även beakta olika aspekter av heterogenitet hos resenärerna dels genom att undersöka förutsättningarna att använda andra fördelningar av önskade avresetider och avgångstider än den uniforma fördelning som används idag. Ett annat viktigt utvecklingsområde som vi identifierat för RDT-linjen är att närmare studera möjligheterna att ta fram en vetenskapligt välgrundad metod för kalibrering.

3) Undersök möjligheterna att kombinera RDT och logit genom att på olika sätt använda beräkningar av generaliserade kostnader från RDT vid estimering och tillämpning av en logitmodell för färdmedelsval och/eller för modellering av total efterfrågan och dess fördelning på relationer.

4) Studera närmare möjligheterna att använda en tidtabellsbaserad modell. Ett första steg kan vara en förstudie som närmare belyser egenskaper mm. hos två existerande sådana modelltyper nämligen Sampers och Visums tidtabellsversioner.

I anslutning till var och en av dessa utvecklingsvägar är det nödvändigt att studera och förhålla sig till hur beräkningar av förändring av konsumentöverskott bör göras.

Vi har pekat på komplexiteten och den mångfacetterade naturen hos det långväga resandet och hur detta resande bestäms. På lång sikt bör man väga in en totalbild av detta i modellutvecklingen, där man i varje fall till viss del kan frigöra sig ifrån tidigare modellmässiga låsningar och investeringar. Därför föreslår vi att ett forskningsprogram för modellering av långväga resande tas fram; detta arbete kan förhoppningsvis dra nytta av det arbete som gjorts i detta projekt för att klarlägga olika aspekter och ansatser som man måste förhålla sig till i ett långsiktigt utvecklingsarbete.

1 Introduction

1.1 Background

This report presents the results of the project “Possibilities to combine passenger transport models for long distance travel”.

The aim of this project has been to find out whether there are possible combinations of elements from the currently used models in Sweden, and possibly also from other sources, that would give substantial improvements of predictions of travel distribution between modes and lines, as well as more precise and valid consumer surplus assessments for transport policy measures directed towards long distance public transport.

During the last twelve years the so-called Sampers model has been developed and used in Sweden. For long distance trips two versions were developed – one using a departure time (“timetable”) approach and one uses an average frequency approach. Only the latter has been used in practical work. That version, called the standard model, uses 3 steps: i) the network model Emme/2 for assignment on lines within each mode and for the estimation of travel time components for each mode, ii) a nested logit model for demand projections concerning modes, destinations, and travel frequency, iii) the Samkalk module for calculation of consumer surplus, revenues, costs etc. for cost-benefit analysis.

In parallel a network model Vips, applying what we denote the RDT-algorithm (stemming from **R**andom **D**eparture **T**imes), has been used for simultaneous assignment across both lines and modes in one step, and for the calculation of revenues, costs and consumer surplus. Since Vips is not a complete demand model the demand matrices for Vips have been generated by Sampers.

One background for the work is that problems related to both Sampers and Vips have been observed over the last ten years, and this has cast some doubts about the relevance and reliability of the application of model results in the context of transport policy evaluation and planning.

1.2 Examples of use of the models

For assessment of public transport the Sampers model has been employed by the Swedish National Railway Administration (BV) and the Swedish Institute for Transport and Communications Analysis (SIKA) for about a decade, especially for the Governmental requirements to produce “action plans” for the transport sector. The model has also been used in a number of studies on various specific road and railway projects.

The Vips model was used at SIKA in the late 1990s for assessment in the context of a Governmental inquiry on whether deregulation of coach services is socially beneficial, also taking into account whether the railway would “suffer” too much. The assessment led to a recommendation of deregulation and that also became the Governmental decision in 1998.

Vips and Sampers have also been used in parallel on behalf of BV, not least for various assessments of high-speed rail in Sweden.

In 2008 a hybrid model was used for a Governmental inquiry on high-speed rail in Sweden. The hybrid involved the use of Sampers for the creation of the travel matrix, Vips for the estimation of demand on lines and modes and for the calculation of travel volume weighted averages of travel time components and price for all origin-destination (O-D) pairs. Finally Samkalk was used for the calculation of consumer surplus and for cost-benefit analysis. In a parallel study Vips was also used for the calculation of consumer surplus and for cost-benefit analysis.

The interesting but somewhat “troublesome” outcome was the substantial differences in the cost-benefit results from the hybrid method using Samkalk and the Vips method based on the same travel matrix. Using the same project cost estimates, the net present value ratio ((Benefits-costs)/investment costs) was 0.15 according to the hybrid model and 0.78 according to Vips. Expressed in billions of SEK the pure Vips method in comparison with the hybrid method gave a change of consumer surplus that was 50 billion greater and a producer surplus change about 15 billion greater.

1.3 Sampers according to users

Since Sampers is the standard model tool in Swedish transport planning it has been used extensively by all transport agencies since it was first introduced around the year 2000. Over the years there has been an ongoing amendment process often triggered by the users’ experience and suggestions. However, while there have been many amendments to the Sampers model since it was first introduced, these have mostly concerned things other than the demand model. The demand model has been updated only once, and even then it was constrained with respect to model specification (Transek 2004). In this context, some specification tests were also carried out.

In 2009 a project on Sampers experience and the long and short term development possibilities was financed by the Swedish Road and Rail Administrations (Algers et al 2009). This work also included interviews with users. Though most of the users considered that Sampers is generally useful for its intended purposes there was also criticism. Among critical viewpoints were that the model cannot realistically enough describe the effects of relevant transport measures, and that it shows unreasonably small shifts between modes, air to train, car to train when high-speed trains are assessed. It also appeared that some users said that they have little experience of running the model, which may be one reason for the critical viewpoints. Some users, according to the report, argued that it may have to be further developed or even replaced by a new system.

One recurring criticism over the years has addressed the currently applied headway-based version of Emme/2 which operates with the assumption that travellers do not know the departure times for long-distance journeys, so that they are assigned in proportion to frequency only (for the set of acceptable lines). Some criticism, although not from the users, has also addressed the problem that the model uses fares per OD rather than considering that various connections may have different prices. However, the version of Emme/2 currently being used remains the same as originally implemented.

1.4 Other models – Vips and Visum

Besides Sampers, which has clearly dominated as a model tool in Sweden over the last decade, it seems as if only the model systems Vips and Visum have been used in Swedish practice, though other traffic demand models are also available in an international perspective.

The main arguments put forward by their users in favour of using Vips and Visum have been these systems' properties for dealing with timetable -based travel, as well as the treatment of combinations of modes and lines in public transport. Note that within the Visum model shell it is possible to implement a timetable-based modelling approach, though this functionality is not available in Vips.

Vips includes a forecast method based on assumed elasticities with respect to change of generalized cost, which gives change in number of journeys per OD-pair but no change of destinations. However, this method has not been applied in this project, and therefore the application of Vips has been restricted to analyzing how a given demand for journeys in various relations is distributed over private and public transport modes and, in the latter case, lines. Hence, applications of Vips have had to use as input origin/destination (OD) matrices of travel demand from Sampers. The combination of two different model systems obviously means that there will be an inconsistency between the mechanisms and assumptions used to generate the demand matrices and the assignment of these matrices in Vips.

The much less widespread use of Vips and Visum has naturally not evoked as much criticism based on user experience as for Sampers. Critical viewpoints about Vips that have been voiced deal with the likely shortcomings with respect to modelling random components of travellers' travel decisions, dealing correctly with car demand, and the seemingly somewhat arbitrary calibration procedure. The latter may bring the model's predictive power into question.

1.5 Reformulating the project aim

Discussions about the development of demand models for long distance travel have often taken as their point of departure the scope and properties of existing model implementations as well as existing practices. An example is the project "Analyses of public transport measures – comparative tests with the model systems Sampers and Vips" (SIKA Report 2009:3).

By and large this project was based on the Swedish experience with transport models – in this case for long distance travel – over the last decade. With this background it seemed natural to ask whether it would be possible to improve the overall performance of modelling long distance travel by finding better ways to combine the use of the model systems that had been used in the Swedish practice as described above.

However, it soon became obvious that this formulation of the project's task was too restrictive. The initial work had made it clear that there were fundamental issues that could not be resolved merely by combining elements from two or three existing model implementations. Moreover, the likely considerable effort on the user side of introducing a new model system was an argument for widening the project's remit to include a discussion of possible amendments beyond the "model combination" issue.

Partly this was because the model implementations of Sampers, Vips, and Visum currently in use were not necessarily seen as suitable as the basis for long term commitments from the transport agencies and traffic companies. The ownership of the Vips system has been taken over by PTV (Germany), the owner of Visum, and Vips is no longer maintained and is no longer commercially available. Visum on the other hand is continuously maintained and developed and also comprises a demand model shell that also includes destination choice.

Based on these considerations it was decided – in consultation with the commissioners – that the project should widen its scope. Therefore, the revised aim of the project is to give recommendations as to the way ahead for the further development of long distance travel demand models. These recommendations should be based on general considerations of user requirements of such models as well as recent research and development within the field.

This note does not include any **detailed** descriptions of existing model implementations in Sweden in relation to an ideal model. However, to find the way ahead it is necessary to make use of the extensive experience with existing model implementations. Therefore brief descriptions of the standard Sampers and Vips models are included. There is also a short discussion of the advantages and disadvantages of these two models. Of course, the results from the in-depth analyses of the current standard Sampers model and the Vips model that have been carried out in the project has been used where applicable.

1.6 Structure of the report

This report includes four parts.

The first part (chapter 2) discusses in general terms the user requirements and main issues in demand modelling for long distance travel. Prerequisites, assumptions, functionality, and scope that should characterise an "ideal" model or approximations of such models for long distance passenger journeys are discussed. The meaning of the word "ideal" is here assumed to be related to the perspective of the users of the model.

The second part (chapter 3) describes two of the model implementations that are currently in use in Sweden: The standard Sampers model, and VIPS. It then discusses some properties of these models partly based on real test runs of their current implementations. In this chapter we also try to highlight similarities and differences between the model implementations and to make observations on particular properties of the model systems that are either important to preserve or need to be amended in the further development of demand models for long distance travel.

In the fourth part, chapters 4 and 5, we try to create an analytical foundation for the development of Swedish models for long distance travel, and to define the way forward. Based on earlier discussions on requirements in part 2 and the problems with the existing models that we have identified in part 3, we analyse in chapter 4 some key issues and how they could be addressed in further research and development work. One aim here is to find out if, and how, elements from the models in use can be combined in order to achieve substantial model improvements. Another aim is to discuss whether and how substantial modelling aspects that are in neither of the existing models might be added. This will point forward to suggestions for the further research presented in chapter 5

This report is intended to reflect consensus views among the authors. In addition to this report, work has been undertaken by different authors to provide input into the research process. These inputs have been collected in a separate report,"Descriptive and theory report", including individual contributions of some of the members of the project group. This report is available as a working report and reflects the views of the various respective authors: it is not necessarily the consensus view of all the authors of this report. The contents of this background report with indications of authors various parts are to be found in appendix 3.

2 Issues that have to be addressed in approaching an ideal model for long distance travel

2.1 The role of an ideal model

An ideal model would allow the users to analyse the policy-relevant effects on demand for long distance travel of various measures relating to infrastructure, traffic, and transport policy with validity as well as sufficient reliability and precision, all at an acceptable cost. The yardstick against which the adequacy of the model should be measured is of course actual travel demand. Is the model capable of replicating actual demand at the required level of detail, as well as projecting demand changes in response to a set of policy measures? Since data for all relevant components of actual demand are not necessarily available or even observable, e.g. details of demand for certain lines and departures as well as future demand, the reference yardstick is in most cases incomplete. Therefore it is mostly not possible to fully establish the degree to which the model has the desired properties.

Transport demand is often a key variable in transport policy analysis, but the degree to which demand has to be analysed at a disaggregated level depends on the type of policy that is analysed. Though a model for transport demand at the most disaggregated level will usually allow demand to be summed to more aggregate levels, the mechanisms and data requirements for a disaggregated model may be unnecessarily cumbersome and expensive for analyses of certain policies, e.g. fuel taxes, which only require demand data at more aggregated levels. Different models may therefore be “ideal” for different situations.

We cannot therefore be sure that it is possible to devise an ideal model in the above sense. We have to draw on an existing body of knowledge about travel behaviour and various sets of axioms, other assumptions, and mechanisms to model such behaviour. We may be further restricted by availability of data. It would be attractive, of course, to be able to try out a large set of promising combinations of axioms, other assumptions, and mechanisms in order to find out which combination shows the closest resemblance with the perfect (ideal) model and thus actual demand. However, the process of defining and implementing a fully consistent demand model is far too time and resource consuming for this to be a viable option. Instead of a comprehensive model search and evaluation we will have to rely on partial judgements about specific assumptions, features, and model mechanisms.

Obviously it is difficult to make assumptions and to define features of an ideal model without any influence from the long modelling tradition in the field. Such influences may in fact give indispensable or very useful information about the properties of certain model components. Existing model implementations provide examples of combinations of model features that may give important information about the properties of these combinations. Such information will be useful to a process aiming at coming closer to the ideal model.

This chapter, which contains the first part of the report, is structured in sections in the following way. Section 2.2 contains a very brief discussion of user requirements. Section 2.3 deals with the scope and purpose of the model. In section 2.4 we discuss travellers’ preferences and wishes, and in section 2.5 the properties of traffic supply and suppliers’ behaviour. Important issues in the modelling of long distance passenger travel are discussed in section 2.6, in the light of experience from existing model implementations and their application. Considerations about data requirements are included. In section 2.7 we move towards model design, and start by discussing behavioural assumptions. We then continue to discuss the pros and cons of different approaches for dealing with the various modelling issues.

2.2 User requirements on model support for long distance travel

The users are involved in public sector planning processes, design of regulatory policies, and design of economic support/subsidies, as well as in the planning of traffic supply and supply policies. From time to time alternative individual or packages of measures are considered for implementation – both separately by public and private sector actors and sometimes in public-private co-operative programs.

The role of the long distance travel demand models is to supply relevant and reliable information about various aspects of travel demand to the users' decision processes. Of course, it is conceivable if not likely that the need for information differs between various users, e.g. between users interested in infrastructure development, public regulation viz. private public transport companies e.g. analyses of competition in transport markets. Therefore, it may well be that there is a demand for models with different properties from different users or user categories. It is also possible that the needs for each user category differ between different decision situations, which may lead to demand for models with wider scope and versatility, or more than one model.

For the further discussion of an "ideal" model it is necessary to be clear about which user or users are addressed here.

We define the following user categories:

- Public Infrastructure Agencies (previously these had a "modal" profile – Vägverket (VV) (bus, car, road ferries), Banverket (BV) (train), Sjöfartsverket (maritime), Luftfartsverket (LV) (Air), but now Trafikverket (TV) will handle road, rail and certain maritime and air planning issues).
- Other public agencies involved in financing and planning of long distance passenger traffic (Eg. Rikstrafiken, Samtrafiken, Regional Public transport agencies, Kollektivtrafikmyndigheten, Trafikanalys)
- Regulatory body for use of the transport infrastructures (Transportstyrelsen,)
- Public transport companies that are run by the public or the private sector

The public sector actors typically are confronted with decisions about:

- Dimensioning of capacity and reliability of the transport infrastructures
- Dimensioning of capacity and properties of transport (service) supply
- Distribution of infrastructure and service capacity between different users
- Efficient allocation of public funds between alternative projects or schemes
- Competition in transport markets
- General distributional issues relating to transport policy measures

A key element of the approach to decision-making in the cases mentioned above is information about travel demand, as well as how travel demand is affected by alternative courses of action.

2.3 Scope and purpose of the model

Obviously an ideal demand model seen from the model user's perspective must cover all aspects of long distance journeys, i.e. the overall level of demand, the distribution of travel demand on lines and modes for all origin-destination (O-D) pairs. From the user's perspective it is essential that the demand model also allows a consistent calculation of the change of consumer surplus (CS) resulting from various transport policy measures. However, in this context we are focussing on only a part of the overall demand modelling problem, namely the distribution of **long distance** travel demand over lines and modes, while keeping in mind that this model component also will have important interactions with other parts of a comprehensive demand model. In those parts of our discussions where we compare existing model implementations, the limited scope of some models makes it convenient to assume that travel demand expressed as a (set of) flow matrix(ces) is exogenously given from other model modules.

The definition of "mode" is problematical, for two main reasons. Firstly, while the broad descriptions of "air", "car" etc. are clear, more detailed descriptions raise issues as to what the essential qualities of a "mode" are (e.g. is high-speed rail a different "mode" from standard rail?). Secondly, a typical journey may involve combinations of "modes", so that "mode" may be seen as a description of particular services, rather than of the whole journey. These difficulties will be elaborated later in this document.

There is also potential confusion in the definition of other concepts used to characterise public transport (PT) such as line, route, transport services, travel path etc. We will deal with these definitions in the relevant context below.

The model should cover traffic by all current and new lines and modes for long distance travel within Sweden.

The main purposes of the limited part of a demand model we are discussing here are to:

1. Generate a reliable estimate about how total travel demand (the relational flow matrix) will be distributed among modes, lines and routes, given the total traffic supply, existing or anticipated, possibly including new modes.
2. Give output data that allows the estimation of the change of consumer surplus of individual or packages of public or private measures relating to infrastructure or long distance traffic supply in a way that as far as possible ensures consistency between the demand model and the CS calculation.
3. It should be possible to separate demand between different journey purposes such as work trips, business, and leisure trips, as well as between various socio-economic-demographic categories, for analyses of distributional issues.

Long distance travel is a small subset of all travel (even if in kilometre terms it is more important), so that decisions about the appropriate level of detail become critical. At one extreme, one might require demand for individual services at particular times of day, while at the other a broad allocation between "modes" in specific corridors (e.g. Stockholm to Göteborg) might be sufficient.

2.4 Traveller categories and travellers' preferences

Travellers' preferences and valuation of various aspects of travel are expected to influence travel behaviour. The range of preferences and values that are relevant to the travellers' decisions about long distance travel has been investigated empirically but there is scope for further study of such preferences and values. We are, of course, aware that additional empirical evidence could influence the range of factors that should be paid attention to, as well as leading to changes of the judgement of their relative importance. There is also scope for introducing specific empirical evidence about the precise distribution between individuals for each of these factors. Such evidence could have direct implications for the appropriate assumptions about e.g. the form of statistical distributions to use and/or on the choice of structure of traveller segments.

Based on empirical evidence, a priori considerations, as well as experience from earlier modelling and validation efforts the following general assumptions are made here about travellers' preferences and values:

- Ideal departure and arrival time vary between the individuals for each journey relation and journey purpose and can also vary for the same individual depending on day or situation
- Travel decisions are typically taken on a tour basis, i.e. outbound as well as homebound trip legs are considered
- The cost per time unit associated with early and late arrival varies between travellers and for the same traveller, and could moreover vary depending on journey purpose
- Travellers have individual preferences for travel with different modes, both as a constant preference and as one related to O-D relation and/or distance
- Value of time for each travel time component varies over the population of travellers, both with respect to journey purpose and categorisation to socio-economic-demographic group

2.5 Characteristics of long distance transport supply

In the existing model implementations that are used as planning tools in the Swedish transport sector, public transport supply is fully exogenous in all time perspectives, while private car transport supply (car ownership, car availability) to some extent is endogenous. Congestion is disregarded in the context of long distance travel. The treatment of congestion in the Sampers system is explained in section 3.2.For the purpose of this note we stick to the assumption of exogenously determined supply. The influence of supply characteristics on demand will thus be mediated directly as travellers' cost components in the demand functions of the demand model.[1]

However, in order to be able to quantify the cost components that strongly influence demand, it is important to develop a good understanding also of the properties and mechanisms determining transport supply, and to devise procedures for defining and calculating all demand-relevant cost parameters.

---

[1] For the current situation the planner needs to ascertain as far as possible the exogenous supply of each operator. For a future situation the planner will have to exogenously define the transport measures of the supplier that are to be assessed. Further, without specified information by the operators, the planner can only make a qualified guess with respect to changes of other suppliers. Since the measures under assessment may affect the supply of other lines and modes, the planner has to estimate reasonable supply changes of these affected lines and modes based on the impacts of the model estimate of demand changes.

Long distance public transport supply is today mainly produced by the commercial sector (air, long distance bus, long distance train). For this sector it seems reasonable to assume that each company in the long run cannot incur losses. Cross-subsidisation between lines within a company will only take place if there are tangible network benefits that could be assumed to outweigh the financial loss on certain lines. In the absence of such benefits each line will have to be run to break even financially.

Competition will limit the market power of supply companies; however, considering the fact that in many cases there are only few and large actors, it is reasonable to assume that some monopolistic pricing will continue to exist in many relations in the long run.

Timetables and fare structures will therefore increasingly be optimised per company as well as per line and per departure in order to maximize profits and to ensure the long term financial viability of each transport supply company. We note that this feature has the potential of leading to rather complex fare and price differentiation structures.

Over time there will be considerable structural changes in the supply markets. The impacts of such changes should be carefully considered in model development.

Certain parts of the traffic supply are produced in subsidised and regulated sectors (Regional traffic that could take on a role for long distance travel, as feeder services for example, as well as particular long distance lines that are subsidised at the national level by Rikstrafiken or Kollektivtrafikmyndigheten).

The structure and institutional setting of transport supply as discussed above makes it a difficult task to define those characteristics of supply that should be the basis of calculation of the cost parameters going into the demand functions. Any supply scenario will be uncertain, causing also uncertainty for the cost parameters. To deal with this problem it might become necessary to carry out sensitivity analyses with alternative scenarios. For such an approach to be realistic, however, it is necessary that a demand model is not too cumbersome, costly and time consuming to apply.

2.6 Modelling issues

2.6.1 Introduction

The task of the modeller is to design a model or a set of models that gives sufficiently reliable information about different aspects of long distance travel demand at the lowest, or, in practice, at a reasonable cost. The model, or the set of models, should provide an acceptable approximation of the rather complex reality of long distance travel that was outlined above. The dimension of development time also has to be considered. The cost of model development has at least two dimensions, namely development cost – including cost of data for model estimation/validation/ calibration, and the cost of using/running the model in terms of computer run time, time to analyse data etc.

Before embarking on the modelling exercise it is essential to develop a good understanding of the issues that need to be tackled, as well as about the specific characteristics of these issues. The discussion in this chapter aims to take a step towards developing such an understanding.

2.6.2 Important characteristics of long distance travel

It should be observed that the items listed below as particular characteristics for long distance travel decisions are at least partly behavioural hypotheses that may be useful to test empirically before embarking on a modelling effort conditioned upon the validity of any of these assumptions.

- Compared to short journey travellers, long journey travellers could be expected to spend more time for planning the journey. This will in all cases include seeking information about travel time, timetables, fares conditioned upon departure/arrival time, considerations about combinations of lines and modes and access/egress legs of the door to door transport chain.
- Tickets are often constrained to a particular departure. Price and conditions may depend on the time of purchase in addition to the time of departure. The planning (booking) process may have to be addressed to take this into account.
- Travellers could be expected to explicitly consider departure/arrival time in relation to the traveller's preferences for arrival at the final destination. The importance of an exact arrival time is likely to vary considerably depending on the traveller as well as the travelling situation or purpose.
- The current Swedish definition of long distance travel as all journeys of more than 100 km (one way) will lead to a considerable heterogeneity with regard to journey length.
- For longer journeys, one would expect that dimensions of the journey other than travel time and fare, e.g. comfort, safety, or travel company considerations will play a greater role than for short distance travel. e.g. luggage may influence travel decisions or for some groups of travellers it may be important to have access to personal service and contact with staff involved in production of the transport services.
- The duration of the stay may be very different compared to short journeys. This may influence preferences, for example by different time constraints.
- The duration of the trip may vary substantially with respect to mode. Switching from a fast mode (air) to a slower mode (car) may imply extra costs like overnight costs that are not associated with the mode as such, but are important implications of choosing the slower mode.
- Variability of travel terms and conditions as well as fares of public transport supply is generally greater for long distance travel

Some important issues in the model users' perspective have been touched upon above in section 2.2.To deal adequately with the users' requirements the modellers need to carefully observe the specific aspects of this particular type of travel. Some examples are given below

- How to deal with heterogeneity of preferences in many dimensions (see above) between traveller groups and purposes. How to categorize?
- How to deal with supply dimensions other than timetable, such as capacity, comfort, fares/fare structure, reliability, personal service, connections to other means of travel etc.?
- How to deal with timetables? By using a full list of all timetables, which would be difficult for future situations as well as cumbersome for any references situation, or by using a simplified representation such as only specifying the ride time and headway for each line (per mode)?

- What combinations of travel route alternatives should be considered for each travel decision and how should travel demand be distributed between various travel alternatives i.e. what type of model should be used for this distribution?
- How to deal with issues relating to calculation of consumer surplus and overall social benefits and cost?

Each of these issues will be discussed briefly below in this chapter as a background to the further considerations about the alternative approaches to modelling.

2.6.3 What traveller characteristics should be included in a model and how?’

First, the decision maker (the traveller) needs to be defined. Although **individual** trip making is often implicit in models for short trips, private long distance travel is often undertaken by a group of people (usually referred to as the travelling party), often a whole family or some other group of people travelling together. In addition, in contrast to short trips, business trips form a considerable share of travellers on some modes. Business travellers may be subject to constraints imposed by the employer in such a way that their behaviour at least partly would reflect employer preferences, as given in travel policies and travel cost remuneration schemes. Consequently, we may need to think of travellers as travelling parties and/or employer-constrained employees.

Traveller characteristics such as preferences, values and categorisation on socio-economic-demographic groups were discussed above in section 2.4. There is considerable knowledge about how various preferences and values influence travel choice, but this knowledge is by no means complete and definite. There is scope for argument as to what aspects should be included in a model and in what way that should be done. Traveller preferences that must be considered by the model developer specifically include:

- Variations of values of time among travellers
- Preferences regarding which modes to use for long distance travel
- Preferences regarding time of departure and arrival

What do we know empirically about all these aspects and their distributions over travellers? Should new empirical data be collected for some aspects? How useful are conventional socio-demographic categorisations as a means of generating segments of travellers that could be expected to take travel decisions based on similar criteria and values? One implication of the travelling party concept is that socio-economic categorisations describing the party may be more relevant than those describing one of the individuals in the party. Another implication is that travel costs should be specifically considered for the whole party.

There is considerable heterogeneity of the collective of long distance travellers regarding socio-demographic-economic characteristics, as well as the variability of various preferences and values, which may or may not be related to socioeconomic groups. There are reasons to believe that this heterogeneity also influences travel choices in various ways. One aspect is, for instance, that different economic categories could have different marginal utility of money. The variability discussed here makes it important to consider whether the models’ quality could be improved with appropriate differentiation of parameters for different population segments, in place of applying a uniform utility maximization approach over the whole population of travellers. The wide range of the duration of the stay may imply a wider range of time constraints as compared to short trip making. This has previously been treated by segmentation on the duration of stay.

2.6.4 Issues relating to description and representation of supply

In principle one could give a full account of various aspects of supply that transport companies define and that are related to aspects which could be assumed to have some impact on the travellers' choices, e.g. the network (nodes and links), the lines, timetables, fares, capacity, reliability, comfort, feasibility of certain types of travel etc and other characteristics of supply. The task for the modeller aiming at the "ideal" model, however, is not to include every possible supply dimension but to pick and/or generate a limited number of key supply variables that are consistent with the required properties of the model. The model implementations that are currently used in Sweden have a common network basis, but deal with the further details of supply in different ways as well as at different levels of detail. This is not to say, however, that **any** of these implementations is using a sufficiently good representation of supply, so that the issue is well worth considering in the course of defining the ideal model.

As pointed out above, there are good reasons to assume that timetables are used by the travellers in long distance travel by public transport. The timetables are defined by the transport service companies with the view (it may be assumed) of suiting travellers' travel patterns in time and geography, and operators are using higher frequency as one of their competition parameters, which is due to the fact that the more departures there are, the more likely it is that travellers will find a departure that suits their ideal departure or arrival times. Sometimes timetables may be restricted by the availability of infrastructure, such as "slots" for air departures as well as rail departures. Since timetables may vary by weekday, season of the year, holidays, as well as due to other ad hoc reasons, the total amount of data to fully describe all timetables that are relevant for long distance travel is really large. However, it should be noted that in recent years there has been some development of more or less comprehensive databases with timetable data, and these may be useful also in the context of demand modelling.

The modeller must consider how to model the overall timetable structure as well as the appropriate level of detail, in order to be able to model travellers' choices of travelling path (route) with sufficient precision. One extreme is to include the full timetable information in the model and its database, and from this extreme various simplifying approximations should be considered.

2.6.5 Issues relating to travellers' choice sets and choice behaviour

Different assumptions about the choice set (travel alternatives that are available for traveller consideration), as well as about the structure of the decision process, could in practice co-exist in the model with different choice criteria and decision rules based on alternative behavioural assumptions. The choice set in model practice is related to assumptions about travellers' decision structure and search process, as well as to the mechanisms used to search for travel alternatives. In the current modelling practice the decision unit is either assumed to be a group of person travelling together or an individual traveller. The choice of decision unit may relate to the type of journey. It may also be relevant to consider influences on the travel decisions that are exogenous to the decision unit, e.g. the influence of an employer's travel policies.

What travel alternatives travellers actually consider for long distance travel, and what their decision structure actually looks like, are by and large empirical questions. Such empirical knowledge should influence the assumptions about the choice set and hence also the mechanisms that should be used to generate alternatives. Therefore, it is important in the context of Swedish long distance travel to analyse what we know empirically about actual choices and travel paths – c.f. TDB (the Tourist Data Base) and RES (the National Swedish Travel Survey. Some results from our investigations about how combinations of modes are used in practice in Sweden will be given in chapter 4 below.

Travellers (consumers of transport services) are often assumed to maximize their utility. However, travellers' behaviour in reality could involve more satisficing than optimizing. An overall optimisation for the traveller that also includes the time and cost of searching and processing information (sometimes called "executive efficiency" in literature) may have the effect that travellers could be inclined to stop searching for alternatives when they have found an acceptable number of feasible travel paths that are consistent with their basic preferences etc. In fact, this remark highlights the fact that behavioural assumptions are themselves simplifications intended to facilitate modelling.

This is not to say, however, that a model assumption about satisficing instead of cost minimisation behaviour will necessarily have a significant influence on the model projections of travel demand. One reason is that alternatives that are far from the best ones will usually get very low choice probabilities, which means that the total effect of including all alternatives will be negligible. There would be more significant differences if a satisficing search process were to systematically end up far from the objectively optimal alternative(s): this is probably not very likely, however.

Based on the principles discussed above, various ideas and arguments about the traveller choice set have to be scrutinized.

2.6.6 Issues relating to calculation of consumer surplus and overall social benefits and cost

As was mentioned in the section about user requirements, one important task of a model for long distance travel is to provide – directly or indirectly – reliable and valid data to allow consistent estimates of the social benefit of various policy measures and actions taken to influence supply of transport services for long distance travel. The generally accepted framework for such estimates of social benefit in the context of transport policy is based on economic (welfare) theory including theory about consumer (traveller) utility. The benefit/disbenefit of some measure in this context is normally defined as the sum of changes of travellers' utility resulting from the measure, expressed in monetary terms.

The utility change is generally estimated numerically as the change of travellers' consumer surplus, based on the links between the utility function, the demand function and consumer surplus. Note that "*Adding the consumer surpluses for the market as a whole involves assuming that the marginal utility of money is equal for all consumers; if there are wide differences in income this is not plausible*"[2].

Aggregate change in consumer's surplus can be positive, but the rich can still be better off and the poorer worse off after the change, especially when multiple changes are involved simultaneously. A positive change in the aggregate measure of change in consumer's surplus only tells us that the "winners" - in principle – should be able to compensate the "losers", but compensation rarely takes place.

[2] Black, J, Dictionary of Economics, Oxford

The argument that makes aggregation defensible – given that compensations are not paid - must resort to the concept of a welfare function reflecting the prevailing political preferences i.e. we assume that the existing income distribution is in accordance with the prevailing welfare function. An "optimum" income distribution will also mean that the "welfare function" is *indifferent to marginal changes* in the existing income distribution.

However, the prevailing income distribution is generally not considered to be "optimal" for the purpose of calculating the aggregate change of consumer surplus due to transport policy measures. Therefore the aggregation problem is often in practice addressed by using values (e.g. values of time) that are decided politically in the context of normative C/B calculations; these values may differ from the "behavioural" values used for demand estimates.

In transport modelling, the concepts generalised cost (GC) or generalised price (GP) are often used. These concepts aim to include the cost of all inputs carried by the traveller for production of the travel service, part of which may be captured by the fare paid, and including the input of time, discomfort etc by the traveller.

For a specific traveller this can be formalised using the following notation:

P_i	Price of mode i
T_i^j	Travel time component j for mode i
V_i^j	Value of time for time component j for mode i
A_i	Value of other attributes for mode i expressed in monetary units

The GC expressed in monetary terms for mode i can be written as follows:

$$GC_i = P_i + \sum_j V_i^j \times T_i^j + A_i$$

V_i^j can thus differ between modes, i.e., reflecting the comfort of each mode.

It is in most cases reasonable to assume that the *change* in generalised cost is equal to the *change* in utility, but this may be modified by concern for variations in the marginal value of income.

Calculations of consumer surplus mainly play a role in decision making e.g. in the choice between alternative lines of policy action. In this context it may be required by the decision makers that e.g. the value of time savings should be considered equal for all travellers, though we know that in reality value of time savings differ widely between groups and individuals.

2.7 Towards model formulations – alternative approaches to specific modelling issues

2.7.1 Introduction

Over the years different approaches and tools have been used to develop and implement approximations of traveller behaviour. The usefulness of various such approaches is inter alia influenced by the availability of general software, available computer capacity and speed.

In the discussion of the development of models it may be useful to make the following distinctions regarding the level of precision of the model formulation:

1. **Conceptual models.** This is when we describe models with functional symbols and “abstract” variables. We may also use boxes and arrows at this stage. We usually make a minimum of assumptions about these functions, possibly only the sign of the derivatives for the variables.
2. **Mathematically fully specified models**. This is when we have chosen specific mathematical functions that may be more or less restrictive and convenient.
3. **Numerically implemented models**. What appears as parameter symbols in (2) now has numerical values and we need precise definitions of the variables that may have rather vague meanings under (1) and (2). The variables should also be observable and measurable.

Numerically implemented models (3) are what we need for practical work, but to arrive there we usually go through (1) and (2) and the different steps of estimation and implementation.

Behavioural assumptions, characteristics of travellers and transport supply, as well as the characteristics of long distance travel decisions, are expressed formally in the models with suitable approximations that do not compromise the quality of output information about travel demand too much.

Before we enter a more detailed discussion of model assumptions and approaches it is appropriate to make a few general observations on the practice of modelling transport supply and demand.

The conventional treatment in transport models is to use network models (typically referred to as “assignment” models) to provide cost estimates for different alternatives, and to use discrete choice models (typically of the logit form) to allocate demand among alternatives on the basis of these costs. Because mode choice is usually an important feature of transport demand, this implies that the network models must deliver costs for each mode, separately for each O-D movement. Typically, the network models are developed for different periods (e.g. peak and off-peak, etc).

In practice, in constructing the costs the network models consider the choice of route through the network, even though this can be viewed as an aspect of demand. The choice of route may depend on the characteristics of the traveller. For highway demand it is usually assumed that travellers will use the minimum generalised cost route, though there may be multiple routes which deliver the same generalised cost (so-called “equilibrium assignment”).

The case is more complex with public transport, where there is a need to represent separate “lines” and the frequency with which they operate. Because of the discrete (in time) nature of public transport services, it is necessary to make some allowance for the temporal distribution of demand within the period being modelled. In this way, travellers are no longer all allocated to the alternative that is the minimum cost alternative without regard to the temporal distribution.

In practice, the relative spacing of the **services** is likely to be important, especially when there is a choice between “fast” and “slow” services. To give a very simplified example, if we have two services both operating on a 20 minute frequency, with one taking 10 minutes longer than the other, then no-one knowing the timetable would take the slower service if it was timed to depart more than 10 minutes after the faster service, whereas if it departed less than 10 minutes after the faster service, it could be expected to capture some of the demand.

The only way to deal with this rigorously is to adopt a **timetabled** approach. Algorithms are now available to do this, but they are demanding both in terms of initial data and computation time. Most of the use of public transport assignment methods is in the urban context, where the frequencies are sufficiently high that the finer points of timetables are less important. In such cases, it is usually considered sufficient merely to represent the frequencies of the competing services (in addition, of course, to their costs and in-vehicle times). We can distinguish between use of timetable information ex ante (before departure from the origin) and ex post (the information is revealed at the boarding point by displays or posted timetables etc). The ex post case may be the most realistic for urban situations.

Even here, there are alternative assumptions which can be made to deal with frequency. Probably the two approaches most often used are a) the "optimal strategy approach" (OS) (Spiess and Florian 1989) (as implemented, for example, in Emme/2) and b) the "random departure times" (RDT) algorithm (Hasselström 1981), implemented in Vips/Visum. Both of these assume that the exact departure times of services are not known by the modeller, only the frequencies and ride time.

A further issue for public transport is the issue of "sub-modes". In the context of long distance travel this may relate e.g. to different types of trains, e.g. high speed, comfort, ordinary, regional trains or for air transport lines with different standard and/or location of terminals, being more or less peripheral. The option is either to have a separate network for each sub-mode, allocating between sub-modes on the basis of a discrete choice model, or to use an integrated network in which all sub-modes are represented (though potentially with different characteristics and cost parameters), and allow the allocation algorithms to load people on to the different sub-modes, often for only parts of their journeys. Within the urban context, this latter option is more commonly used.

It seems as if experience of public transport network (assignment) models in the long-distance context, where frequencies are typically low, internationally is very limited, though for Sweden and in some other countries there is considerable experience of practical use of such models; in Sweden during 15-20 years. While it would be convenient if the same approaches could be used as are common in urban contexts, in practice the choice of approach would benefit from choosing the approach for long distance travel independently of the approach used for the urban context.

2.7.2 Behavioural assumptions

Some general remarks

In the context of transport demand model development, the purpose of behavioural assumptions is generally to simplify or even make it possible to define various model components. It is by no means obvious which behavioural assumptions to make in the course of transport demand model development. While behavioural assumptions generally aim to simplify the modelling process, there is also a risk that certain assumptions make the model flawed, since they are simply wrong or too far away from real world behaviour.

Generally modellers use behavioural assumptions that allow the use of certain analytical techniques and algorithms that are available from research. Furthermore, assumptions are mostly checked for consistency with (some) behavioural theory. Due to the general evaluation framework of transport policy it might be convenient to assume that travellers are utility maximising, though such consistence between behaviour assumptions and the evaluation framework does not per se guarantee a better forecasting ability for the model.

It must be remembered that behavioural assumptions may incur weaknesses in the model as well as criticism as to the model's validity. A notable example is the assumption in the first bullet point below in the next section, based on the notion of the utility maximising consumer or "economic man" – a notion which has been criticised from various scientific perspectives.

Behavioural assumptions about each individual traveller in the context of long-distance travel

- Each individual traveller (or group of travellers travelling together) is assumed to minimize the generalised cost (GC) of each tour[3]. Individual preferences for different modes affect the travellers' perception of GC. The generalised cost of travellers is also assumed to include a component giving the "schedule delay" cost of early or late arrival
- Travellers are assumed to have full access to, and to use, timetable information about all combinations of modes and lines in the course of planning and deciding on their journey route (travel path) in order to minimize GC.
- Travellers have full information about the (sum of) fares or other costs corresponding to money outlays for each alternative combination.
- Travellers could acquire and have some knowledge about the probability of delay and/or cancellation. However, the possible implications for modelling travellers' decisions are not discussed further in this paper.

Some of these assumptions are obviously rather unrealistic, e.g. the assumptions that all travellers have access to and use complete information about supply in deciding their optimum route[4]. The assumption of perfect information is of course very convenient from a modelling point of view. However, in modelling practice one has to allow for the fact that traveller information may be less than perfect, which can be dealt with by means of a random component.

Behavioural assumptions relating to suppliers of long distance traffic

Since we assume for the purpose of this project that supply is determined exogenously, there is no need to make any explicit assumptions about supply here. Let us only mention in passing that we assume that there are no supply restrictions that influence demand e.g. via congestion, or simply because there is no place available to accommodate demand for a certain transport route. Neither is there a guarantee that the supply that is offered is used at a reasonable rate. This means that the exogenously defined supply may well be inconsistent with the demand that comes forward at the exogenously given prices, available substitutes etc. Therefore it is necessary to ensure that supply and demand match reasonably well. In practical applications one ought to run the model iteratively and adapt frequencies and/or number of carriages per train (etc.) to make the load factors reasonable. All such consistency checks have to be carried out separately and are not assumed to be a part of the modelling discussed in this note.

---

[3] In the wider travel demand context which also includes choice of travel relations, the wider and consistent assumption is that each traveller maximises her/his utility.

[4] The assumption of complete information is perhaps not crucial. Usually less than complete information on the attributes of alternatives is sufficient to deem some (or most) as inferior to at least one other alternative. Travellers don't need to make a ranking of all feasible alternatives, they only need to find the best, and that may in many cases be less demanding.

2.7.3 Approaches to heterogeneity of traveller characteristics, preferences, and values

An ideal model has to deal with the issue of heterogeneity, which is complex as well as multi-dimensional. Approaches that have been tried are:

- Segmentation on population groups based on socio-demographic/economic criteria
- Estimating separate models for different segments, or allowing for different systematic influence of segment variables in model estimation
- Random influence on travellers' choice from one or more sources captured in one term or in more than one random term. If this approach is used the functional form of the random variable(s) becomes an issue.
- Endogenous segmentation where the segments are treated as latent classes to which travellers belong with a certain probability, for example defined by a logit model. These classes will reflect the most efficient subdivision of the sample with respect to preferences related to travel behaviour.

Very detailed segmentation requires large amounts of data. However, the developments of computing power and data storage capacity during recent years have made the amounts of data per se less of an issue. It is today possible to store, retrieve and analyse huge amounts of data without excessive use of time and cost. The more important restriction on such a detailed approach to segmentation is the availability of relevant data about the population of travellers. When such data are available today, e.g. about the variability of value of time savings between different groups, the information is often based on small samples. In such cases it is necessary to be careful with statements about small subgroups. In addition, the principle of "parsimony" in model building should argue against incorporating excessive amounts of detail.

There is limited knowledge today regarding the stability of preferences and values over time. It is possible or even likely that values and preferences change. Ideally this aspect should be observed in model development. In recent years some studies relating to long term changes of value of time have been carried out in Sweden.

2.7.4 Approaches to description and representation of supply

There is a wide range of possibilities when it comes to the representation of long distance transport supply. The modeller has to consider the following issues:

- Level of detail of the long distance network including infrastructure and terminals
- Level of detail of the description of lines and departures
- Level of detail when it comes to dealing with travel from door to door, e.g. whether a separate access/egress function should be used to relate journey origins and destinations to the long distance traffic network and in this case at what detail access/egress should be represented
- Timetables; how to include and at what level of detail
- Fares, fare structures, discounts (e.g. for students, retired)
- Operators and their roles and characteristics

Considerations that have to influence the decisions implicit in the bullet points above relate to the feasibility and cost of providing relevant input data for existing supply, as well as for future supply, when the model is applied for projections of future travel. The costs have to be balanced against the benefit in terms of validity, precision, and reliability of the model.

Network models are a well-established tool for dealing rigorously with the complexities of large traffic and infrastructure networks. However, the properties of the network model that is used in the context of modelling long distance travel must be compatible with its intended role in the overall demand model. Beside network models, one could consider using auxiliary dynamic programming approaches e.g. to identify feasible transport chains.

For a demand model structure that uses separate models for certain parts of the modelling tasks, e.g. modelling total demand, distribution over relations, and distribution over certain transport chains, it is important to keep in mind that the network model often has dual tasks. Firstly it is normally used to provide input data about level of service to the estimation of parameters of models that are used in other modules of the model system: errors in these estimates may cause subsequent estimates to be biased. Secondly, the network model is used in model application to assign demand to the relevant transport network

2.7.5 The model's level of detail

An important aspect for the model specification and design relates to the required level of detail. Of course, the choice of level of detail depends on the requirements of the users. The issue was briefly discussed above where we envisaged the co-existence of multiple models designed for different purposes. For models primarily intended for analysis e.g. of climate effects of the activities in the transport sector, other aspects and details may have to be in focus than for a model that is primarily designed to deal with infrastructure development and/or traffic supply. Are we interested in different times of day? Are we only concerned with the long-haul section of the journey, together with a more rudimentary approach to access/egress? Are we interested in the distribution of air travel over low cost vs. scheduled airlines etc. for the analysis of certain transport markets and competitive behaviour in these markets? It is important to answer these and other related questions at an early stage – already at the conceptual model level.

2.7.6 Approaches to choice sets that are assumed to be considered by travellers and to the structure of travellers' decisions

The most interesting aspect of the travellers' choices in this context is of course the outcome. What the modeller therefore should aim at is to use a combination of decision structure and choice set that has the capability of giving reliable results as to the outcome of choices. We consider the following alternatives:

a) Pure deterministic "optimization" of travel path based on basic behavioural assumptions for (possibly very small segments of) homogeneous groups (for this approach to work it is necessary to be able to translate all aspects of traveller choice into the same unit, e.g. time or cost). As mentioned above, the availability of data restricts the level at which the segmentation could be applied in practice.

The optimisation could be done over more or less comprehensive choice sets. Under all circumstances the composite GC of the relevant set of travel paths has to be calculated. Obviously, most travellers are not in a position to identify more than a very limited number of feasible alternative travel paths. Most models that are used in practice apply simple criteria to delimit the amount of travel path alternatives that are considered feasible. Such criteria are normally based on the assumption that travellers (or groups of travellers forming a decision unit) behave as if they were economically rational and at this stage no allowance is normally made for the influence of random traveller preferences, even though random effects are incorporated at other stages. See further the description below of Sampers and Vips.

b) Optimisation as under a) above but including regard to the influence of heterogeneity that is not captured by segmentation

c) Satisficing feasible solutions

Typically, approach a) is used in assignment models. However, the pure approach a) is seldom considered plausible for comprehensive demand models, due to the wide range of factors that may influence travellers' decisions. Instead alternative b) is mostly preferred in these cases. Alternative c) has also been analysed but to our knowledge has hitherto not been tried in full scale transport demand modelling.

For approach b) there are a number of points that need to be made. Essentially, the aim is to retain the notion of optimisation (i.e. minimise generalise cost, or, equivalently, maximise utility), but to allow for the fact that some elements are unknown. As Andrew Daly [ref: **Notes on Public Transport Models**, 18 March 2010] has said:

The standard approach to unknown elements in modelling is either to ignore them, or to allow them to be represented by an appropriate distribution. Typically, rather simple assumptions are made, using a mean and a variance, for a standard mathematical distribution. As Daly (op.cit) also notes:

"sources of variation around the mean utility arise from

- the distribution of preferences in the population, including preferred trip timings and
- measurement error concerning, for example, the exact location of households;

"these sources of variation are almost invariably treated as homoskedastic for practical reasons, but this implies that non-linear functions may be required for the average utility"

In the current context, there are three particular sources of randomness which need to be considered. The first is randomness in **tastes**: this may be dealt with by means of random parameters in the utility or generalised cost functions, such models being generally referred to as random utility models (RUM). In the simplest case, the randomness is effectively confined to the mode constant (or, more generally, the alternative-specific constant – ASC), and with a particular assumption about the distribution, this yields the well-known logit model.

The next aspect of randomness, which assumes particular importance in the context of scheduled public transport services, relates to the particular timing preferences of individuals. In the literature this is mostly characterised by the "preferred arrival time" (PAT), but can (and probably should) be extended to considerations of preferred departure time as well (see Tseng Y-Y, Verhoef E T. (2008), Value of time by time of day: A stated-preference study, Transportation Research Part B 42 607–618). In the absence of any information, the distribution of preferred times is typically assumed to be uniform, at least over defined intervals of time.

The final aspect is related, but essentially on the supply side: this relates to the actual timing of individual services. With a full "timetable" approach, this should be known, but very often only the in-vehicle time and frequency/headway are available. This means that further assumptions need to be made about the actual spacing of the competing services.

Random utility models

Random utility models have been used extensively to deal with travel decisions. For computational reasons the logit model is the most frequently used among alternative RUM models. The RUM framework presupposes a decision structure with choices among a number of (potentially large) discrete travel alternatives that are mutually exclusive. RUM models are consistent with the assumption that travellers maximise utility, but allow for influences on travellers' decisions that are random to the modeller. This result is that the model predicts the probability by which a certain alternative is chosen, and not the choice. An estimate of the number of times an alternative is chosen is given by aggregating the choice probabilities over decision makers.

The logit model

For a thorough presentation of the logit model, the reader is referred to Ben-Akiva, Lerman (1985). Here, we will just briefly describe some properties of this model.
The simplest version is the standard multinomial logit model (MNL). A general assumption is that an alternative i is chosen if the utility of that alternative U_i is larger than the utility of any other alternative U_j, or
$P(i|U_i > U_j) = 1$ for all $j \neq i$
The RUM assumption is that the utilities can be described by observed parts V_i and V_j and unobserved parts ε_i and ε_j. Examples of unobserved utility include taste differences among decision makers, omitted explanatory variables and measurement errors. We can then formulate the probability of choosing a specific alternative i:
$P(i) = P(V_i + \epsilon_i > V_j + \epsilon_j)$, for all $j \neq i$
The specific MNL assumption is that ε is independently and identically (IID) Gumbel distributed for all alternatives. Then the familiar logit model equation can be derived:
$P(i) = \frac{e^{\mu V_i}}{\sum_{j \in C} e^{\mu V_j}}$ where μ is a scale parameter which is inversely proportional to the variance of ε. In practice, the scale parameter cannot be separately estimated from the parameters. The utility function is usually linear in parameters as shown below but the variables can be transformed in different ways.
$P(i) = \frac{e^{\beta \mathbf{x}_i}}{\sum_{j \in C} e^{\beta \mathbf{x}_j}}$ where x is a vector of variables in the utility function, β is a vector of parameters to be estimated and C is the choice set containing all the alternatives. The β includes the scale parameter.

Estimation and calibration

The parameters are estimated using the maximum likelihood method on a sample of decision makers, for which the choices and x variables have been observed. The method implies that the calculated β values are those that maximise the probability of the choices in the sample. The maximum likelihood estimator has statistical properties enabling statistical tests of individual parameters and different model specifications.

The estimation procedure will produce parameter estimates that will replicate the shares of the alternatives in the estimation sample to the extent that alternative- specific constants are used. This may not be sufficient to replicate known market shares in more detail after model implementation. Therefore calibration may be needed to make the model results consistent with observed market shares. Calibration can be done with respect to calibration targets in terms of the different alternatives or aggregates of the alternatives by adding constants but leaving the estimated parameters unchanged.

Elasticities

The direct elasticity, E_i, with respect to utility V_i or any component in V_i, is proportional to the level of V_i or any other component, and proportional to the scale. $E_i = -(1- P(i))\mu V_i$. It is also proportional to 1 – P(i) which implies higher elasticities for smaller market shares.

The cross elasticity E_{ij} with respect to utility V_j or any component in V_j, is uniform, i.e., the cross elasticity of the probability of alternative i with respect to a change of V_j are equal for all alternatives $i \neq j$. $E_{ij} = P(j)\mu V_j$. This means for example that if the generalised cost of rail is reduced, then the demand of the other modes air, bus and car is reduced by the same percentage.

Independence of Irrelevant Alternatives (IIA)

By dividing the probability of alternative i by the probability of alternative j we find that

$$\frac{P_i}{P_j} = e^{\mu(V_i - V_j)}$$

in which expression only the utilities of alternative i and j are included. This is called the independence of irrelevant alternatives (IIA) property. An advantage of this property is that additional alternatives can be easily included. In such a case, equal shares of all alternatives will be attracted to the new alternative. However, when added alternatives appear more similar to some of the existing alternatives than to others, this seems more like a disadvantage. A standard example of this is the so-called red-blue bus problem. This example demonstrates that in a car – bus choice model (where the utilities are the same and thus result in equal choice probabilities), an added identical blue bus alternative will attract as much demand from the car alternative as from the red bus alternative. Intuitively, one would expect more demand to come from the blue bus alternative.

Independent and Identically Distributed (IID) error terms

The root of the red-blue bus problem is the IID assumption on the random component ε. Because the red and blue bus alternatives are so close, they would be expected to share all unobserved utility (except for the colour). In that case the assumption that the random components are independently distributed is not valid. There will be a very high correlation between the random components of the red and blue bus alternatives. The most common way to deal with this problem in mode choice applications is to apply the so-called nested logit model. In this model alternatives that are perceived to be more similar are grouped in separate nests, each being a separate MNL model.

The nested logit model

This concept corresponds to a decomposition of the error terms into one part that is equal for the alternatives in the nest, and one part that is not. An example for long distance travel could be IC and X2000 trains, which are both train alternatives. For these two alternatives, unobserved variables like the service operator may be the same. In a nested model, an MNL model would be used for the choice between these two alternatives, and then a composite train alternative would be formulated in a MNL model for the choice between train and the other modes. The nested logit model offers a simple way of formulating such a composite alternative. The property of the Gumbel distribution implies that the log of the denominator of an MNL model (the so-called logsum) represents the expected maximum utility of the alternatives in the model (the IC and X2000 trains in our case). The nested logit model can be written

$$P(m) = \frac{e^{\beta x_m + \theta \ln(\sum_{s' \in S} e^{\beta x_{s'}})}}{\sum_{m' \in M} e^{\beta x_{m'} + \theta \ln(\sum_{s' \in S} e^{\beta x_{s'}})}} \quad P(s|m) = \frac{e^{\beta x_s}}{\sum_{s' \in S} e^{\beta x_{s'}}}$$

$$P(ms) = P(m) * P(s|m)$$

where m and m' are indices for main modes belonging to the choice set M (of which train is one in our example) and s and s' are indices for sub modes belonging to the choice set S (IC and X2000 in our example). β and x are vectors of parameters and explanatory variables for the different modes. The θ is the ratio of the scale parameters of the main mode model and the sub mode model.

In this model, the logsum from the sub mode model will in the main mode model represent the expected maximum utility from the sub mode alternatives. All the parameters β, γ, and θ can be simultaneously estimated using the maximum likelihood method. The nesting can be extended to several levels (one application could be nesting levels related to trip frequency, destination choice, mode choice and choice of departure time in separate nesting levels).

The logsum

The interpretation of the logsum to represent the composite (expected maximum) utility is useful also for economic assessment. The logsum can be rescaled into monetary units by dividing the logsum by the cost parameter. The value of a change in infrastructure to an individual can then be described by the change in the rescaled logsum.

The property of the logsum concept can be further illustrated by the following example:

Assume that originally there is only one alternative, 1, with the observed utility G^1. In this original situation the composite G according to the logsum is simply:

$$G = -\frac{1}{\mu}\ln(e^{-\mu G^1}) \equiv \frac{1}{\mu}\mu G^1 \equiv G^1$$

Assume now that we double the number of alternatives so that there are two alternatives with the same G. The new composite G* is then:

$$G^* = -\frac{1}{\mu}\ln(2e^{-\mu G^1}) \equiv -\frac{1}{\mu}\ln 2 - \frac{1}{\mu}\ln(e^{-\mu G^1}) \equiv G^1 - \frac{1}{\mu}\ln 2$$

The change of the composite G* is thus (1/μ)ln2. If we added k alternatives with the same G the new composite G* would increase by (1/μ)ln(k).

This example tells us that

- As there is unobserved utility, increasing the number of alternatives implies a better chance to find an alternative with a higher total utility. Even if an alternative with a lower utility were introduced, there would be a gain (but to a lesser extent).
- The gain from increasing the number of alternatives depends on the variance of the unobserved utilities – the larger the variance, the smaller the scale factor μ and the higher the increase of the expected maximum utility
- If there is no unobserved utility, then the scale parameter μ will approach infinity, and there will be no gain by increasing the number of such alternatives (as in the case of two identical alternatives)

- The increase does not depend on the observed utility G^1 – the variance is homoskedastic

Specification tests
It seems relevant to briefly summarise some of the different specification tests of Swedish long distance logit models that have been carried out as research extensions of practical estimation work. In one of these tests, the assumption of the identically distributed error term was tested by applying Heteroskedastic Extreme Value (HEV) logit models. This setting allows for identifying mode specific relative scale parameters. The test was based on data from the 1984/85 Swedish national travel survey, and found that the MNL assumption could be rejected.

In another specification test on the same data as was used for the first version of the Sampers private trip model, a nested structure including car and the two train modes IC and X2000 in the same nest was found to improve the model. The effect of this improvement was however small compared to the effects of nonlinear transformations of in vehicle time and cost variables. The nonlinear specification test was carried out on the data used for the current models. Nonlinear utility specifications have also been found important in national models for long distance travel for other countries.

It is not possible to “add up” the results of these specification tests, as different specification will influence error terms and hence their distributions. Further research needs to address model specification in a comprehensive way.

Other logit models
Formulating random utility models that are sufficiently tractable for practical application has required constraints on the model formulation corresponding to assumptions on the correlation between unobserved utilities. In the last decades, much research has been directed towards relaxing such constraints and to develop more behaviourally realistic models. The development of estimation procedures (like simulation techniques) has also contributed to this. Today, more elaborate models are at our disposal (such as the mixed logit model). It is however not within the scope of this paper to describe this development.

RDT
The RDT approach is based on two assumptions (in addition to the general assumption that the traveller chooses the alternative with the lowest generalised cost). Firstly it assumes that the “ideal” departure times are uniformly distributed for the period of time (a whole day, peak hours or non-peak hours for example) we are analysing. Secondly it assumes that the departure times of specific services are uniformly and independently distributed with a maximum equal to the corresponding headway. On this basis the “schedule delay” (difference between ideal departure time and actual departure time for any service) must also be uniformly distributed with a maximum equal to the corresponding headway.

Since departure times of all lines are assumed to be known, all lines **and stops** are considered simultaneously, but all may not be acceptable. Assume that different lines i have travel time R_i, Access/egress time A_i and headway H_i. The basis for the choice of acceptable lines is walk time to the stop A_i plus travel time after boarding, R_i, plus all of the headway, H_i. Assume that line 1 is best, with the lowest value $A_1 + R_1 + H_1$. Other lines m are acceptable if $A_m + R_m < A_1 + R_1 + H_1$. This means that it is not worthwhile to wait for a line for which the difference in access + ride time with reference to the best line is greater than the whole headway of the best line. Note that this definition of the attractive set is different from that used in the Sampers implementation of Emme/2, and corresponds to that cited in the WebTAG guidance in §3.2.

Below we provide a mathematical formulation of choice probabilities and composite schedule delay time in the case with only two travel alternatives. In this case, we ignore the access cost, assuming it to be the same for both services. Note that it is also possible to include the fares (which might be different for the two services) provided they can be converted to travel time minutes (using an appropriate Value of Time).
Because most of the discussion is related to time issues rather than monetary cost issues, we choose to use time units instead of monetary units in our generalised cost definition in this section as well as in subsequent sections. This implies no loss of generality.

Ride time and wait times have their specific values of time.

Different modes may have different values of time for ride time because of varying experienced comfort. These values may also differ between traveller segments. The value of time for one mode is chosen as the norm and is given the ride time weight 1. A more uncomfortable mode could be given the weight 1.2. I for example an IC-train is the norm with value of time equal to 70, it is given the weight 1. If a bus mode is more uncomfortable it could have the value of time equal to 84, which means that it has the weight 1.2. We denote the ride time weights α^1 and α^2 for the two modes respectively.

The wait time that is spent at home typically has a value of time below the ride time value since the opportunity cost is lower for waiting at home than for riding. Assume that the weight of wait time in relation to the norm weight of ride time is 0.5. This means that the value of time for waiting is 35 according to the example below. The norm weight of ride time over wait time is thus 2.0 in this example. We denote with β this relation norm ride time weight over wait time weight.

Notation
H^1 headway of line1
H^2 headway of line2
R^1 travel time (including price expressed in minutes) of line1.
α^1 weight of mode ride time in relation to the norm weight 1, for the mode where line 1 belongs
R^2 travel time (including price expressed in minutes) of line 2.
α^2 weight of mode ride time in relation to the norm weight 1, for the mode where line 2 belongs
x^1 time to departure (schedule delay) of line 1.
x^2 time to departure (schedule delay) of line 2.
β relation between the norm ride time weight and the weight of waiting time ("at home")

The joint (composite) generalised cost is then for a segment with individuals i:

$$G = E\left[\min\left[\alpha^1 R^1 + \beta x_i^1, \alpha^2 R^2 + \beta x_i^2\right]\right]$$

It has been shown, see Hasselström (1981) and Jansson, Lang and Mattsson (2008), that, given the assumed uniform distributions for x^1, x^2 over the respective ranges $[0,H^1]$ and $[0,H^2]$, the probability of choice of alternative 1, Pr(1), is:

$$\Pr(1) = \frac{1}{H^1H^2}\int_0^{H^1}\int_0^{H^2} h\left[\alpha^1R^2 - \alpha^2R^1 + \beta x^2 - \beta x^1\right]dx^2dx^1$$

where h(s) is the Heaviside step function defined by:

$$\mathrm{h(s)} = \begin{cases} 1 & \text{if } s > 0 \\ 0 & \text{if } s \le 0 \end{cases}$$

Note that the larger β is, i.e., the larger the weight of ride time is in relation to the weight of wait time, the larger will the probability be for lines with long headways.

This formula generalises to multiple competing lines. Note that the probability for choice of a specific line depends on travel times, prices and intervals of all acceptable lines. In fact this calculation is done for the set of all travel paths between origin and destination, including a number of combinations of lines and modes, but the formulae are for simplicity including lines only.

Hasselström (1981) and Jansson, Lang, Mattsson (2008) also show that the expected schedule delay, V, is:

$$V = \frac{1}{H^1H^2}\int_0^{H^1}\int_0^{H^2}\left(h\left[\alpha^1R^2 - \alpha^2R^1 + \beta x^2 - \beta x^1\right]\left(\beta x^1 - \beta x^2\right) + \beta x^2\right)dx^2dx^1$$

The average expected travel time when there are several acceptable lines is found by the weighted travel time for all lines where the weights are the calculated probabilities. If there are j acceptable lines and the travel time for line j is R^j and the probability of choice of line j is denoted Pr(j), the average expected travel time, R, is:

$$R = \sum_{j=1}^{k}\Pr(j)\mathrm{R}^j$$

The average composite generalised cost is simply the sum of the average composite schedule delay and the average expected travel time: G=V+R.

2.7.7 Approaches to meeting the requirement to generate data for calculation of consumer surplus and overall social benefits and cost

A common approach by transport economists is to evaluate the benefit of some measure that changes the travellers' GC by means of the change in consumer surplus (CS). In most cases this is calculated using the "rule of one half" (ROH) approximation. This means that the expected total benefit of the measure is calculated as the change in GC multiplied by the average of the demand before and after the measure is implemented. Of course this approach relies on a number of assumptions about the population of travellers reflected in the demand curve as well as using a correct estimate of the change of GC[5]. Therefore, in modelling long

[5] The trapezoid formulae (ROH) *at the level of the individual* is based on:
- A second order approximation of the indirect utility function.
- A first order approximation of the demand function.

distance travel it is important to check that the prerequisites for applying the rule of half are fulfilled before using the approximation in the normal fashion.

When demand is modelled by means of a logit model, changes of consumer surplus may be calculated using the so-called logsum. As explained above the logsum could be obtained as a standard output from such models and thus be used for calculations of change of CS of different transport policy measures.

It is important to observe that for consistency between assumptions underpinning demand projections and the calculation of Consumer Surplus the mechanism for calculation of CS should be compatible with the demand function: by mixing approaches there is a risk for confusion and error. However, political restrictions on the evaluation process of policy measures may make it necessary to depart from this compatibility requirement, which in most cases will lead to some inconsistency between the demand and CS calculation. See a discussion on consumer surplus calculations in appendix 2.

2.7.8 Issues in public transport assignment – Public Transport Network Algorithms

Public Transport network assignment is carried out according to two main approaches – schedule (timetable)-based and frequency-based. While timetable-based assignment may ultimately be the best way forward, neither of the current Swedish models adopts this approach, and hence the discussion in this section is confined to frequency-based assignment.

Like a highway network, a public transport network operates with a set of zones, nodes and links, with links being connected at nodes, and with nodes being connected to zone centroids. However, in addition, as a bare minimum, it is necessary to represent public transport lines and terminals/stations/stops, and the frequency between services on the same line. There may also be variations in fare, comfort, reliability between different services: however, although these may also influence choice, we will ignore them for the purpose of explaining the basic issue. We will also ignore the possible need to interchange between services.

In considering the range of services available for making the journey between two zones, the key issues are: location of terminals/stations/stops (and the time to access them), time spent on the vehicle ("ride time") and the service frequency. The traveller's choice between alternative services will also depend on whether they know the timetable, and the modeller's ability to allocate travellers among services will depend firstly on whether it is assumed that the travellers know the timetable, secondly on whether the modeller knows the timetable, and thirdly whether the modeller knows the distribution of desired times of travel!

In order to set out the possibilities, we begin by assuming that all travellers for a particular O-D pair must use the same start (and end) point: thus there is no choice of terminal/station/stop. If there is only one service, running at regular intervals (say, 30 minutes), then the only issue is "which instance" of the service (e.g. the 9.05 or the 9.35) will a particular traveller board.

- An assumption of constant marginal utility of income within the range of changes considered.

On these conditions - the trapezoid formulae is an estimate of the equivalent change in income.

- The approximations should be good provided that the price changes are sufficiently small and/or the curvatures of the functions involved are "small".
- As long as all consumers face the same price changes, the trapezoid formula applied to the aggregate demand function gives an estimate of the sum of equivalent income changes over all affected. If they don't face the same "price changes" we must operate with different segments and aggregate ΔCS over the segments at the end.

As noted, it is often assumed, in questions of "departure time choice", that the traveller has a preferred arrival time (PAT). For convenience we will assume this, though it might also be acceptable to have at least a preferred "window" for departure. Suppose the PAT is 10.00, and the travel time is 30 minutes. The services depart at 5 minutes past the half-hour (i.e. 8.35, 9.05, 9.35 etc).

Firstly, assume the traveller knows the timetable. Then he/she will choose either the last early arrival (i.e. depart at 9.05 and arrive 25 minutes early) or the first late arrival (i.e. depart at 9.35 and arrive 5 minutes late). Which is chosen will depend on the relative disutility of early and late arrival ("schedule delay"). If we know this, we can calculate the choice. Further, if we assume that all travellers have the same relative disutility of early and late arrival, then we can "translate" the distribution of PAT into a pattern of the choice of each "instance", and we can also calculate the average disutility, taking account of the early and late arrivals as well as the (constant) ride time.

Now suppose the traveller does not know the timetable, only the ride time and frequency. In this case, to arrive no later than PAT, he/she must arrive at the terminal/station/stop no later than 9.00. From the traveller's point of view, there is a uniform chance of the service arriving anytime between 0900 and 0930: on average, he/she will arrive 15 minutes early. Again depending on the relative disutility of early and late arrival, it would be preferable to arrive at the terminal/station/stop somewhat later, and accept some probability of being late. What is clear, however, is that without knowing exactly when the service is scheduled, the traveller will incur additional waiting time and, potentially, more disutility (schedule delay) from early or late arrival.

Note that in order to model this situation, the modeller needs to predict the time at which the traveller will arrive at the terminal/station/stop, **and** which "instance" he/she will catch. If the modeller does not know the actual timetable either, he is in no better position than the traveller.

We now introduce a second service between the same two terminals/stations/stops. Assume in the first place that the ride time is the same for both services, but the frequency is different. On this basis, the traveller will be indifferent between the services, and the allocation (and resulting schedule delay) will depend entirely on the particular timing of the instances of both services and the distribution of PAT. In the case where the timetable is known, travellers can be assumed to allocate themselves to particular "instances" so as to minimise schedule delay, as before.

When the timetable is not known, one could proceed in two ways. The traveller could operate on the **combined** frequency, and effectively treat it as if there was a single service, and on this basis decide when to arrive at the terminal/station/stop. This is likely to be an inefficient approach to some extent, since it implies that the instances are equally spaced, which they will certainly not be if the frequencies of the two services are different. An alternative is to assume that the services operate entirely independently, so that at any time the chance of service 1 arriving is $1/H_1$ and the chance of service 2 arriving is $1/H_2$, where H_i is the service headway, and to calculate the probability of the first service arriving in some given time interval. The second method seems likely to give a better estimate of average waiting time, and on this basis the decision can be taken when to arrive at the terminal/station/stop.

Note, however, that in practice whatever timing decision is made, the traveller will take the first service which arrives. The two methods (and other possible variants) assume more importance a) for the modeller, who has to allocate between the two services and find the average waiting time and schedule delay, and b) for the traveller when there is a choice of terminals/stations/stops and, without a knowledge of the timetable, it is necessary to decide in advance which to choose.

We now have to take into account the case where the two services have different characteristics. For simplicity, we confine this to the ride time (R_i). Once we allow the ride time to vary, it becomes clear that some services, or instances of services, might never be chosen (for a particular O-D journey). If an instance of one service departs earlier and arrives later than an instance of another service, the first will never be chosen if the timetable is known. There is a corresponding condition even if the timetable is not known: if the headway of the faster service is less than the difference between the ride times, it will be worth waiting for the faster service even if the slower service arrives first.

This leads to the notion of an "acceptable" or "attractive" set of lines: essentially, the aim is to rule out possibilities which are inferior. There are different ways in which this set can be defined: the following text is taken from the UK Department for Transport's guidance (WebTAG Unit 3.11.2):

> 7.3.2 The process is illustrated in Figure 3. The best path (excluding the origin wait time) is Path 1. On the other hand the path with the lowest maximum generalised cost (generalised cost excluding waiting time plus headway (headway=maximum waiting time)) is Path 2. Paths 2-4 all have a generalised cost (excluding the initial wait) that is less than the generalised cost of best path plus the headway of the best path; they are therefore considered acceptable paths, i.e. in some circumstances it would be better to take one of these paths rather than wait for the service on Path 1. Path 5 is not acceptable because the generalised cost is too high – it would *always* be preferable to wait for the service on Path 2.

Figure 3. Identification of acceptable paths (Method 1).

The normal interpretation of this is that travellers arrive randomly at terminals/stations/stops and will take the first service **within the attractive set** which arrives.

The same issues arise as before, both for the modeller, and for the traveller in evaluating alternative terminals/stations/stops – how to allocate to the different services, and hence how to calculate the average waiting time and schedule delay. In addition, because the ride times vary, it will be necessary to calculate an appropriate average ride time (typically flow-weighted over the services used).

Finally, we relax the condition that all services depart from the same terminal/station/stop. In this case, there is essentially a further choice, for both traveller and modeller, as to which terminal/station/stop to go to. Certainly if the timetable is not known, it will generally be sensible for the traveller to go to the terminal/station/stop which conveys the least expected generalised cost (inclusive of the access and egress costs): note, of course, that the calculation of this expectation involves the same problems about waiting time and schedule delay as discussed above.

From the modeller's point of view, if the distribution of generalised cost at any one terminal/station/stop was quite wide (in terms of its dispersion from the average), It might be appropriate to allocate to more than one terminal/station/stop, taking account of the difference in the means. In other words, if the expected generalised cost at terminal/station/stop 1 was only one minute less than that at terminal/station/stop 2, **and** the variation in generalised cost from the different services at each terminal/station/stop was substantially more than one minute, one might expect the allocation to terminals/stations/stops to be relatively even, rather than allocating all travellers to terminal/station/stop 1.

The concentration in this section has been on those aspects of public transport assignment which are directly relevant to the current methods. These are:

definition of attractive set of lines
assumption about traveller's knowledge of timetable
assumption about modeller's knowledge of timetable
allocation of travellers between competing services at the same terminal/station/stop
calculation of average waiting time and schedule delay
allocation of travellers between competing services at different terminals/stations/stops
calculation of average waiting time and schedule delay over all terminals/stations/stops

In the subsequent discussion these issues will be further discussed at some length.

Note that we have said very little about the actual distribution of preferred arrival, or departure, times. It is standard to assume that these are uniformly distributed, though this is done more for reasons of convenience than because of any appropriate evidence.

2.7.9 A schematic overview of issues in modelling long distance transport demand and alternative approaches to tackle these issues in modelling

In this section we try to summarize the discussion in the preceding sections – with considerable simplification – in schematic form. In the table below there is an overview of alternative approaches to deal with different issues in the modelling process.

Table 2.7.9 Alternative approaches to modelling issues: schematic provisional overview.

Issue	Sub-issue	Alternative or supplementary approaches (examples)
Modelled object	**Concept**	**Tour**
		Trip
Behavioural Assumptions	Travellers	Generalised Cost minimising
		Satisfying a set of restrictions
		Risk aversion or other decision strategies
	Suppliers	Exogenous
Representation of supply	Timetables	Rough representation by headways and ride time only
		Average exact timetable or “representative” timetable for a certain traffic period, e.g. weekdays
		Mix of headways and timetables using the latter for major lines
		Exact timetables including all seasonal and other variations
Traveller characteristics	Socio-demographic-economic	Segmentation
		Endogenous segmentation
		Separate models per category
	Valuations	Estimated mode specific constants to cater for separate valuation of modes
		Exogenously estimated values of time per mode equal for all segments
		ditto but with separate values per segment
		Continuous value distributions
	Preferred arrival/departure	Uniform distribution common to all travellers
		Uniform distribution per time period (time slice) and day
		Distribution profile over the day average for all segments
		Distribution profile (empirical) separate for segments per day
	Fares	Average fare per mode for each relation
		Differentiation of fares per line and mode and segment
		Full price list

	Travel terms and conditions	None
		Full set of terms applied to calculation of generalised price
	Modes	Main modes rail, air bus coach, car, other
		Within each main mode separate “sub-modes” with different characteristics
Choice set and choice behaviour		Full information about choice set from route planner applied to traveller segments.
		Structured “stepwise” choice with random influence
		Search of full set of feasible alternatives from network model and choice according to min GC
		ditto with random influence
Social benefit/Change of consumer surplus		Rule of one half (ROH)
		Change of overall average GC (RDT)
		Change of logsum
		Other demand model related measure

The table makes clear that the modeller’s task is multi-faceted and complex. Of course a lot has been learnt from research and earlier development, and the use of more or less comprehensive and sophisticated models for long distance travel. However, it is necessary to look carefully at the user requirements in order to concentrate on the right issues. Though development should be ambitious and consistent, it will always be necessary to make approximations and to strike balance between, on the one hand, detail and perfection in the representation of conditions and mechanisms behind supply and demand and, on the other hand, development cost, data cost and the cost and benefit for users of the model. The table draws attention to the basic needs to consider a range of approximation issues in further development work.

3 Brief description of Sampers and Vips and discussion of some properties of the models

3.1 Introduction and background

The current models could contribute to model improvements

One aim of the project is to investigate whether the existing implementations of the standard Sampers and Vips could contribute usefully to improving the model tools for long distance travel that are available to users. For this reason it is necessary to describe these models as well as to discuss their properties in relation to the general requirements and issues that have been raised in the preceding chapter. Hence in this chapter we will give a descriptive overview of the models and discuss in some detail how the current models Sampers and Vips have dealt with some of the important modelling issues brought forward in the preceding chapter. Hopefully this discussion will provide a basis for a judgement as to how far these models could contribute to viz. be consistent with the requirements of an “ideal” model

In fact, the basic algorithms in Sampers and Vips/Visum have to some extent the same theoretical background in the sense that each individual is assumed to choose the alternative with minimum generalized cost (GC). This GC is, however, not the same for each individual due to the variation of individual preferences – stochastic influence, and this stochastic influence reflects different aspects in the two models.

In Sampers there is a variation among individuals due to taste, measurement errors, omitted variables etc. In Vips/Visum there is an assumed uniform variation among individuals due to ideal departure or arrival time in relation to actual departure or arrival time. It is also assumed that departure times of alternative lines are uniformly distributed. Both the departure times of travellers and lines are thus assumed randomly (uniformly) distributed, which has given rise to the term RDT (Random Departure Times).

As noted earlier, the Sampers model uses three steps: i) the network model Emme/2 for assignment on lines within each mode and for estimation of travel time components for each mode, ii) a nested logit model for demand calculations concerning modes, destinations, and travel frequency, iii) Samkalk for calculation of consumer surplus, revenues, costs etc. for cost-benefit analysis. For each scenario Vips/Visum uses one step for simultaneous assignment on all combinations on lines and modes and for consumer surplus computation.

3.2 The Sampers model – principles and description

Long distance travel in the Sampers context

The Sampers system was designed to comprise all personal travel in Sweden, covering the range from local travel to international travel. To meet the requirement of enough spatial detail for local and regional travel and to avoid an excessive number of traffic zones (about 9000 for the whole of Sweden) in one application, Sweden had to be divided into 5 regions. For each of these five regions a separate application of the model for local and regional travel was defined. For long distance travel, less spatial detail is required. Therefore a single application was defined, comprising about 700 zones.

The subdivision into different sub models has implications for how congestion is handled, and for how public transport mode combinations are handled. It is obvious that long distance car travel in principle will affect road congestion, and also will be affected by congestion. The magnitude of this impact was considered not to be important enough to require iterative runs with regional models when forecasting long distance travel (which would be very time consuming). Instead, it was decided to use the long distance model without congestion feedback, but to allow for long distance traffic to affect congestion in the regional forecast by disaggregating the long distance car matrix to the regional spatial level. This requires the long distance model to be run prior to the model for local/regional trips.

Regarding public transport mode combinations, it is obvious that people using public transport for the whole journey will also use local and regional lines. For lines connecting to long distance mode terminals, the share of long distance travel may be significant. This can be true also for road links, where car is used as an access/egress mode. For this reason it was decided to model access/egress mode choice at the local/regional level. To generate the demand for access/egress trips, the long distance matrices were disaggregated to the regional level giving the demand to/from terminals and origins/destinations.

In the following sections, the long distance model will be described in more detail.

3.2.1 Basic principles of Sampers standard procedure for long distance travel

Originally, the model for long distance trips was designed to include the choice dimensions of trip frequency, destination, mode, departure time, and ticket type. In addition, a less comprehensive model excluding the departure time and ticket type choices was developed and implemented in the Sampers system. For various reasons (computation time, time to construct timetables for future scenarios), only the less comprehensive model (here called the standard model) has been maintained and used. It is the latter model that we have been studying in this project.

The purpose of the Sampers long distance model is to generate demand for all long distance trips in Sweden, to allocate them to main mode and destination, and then (as described above) to load them on the regional road network and the long distance and regional public transport networks. To fulfil this purpose, a procedure to describe travellers' behaviour with respect to trip frequency, destination choice, mode choice and line choice had to be developed. This has been done using the nested logit model framework for generation and distribution on destinations and modes, and using the so-called Optimal Strategy (OS) approach (see 3.2.2 below) in Emme/2 for line choice within public transport modes.

As said above, long distance trips may involve combinations of different modes. A long distance public transport trip normally includes an "access mode choice" for travel to a train station or an airport, before the trip on a "main mode" (i.e. the mode that will be used for most of the trip) can take place. After embarking the main mode, an "egress mode choice" is also normally needed. If it is assumed that the main mode choice is more important in terms of time and cost, it may be defendable to simplify the choice model to explicitly include only the main modes. The basic principle of the standard Sampers model can be described as a "main mode" discrete choice approach, comprising car, long-distance bus, train and air modes as main modes. A requirement for this approach is that combinations of main modes are rare. Whether such an assumption is defensible or not is discussed further in chapter 4 below.

In Sampers the influence of access/egress modes on main mode” choices is treated in a simplified way, by means of a distance variable. This means that changes in the public transport supply for access/egress will not influence the main mode choice. It is perfectly possible also to include access/egress mode choices in a main mode logit model (this had been done in a previous model for Swedish long distance travel), but this was not done in Sampers because of time constraints. Instead, an access/aggress logit mode choice model is applied at the local/regional level without feedback to the main mode choice. The access/egress mode choice model includes the following alternatives (different for access and egress):

Alternative	Origin, Private	Destination Private	Origin, Business	Destination Business
Car parked at station/airport	x		x	
Kiss & Ride	x	x	x	x
Taxi and hired car	x	x	x	x
Bus	x	x	x	x
Train	x	x		
Walk and Bicycle	x	x	x	x

The access/egress models concern trip using the main modes train and air only. For bus trips, the terminals are not well known, and to a large extent bus trips do not have access/egress trips, specifically in the destination part of the long distance journey.

The procedure for applying the Sampers standard main mode model can be described as first generating the main mode time and cost components, then calculating the demand – first for each main mode, then for specific travel paths within each mode. Finally the cost benefit calculation is performed in the Samkalk module. These steps will be explained in more detail below.

3.2.2 Step 1: The optimal strategy approach in Emme/2 in Sampers

The resulting travel time components used in the context of the national model system are based on travel time and frequencies of services which have been included in the travellers’ choice set, separately for each (main) mode. These frequencies are calculated from real timetables or generated in scenario descriptions. The travellers are assumed to know the travel time components and headway of all lines, but not the timetable (the actual departure times) or behave as if they do not. The basic behavioural assumption is that travellers choose the alternative with minimum weighted travel **time** components, i.e., generalised cost except price (fare).

This is done because price is assumed to be equal for all alternatives in the (modal) choice set. The price variable is specified as an exogenous O-D matrix included in the implementation of the multinomial logit models; this way of handling the price variable has the effect that possible price differentials between different services within one mode are not taken into consideration.

Because of the assumption that travellers do not know the timetables, they need to select the main mode boarding node in advance, and it is assumed that they select the node that represents the shortest expected total travel time, as explained below. The distribution over different services at the boarding node will be proportional only to the frequency of the services included in the traveller’s choice set.

The set of acceptable lines of a mode at a stop is selected as follows. Assume that different lines j have total travel times R^j and headway H^j

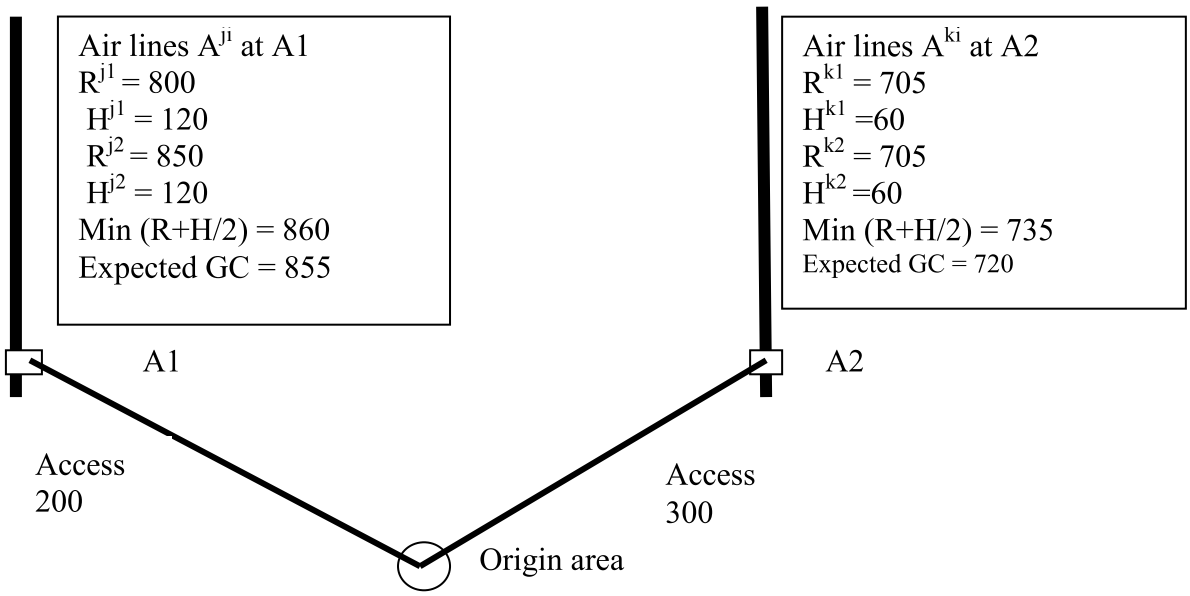
The efficient (or attractive) set of lines will depend on the waiting time weight used in assignment, and the efficient set is defined by $R(j) < R+w*Wait$ where R = mean time for lines already in the efficient set, w=weight on waiting time and Wait= combined waiting time of lines already in the efficient set. Lines are added to the attractive set as long as the inequality is satisfied for lines sorted in increasing order of R.

Note that in Emme/2 new lines are incrementally updated. The acceptance mechanism could be regarded as a way to take into consideration that travellers do not compare all alternatives in detail. They are looking at the shortest travel times first. The mechanism may cause that a third or a fourth line will not be included in the acceptable set even if it would be if all lines were regarded at the same time.

The best stop is the one giving the lowest expected generalised cost (ignoring fare but including access/egress cost), i.e. giving the "optimal strategy" for the trip. The line choice issue is of course less sensitive for the within mode case (as in the main mode approach) than when all modes are included, which makes the consequences of simplifications less severe.

The access/egress parts are represented only by distance and a speed factor in the main mode line choice model. Due to long headways, it may be that the access/egress mode in some cases will appear to be the best to use all the way, or for an unrealistically long part of the trip. Therefore the speed factor was set to a very low value, to prevent such effects. This also influences the selection of the best stop. The effects will be present also in the egress part with the same consequences.

The assignment procedure is illustrated in the following figure for air lines at two alternative airports A1 and A2. In the figure below H stands for headway and R for all travel time components after boarding plus access and egress links but fare is not considered.



In this case all travellers will choose the strategy to go to airport A2 with the minimum expected generalised cost equal to 720, and to choose one of the two similar air lines in proportion to their frequencies. Given its importance, the calculations are explained in detail below.

At A1 minimum R is R^{j1}. We add service 2 to the attractive set if $R^{j2} < R^{j1} + w*(\frac{1}{2} H^{j1})$. By assumption in this case w = 1: this gives 850 < 800 + 60, so service 2 is included. Hence combined headway = 60, and travellers are allocated equally to both services (since headways equal). Hence average R = 825, and average wait = w*30, giving average GC = 855.

At A2 both services are equal so both are included in the attractive set. Hence combined headway = 30, and travellers again allocated equally to both services (since headways equal). Hence average R = 705, and average wait = w*15, giving average GC = 720.
We now add in the access costs, which gives 855 + 200 = 1055 for A1 and 720 + 300 = 1020 for A2. Hence, on an all-or-nothing basis, A2 is chosen

However, if access/egress cost for A1 were shortened by more than 35 units, all demand would be directed towards airport A1. Such deterministic behaviour is not very realistic. Nonetheless, the main mode approach limits this problem to the choice of stop/terminal **within** each mode, and not across the modes.

3.2.3 Step 2: The logit model in Sampers

Sampers applies a standard nested logit model for the choice of trip frequency, destination zone and main mode, using travel time components of each mode generated by the Emme/2 Optimal Strategy approach. The price of each main mode (bus, train, air and car) is specified exogenously for each O-D pair and included in the utility function for each mode. The models are described in the technical documentation of the Sampers model. Here some important aspects of the models will be briefly discussed.

Decision unit

Long distance travel is often undertaken in groups rather than individually. (According to the national travel survey, RES05/06). It is important to consider this aspect in the modelling of long distance travel. The fact that travel in groups is important has been accommodated in Sampers by segmenting the application according to a distribution of the size of the travelling party, and adjusting the cost of travel by car with respect to the party size. Another consequence is that individual-specific socio-economic variables (and segmentation) are avoided, as the party composition is not fully known in the travel survey. An exception is business trips, in which gender, educational variables and individual income is used.

Dependent variable

It is assumed that long distance travel choices are planned on a tour basis, i.e. not only the outbound trip leg but also the homebound trip leg. Intermediate trip legs do also occur, but the occurrence of this in the data was judged to be rare enough to allow for a simplification of long distance trips to a two-legged tour. For long distance travel, it may be that slower modes require additional overnight stay when both trip legs are considered. This may impact on travellers' preferences, depending on the duration of their destination activity. While it would also be possible that different modes are chosen for the two trip legs, this was shown to be very rare in the data, and therefore the use of different modes on each trip leg is not allowed for in the model.

Regard to timetable

In reality it is likely that most travellers make use of timetables to acquire information about different modes, not only to know the travel time and service frequency but also to be able to take schedule delay into account. In this model, explicit potential schedule delay effects are neglected, and only travel time and headway are used. It is however obvious that the more frequent a mode, the lesser the expected schedule delay. Therefore, to some extent, the headway preference takes account of schedule delay. It is often found that, per minute headway, the disutility of headway is a decreasing function of headway. In the Sampers application, the results from the 1994 Swedish value of time study have been used as a transformation of the headway variable (resulting in disutility being a piecewise linear decreasing function of headway).

Segmentation

Segmentation can be undertaken for two reasons. One is to make the implementation as true as possible to the model, which implies segmentation by the decision maker. That kind of segmentation is often treated by sample simulation, implying calculation of the choice probabilities for each individual in the sample. The sample is re-weighted to fit the distribution over categories for which there is available information at the zone level of the implementation. A more aggregate approach is to apply the model to each of a number of categories, which are assumed to be internally homogenous with respect to the model. This is the case for the Sampers system. For efficiency reasons, the long distance model is segmented on the following categories:

Segmentation variables
Trip purpose
Destination activity duration
Car ownership
Party size
Income group
Gender

The other reason for segmentation is to produce results for specific socioeconomic groups, even if they do not differ with respect to calculated probabilities for a specific origin zone. In addition to this segmentation, the procedure is therefore segmented on a number of individual socio economic variables such as gender, age and occupation,

Estimation

The model has been estimated on data from the national travel survey from the period 1994 – 2000. About 15000 observations were used for private trips, and about 2000 observations for business trips.

Calibration

The model has been calibrated by adding constants to the utility functions to replicate information on air trips from the aviation authority and on train trips from the State Railway Company. For bus and car, the calibration target was the number of trips in the travel survey.

3.2.4 Step 3: Samkalk

For CBA analysis, a deliberate decision was taken not to use the preferences inherent in the demand models to assess the consumer surplus effect. This decision was taken for political reasons, and also made it impossible to use the logsum concept. Instead the "rule of a half" is employed. Consumer surplus change is then based on the change of composite generalised cost *over all lines used.* For existing/remaining travellers the change in consumer surplus is the change in composite generalised cost (G) multiplied by the number of existing/remaining travellers. For new or lost travellers on this mode the change in consumer surplus is the change in composite generalised cost (G) multiplied by the number of new/lost travellers, divided by 2.

3.3 The Vips long distance travel application– principles and description

3.3.1 Basic principles

The network models Vips and Visum work in *one simultaneous step for lines and modes,* where all feasible combinations of lines and modes are regarded, and where each line, as well as the car alternative, has a specific ride time and price (kilometre, zone, or stop-stop based). Both line-specific ride times and prices will thus affect the travellers' choice.

3.3.2 The RDT algorithm

The RDT-algorithm is a central element in Vips and Visum. The implemented method used in Vips is found in Jansson and Ridderstolpe (1992).With the RDT-principle in Vips/Visum all feasible lines and modes give rise to a composite generalized cost and the change in generalized cost takes into consideration measures assumed for any mode.

The RDT algorithm is based on two assumptions. In its implemented form it ignores the stochastic element that varies with taste differences among individuals, measurement errors etc. It uses the stochastic element x, the difference between actual and ideal departure time, which may be regarded as a variant of schedule delay Travellers are assumed to know the timetables and thus the schedule delay resulting from each possibility, which in this case merely relates to the inconvenience of not being able to depart precisely at their ideal time.

From the modeller's point of view the calculation of schedule delay is equal to the expected delay based on average frequencies of services and not on exact departure times, which are assumed unknown (NB to the modeller) and therefore assumed to be uniformly distributed. In the current RDT-implementation the schedule delay is only related to the difference between ideal and actual departure time. It could be argued, however, that both early and late arrival times (at the destination) should be taken into account. The algorithm was described above in section 2.7.6.

We illustrate the calculation principle by the use of two diagrams.

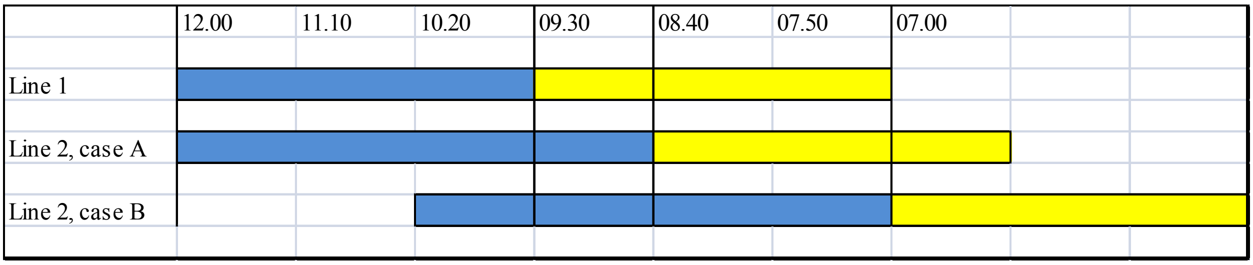
In diagram 3.3.2.1 below the blue parts denote ride time and the yellow parts headway. Line 1 has ride time 150 minutes and line 2 has ride time 200 minutes. Both lines have headway 150 minutes. We assume that line 1 departs at 09.30 (since the interval is 150 minutes the previous departure is at 07.00). For line 2 we illustrate two extreme cases for departure times, phase A and phase. The travel time with line 1 is 150 minutes so a journey with line 1 ends at 12.00. We assume travellers' ideal departure times are uniformly distributed, and consider the range 07.00 to 09.30.

In case A for line 2, we assume line 2 departs at 08.40 and arrives at the same time as line 1. Hence all who want to depart between 08.40 and 09.30 will choose line 1 since line 2 cannot reach the destination earlier. This means that for 50/150 of the travellers choice of line 1 is certainly the best. Travellers who want to depart between 07.00 and 08.40 would be assumed indifferent, since the sum of schedule delay and ride time is the same. However, depending on the assumptions about the departure for line 2, for ideal departure times before 08.40 the probability of line 1 to be chosen is successively decreasing from probability 1.0

What is the lowest probability for the choice of line 1? If line 2 has a departure time relative to line 1 so that it arrives at the same time as line 1, then line 2 may be the best choice (called case A in the diagram 3.3.2.1). If line 2 departs relative to line 1 so that it arrives just before 10.20, then line 2 may also be the best choice (called case B in the diagram 3.3.2.1). The difference between these two relative departure times is 100 minutes. Since line 2 has headway 150 minutes it is the best choice with the probability 100/150. Per unit shift in line 2 departure time, measured in minutes, the probability is 100/(150x150) = 100/22 500. The minimum probability for choice of line 1 for ideal departure times between 07.00 and 08.40 is then 1-100/150=50/150, and the probability per unit shift in line 2 departure time (measured in minutes) is 50/22 500.

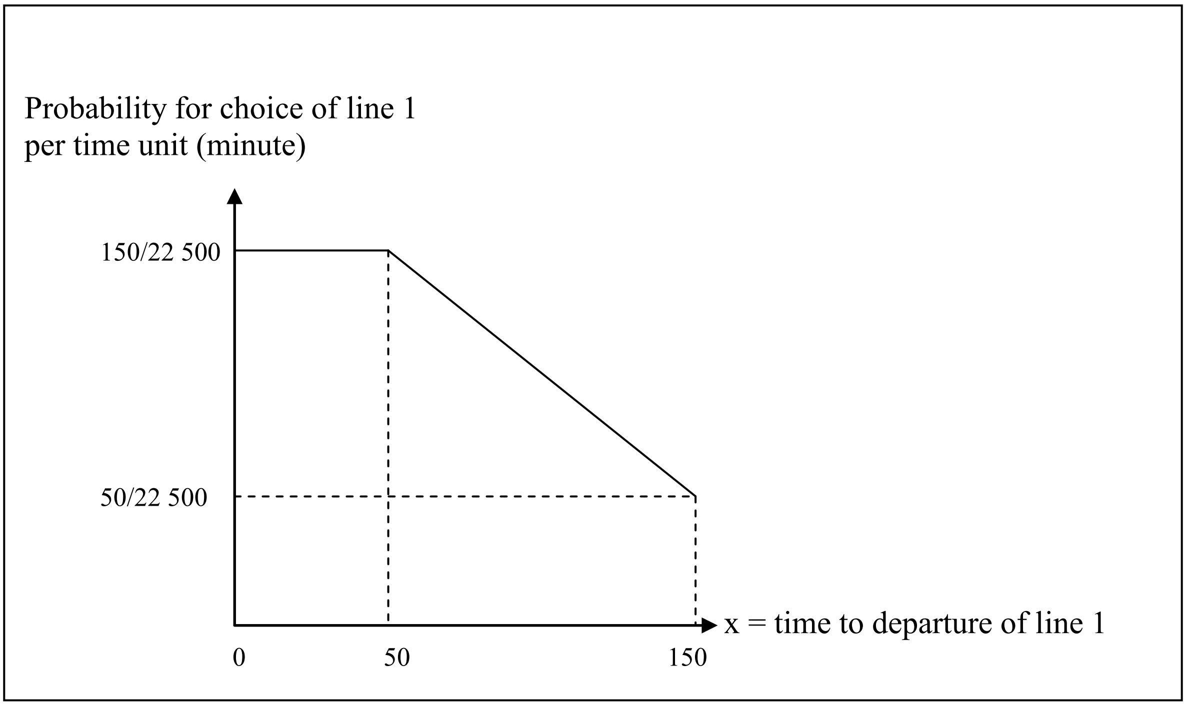
For ideal departure times between 07.00 and 08.40 the probability for choice of line 1 decreases from 1.0 to 50/150, or from 10/150 to 50/22 500 per minute. The probability for choice of line 1 when time to departure is below 50 minutes is 10/150 per minute. Below10/150 is also expressed as 150/22 500.

Diagram 3.3.2.1 Travel time and headway for line 1 and 2 for various relative departure times



We will now illustrate the varying probability for choice of line 1 in a diagram that explains the integral calculus.

Diagram 3.3.2.2 Probability for choice of line 1



The horizontal line and the slanting line correspond to the function (integrand):

$$1 - \frac{x + R_u - R_i}{H_i}$$

The horizontal line shows the probability per minute for choice of line 1 when the difference between ideal and actual departure times is below 50 minutes.

The slanting line shows that the probability decreases the longer the time between ideal and actual departure times for line 1.

The dotted horizontal line shows the minimum probability per minute for choice of line 1.

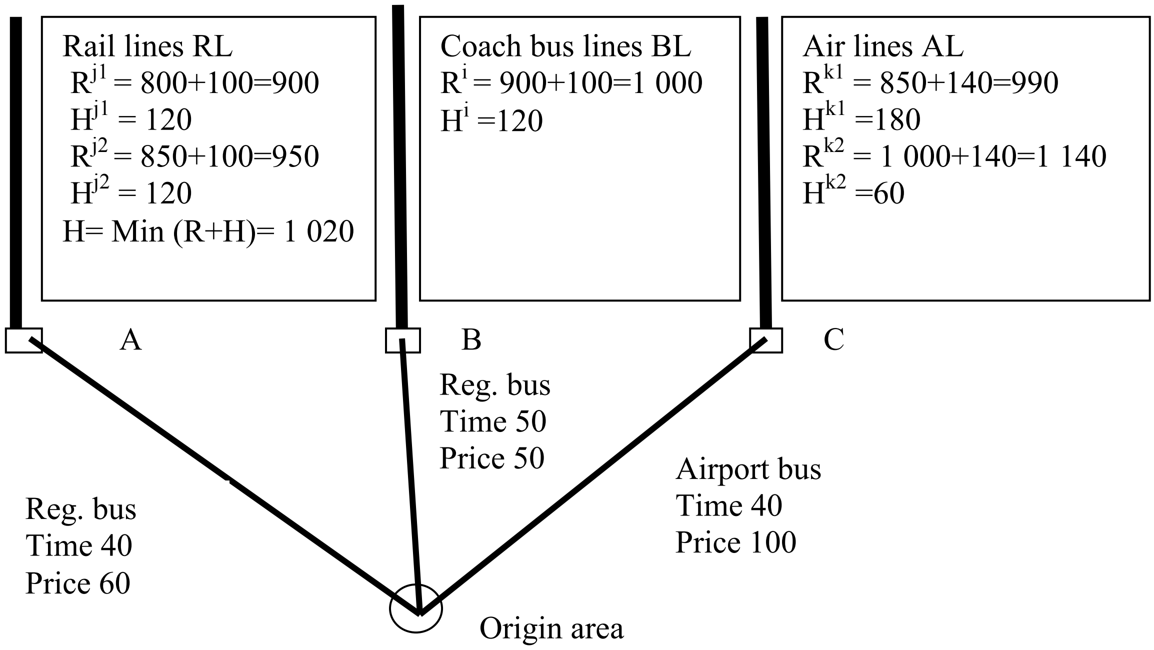
The average probability for choice of line 1 can thus be calculated as the area below the horizontal and the slanting lines (the integral over the function with integrand x):

$$\Pr(1) = 50 \times \frac{150}{22500} + \frac{100 \times 100}{22500} \times \frac{1}{2} + \frac{100 \times 50}{22500} = \frac{15000 + 10000 + 10000}{45000} = \frac{35000}{45000} = \frac{7}{9} \approx 0.78$$

This means that the "average probability for the choice of line 1 integrated over all possible departure times for line 2 in the range (0,150)".

3.3.3 Choice of lines according to Vips

The figure below illustrates the choices according to Vips. Rail lines depart from terminal A, coach bus lines from B and air lines from C. There are two rail lines from R but one coach bus line from B and two air lines from C, where one of them departs more often but is more expensive. Regional bus times and price are assumed to be available from origin area to each of these terminals. In the figure below H stands for headway and R for all travel time components after boarding, plus access to the boarding stop in terms of walk or car, plus the fare. Access costs for RL, BL, AL are 100, 100, and 140. Price is expressed in minutes.



The best line apparently is the rail line RL^{j1}, with the shortest ride time plus headway, equal to 1 020. In this case all the other lines are acceptable except air line AL^{k2} since R^{k2} =1 140 > 1 020.

The best line is RL^{j1} with 900+120 = 1020. Now take each other line and see whether R alone is less than this. This applies in all cases except AL^{k2} where R = 1140. If R for AL^{k1} goes up by more than (1020-990) time units then this "air line" will no longer be in the acceptable set and no one will go by air.

In the logit model there would still be some people going by air. This is the same kind of deterministic behaviour that exists in the optimal strategy approach and which is due to the definition of the choice set. In the RDT context, however, this is likely to happen to lines with lower real choice probabilities as the acceptance criterion is more "generous". On the other hand, this may happen to entire modes (as in the above example), which is not the case with the Sampers approach in which all modes always will have a choice probability.

The acceptable set criterion has some similarity with that of the optimal strategy approach in Emme/2, since the criterion is $R^i < R^{best} + H^{best}/2$ in Emme/2 but $R^i < R^{best} + H^{best}$ in Vips and Visum.

Just like in emme/2 new lines are incrementally updated. Vips fist picks the line with shortest $R^{best} + H^{best}$. Then the line with the shortest travel time (including price) is compared with the best one. If this one is acceptable according to the criterion it belongs to the choice set. Then the two lines are merged and regarded as one alternative with headway equal to the double expected wait time and the average ride time, weighted with probabilities. Then the next candidate line with the shortest travel time of the remaining ones is examined in the same manner and so on. This acceptance mechanism may be regarded as a way to take into consideration that travellers do not compare all alternatives in detail; they are looking at the shortest travel times and prices first. However, the mechanism may in some cases cause that a third or a fourth line will not be included in the acceptable set even if these would be accepted if the travellers actually compare all alternatives at the same time.

That the whole interval is included in Vips and Visum depends on the fact that the travellers are assumed not to accept late arrivals. The schedule delay in Vips is only related to early arrival, i.e., that one must not be too late. However, the possibility of choice between early and late arrival should be seriously considered. With allowance for both early and late arrival/departure, the attractive set of lines will approach
time plus half the headway of the best line. It is just a matter of choosing the proper weight for headway in Vips assignments, but the proper weight is a matter of estimation/empirical evidence.

Assuming that the timetables are known also means that all stops, modes and lines are considered by the travellers, so that more than one stop and mode may be acceptable. The criterion may, however, mean that some modes are not acceptable at all (for particular O-D movements). With the Sampers approach all modes will always have a choice probability due to the random terms that are not available in Vips and Visum.

Based on the assumption of Vips/Visum the travellers are not allocated to travel paths in proportion to frequency. Given the preferred departure or arrival time each individual chooses one alternative, that with the minimum generalised cost, which depends on how the actual departure or arrival times relate to the ideal departure or arrival times for this individual. The best alternative differs between individuals since the schedule delay differs between the individuals, reflected in the stochastic element in RDT, i.e. the differing ideal times and the departure times of the alternatives.

This is an expected schedule delay based on average frequencies of services and not on exact departure times. The RDT approach assumes that travellers´ ideal departure or arrival times are uniformly distributed. It also assumes that departure times of alternative lines are uniformly distributed. The mathematical derivations of choice probabilities, composite schedule delay and composite generalised cost have been given above.

3.3.4 Calculation of change of consumer surplus in Vips

The calculation of CS in Vips is simply:

$$CS = \sum_{OD} T_{OD} . (G_{OD} - G'_{OD})$$

where G is the composite generalised cost before the measures and G′ after, with T being the (unchanged) travel matrix.

3.3.5 Estimation and calibration

In Vips parameters are not estimated by use of any statistical method. Basically the values of time for the various travel time components are used, originating from ASEK 4. These values are not changed. But in the calibration process other parameters may be changed. This calibration process may also be seen as a "manual" estimation process, since one is searching by trial and error the parameters that give demands for lines and modes as close as possible to empirically based demand.

Basically weights for public transport modes have been specified without further changes. Various mode constants were tested, which did not affect the demand pattern very much.

Vips has been calibrated against:

- The Swedish airport administrations (LFV) statistics on number of travellers between airports,
- The Swedish railways (SJ) statistics on number of travellers on specific links,
- Number of passenger km per traveller segment according to the national survey (RES).

The parameters for calibration of Vips have been:

Prices for rail lines and air lines. These prices are originally taken for the various traveller segments (employed, pensioners, students, and business travellers) from the web home pages of the operators. These prices are, however, highly differentiated since prices shift between times of day and from day to day. It is therefore difficult to calculate a "true" average price. For this reason price changes by some plus or minus 10 per cent were tested. It appeared that variation of prices had a significant effect on demand for lines and modes.

Prices for car use. The price was assumed to include fuel price plus an additional 30 per cent more or less. It is genuinely difficult to know how people believe the true prices to be. Some people probably only take fuel price into account, other people include all variable costs. Various prices were tested for the different segments. It appeared that the model is very sensitive for assumptions on car use prices.

Comfort of car. Travellers may perceive car as more or less comfortable than public transport modes. The ride time weight of car in relation to the norm 1 (ordinary intercity train) that was found most suitable was above 1, i.e. that travellers in general seem to regard car as less comfortable than train. It appeared also that the model is very sensitive for assumptions on these comfort weights.

3.4 Summary table with description of currently implemented versions of Sampers and Vips

In the following tables we give an overview summary over how various aspects of the travel demand model have been implemented in the current versions of Standard Sampers and Vips. To facilitate comparison the characteristics of the systems are presented side by side.

Table 2 a Dependent variable

Sampers	**Vips**
Tour (Restricted to return trips only)	Trips

Network representation

Sampers	**Vips**
<u>Lines for each mode</u> Only "main modes" between each O-D pair; Any main mode can only be combined with other modes when the latter are used for access/egress parts of tours; access/egress parts of tours are modelled in a mode choice model including local lines on a more detailed level (regional network) The average (weighted by frequency which in this case is equivalent to market share) ride time and wait time for all feasible lines of each main mode in each O-D pair is used to estimate/calculate mode shares in Sampers demand model (logit)	<u>Lines for each mode</u> Combinations of all feasible lines and modes between all O-D pairs Individual ride times for all feasible lines of each mode in each O-D pair.

Table 2 b Parameters

Sampers (data from Emme/2 for assignment within main modes)	**Vips**
Mode specific constants and parameters for different time components. Different parameters w r t activity duration Same weight for wait time and transfer time in Emme/2 assignments. In Sampers headway has a stepwise weight, which affects mode and destination choice Separate relative values of different time components in Emme/2 assignments and in Sampers logit models => inconsistency	Mode specific constants. Separate weights for wait time and transfer time. Separate weights for wait "at home" (plus a margin) and at stop. These affect both assignment and level of service. No regard to diminishing disutility of headway Convenience parameters per mode (specific ride time weights) and per stop (specific wait time weights) Parameters for wait time and ride time delay at each stop on each line, which affect both assignment and level of service. Consistent values of time since there is one simultaneous calculation step.

Table 2 c Fares

Sampers	**Vips**
No fares in Emme/2 One average fare for each main mode in each O-D pair, for private and business journeys respectively is used. These are coded exogenously as a price matrix for each main mode.	Individual prices for many lines of a mode. Prices differ w r t: a) direct trips, b) transfer trips, c) operators. Distance based fares are coded as progressive or regressive. Stop-stop fares are applied for airlines. The price may then differ between direct air transport and with transfers respectively. The price from A to C may differ from the addition of the price from A to B and B to C. Free transfers within an operator or within a group of lines are specified. The composite fare for a journey for each O-D pair is endogenously calculated as the weighted average of the traveller traffic shares of feasible combinations of lines and modes.

Table 2 d Algorithms

Emme/2 Assumes that travellers have no knowledge of timetables, only of travel times and headways. Travellers choose the stop (for the mode given by the choice model) with minimum generalized cost. At this stop assignment on (acceptable) lines is in proportion to frequency, with no further concern for travel times and price. Wait time is calculated as half the combined headways. **Sampers** Takes from Emme/2 average travel times for each mode plus the exogenous price matrix and assigns on modes by use of the logit model. Calculates the logsum for each scenario of network, which is used for travel matrix generation, but not for consumer surplus calculation. Sampers pays attention to variation w r t taste and measurement errors but not w r t RDT.	**Vips** Assumes that travellers have knowledge of [see earlier remarks!] timetables, travel times and headways. Travellers are assigned to all acceptable lines regardless of stops and modes. Hence allocation depends on headways, travel times and prices of all lines and modes. Assumes uniform distribution of ideal departure or arrival times and uniform distribution for departures of lines. These two features are called RDT (Random Departure Times). For certain stops specific matched transfer time between for example regional buses and trains are specified. Shares for lines and modes and wait times are thus calculated by use of integration over all alternatives (lines and modes) with their respective travel times, prices and headways. Vips pays attention to variation w r t RDT but not w r t taste and measurement errors.

Table 2 e Consumer surplus (CS)

Samkalk The module Samkalk calculates CS by regarding all lines used; Samkalk then distinguishes between existing/remaining travellers and new/lost travellers, where the latter gain/lose half of the existing/remaining. The CS calculated by Samkalk includes regard to destination and trip frequency effects of transport policy measures	**Vips** CS is calculated for all lines and modes simultaneously. One can thus calculate CS for changes of one or several lines and modes at the same time

Table 2 f Segmentation on traveller categories

Sampers	**Vips**
Sampers is segmented w r t calculation of choice probabilities on the following variables: Trip purpose Destination activity duration Party size Income group Gender Sampers is also segmented into the following categories for equity considerations: Age Occupation	Domestic long distance travel is segmented on private and business. Private travel is further segmented on: Working Pensioners Students and for each of these categories specification on available car, no car available Business trips are segmented on: available car no car available For each segment, separate value of time components are given for each mode

Table 2 g Estimation and Calibration

Sampers	**Vips**
Sampers is estimated by maximum likelihood based on the national travel survey. It is calibrated to match statistic from the Aviation authority for air, the State Railway company for train and the travel survey for car and bus.	Parameters, e.g. time values, which are used in the current Vips implementation, are entirely based on exogenous estimates from the official ASEK 4. *Vips has been calibrated against:* • The Swedish airport administrations (LFV) statistics on number of travellers between airports, • The Swedish railways (SJ) statistics on number of travellers on specific links, • Number of passenger km per segment according to the national survey (RES). Calibration has been done using the following system parameters: • Prices for rail and air lines • Prices for car use • Comfort of car No formal statistical procedure has been used or is available for calibration

3.5 Notes on differences between some model functions in Sampers/Samkalk and Vips based on model descriptions and observations from model runs

3.5.1 Introduction

In the preceding sections in this chapter we have tried to describe the principles, algorithms, and the ways by which various model elements have been implemented. We have also tried to highlight the differences by means of a tabular presentation where the two model implementations have been put side by side.

In this section (3.5) we will make additional and somewhat more detailed comparisons of how specific model components work in Sampers/Samkalk and Vips in order to further highlight differences and similarities. We think that a detailed understanding of the key properties of each of the model implementation is of decisive importance in order to understand how the properties of each of the implemented models could contribute to the further development of models for long distance travel in Sweden and elsewhere.

3.5.2 Observations relating to the algorithms for travel distribution over different services (lines)

The general design and implementation of the OS algorithm and the RDT-algorithm is given above.

Similar or identical approaches in the two model systems:

- Both models define a set of acceptable (or feasible) lines
- Both models assume that the travellers minimize their GC
- In both models travel time and headway are given for each line
- For OS/Emme/2 only travel time and headway are given
- Spacing between departures for lines is not given for either of the two model systems. For both applications modellers assume equal spacing between departures of lines.

Different approaches in the two model systems:

For RDT fare is represented for each line, while for OS/Emme/2 only average fare per mode is represented.

The RDT algorithm is based on the assumption that travellers behave as if they know the timetable that is, the exact departure and arrival times. Acceptable alternatives, at all stops, are such that their travel times are less than the travel time plus the *whole of the headway* for the best combined alternative (see above section 3.3.3).

The OS-algorithm does not explicitly consider the fact that long distance travellers could be expected to use timetable information and therefore assigns travellers to lines at the perceived best stop in proportion to frequency. Acceptable alternatives are such that their travel times are less than the travel time plus *half the headway* for the best combined alternative (see above section 3.2.2).

In Sampers it is assumed that departures between all acceptable lines within a mode are equally spaced. In Vips it is assumed that the offset between departures for all lines within a mode has a uniform distribution.

It is important to distinguish clearly between behavioural assumptions about the traveller and the way by which the modeller tries to approximate traveller behaviour in the model.

Wait time and transfer time are calculated in the same manner in Vips. For transfers there are cases where there is a specified matched transfer coded between certain lines in certain directions. This is made in order to take into consideration that there are planned short transfer times, between for example rail lines or between regional bus lines and rail lines. It could be argued, however, that it is the longest headway of the lines in the travel path that matters for schedule delay associated with a path.

3.5.3 Access to stops and modes

In Vips' implementation for long distance travel all lines and modes are coded except local bus lines in cities. From each origin travellers may have several regional buses or coach lines to choose among to reach a rail or air service. They may in addition have several rail services to choose among to reach an airport etc. In addition there are access/egress links that represent going by car/taxi to a railway station or an airport. The time for these is a linear combination of ride time and price, together expressed in minutes by use of a value of time. So, there is a choice between lines including access/egress links.

In Sampers on the other hand, there are only notional access/egress links from origin and to destination to and from "main modes" in the main mode model, not real lines. In order to avoid unrealistically long access/egress parts these access/egress links have been given the speed 5 km/h. The access/egress mode choice is modelled separately in more detail using the regional network in Sampers where all lines are coded. After disaggregation of the main mode matrices to the regional level, the access/egress model is applied to the trips departing/arriving at the different terminals from/to their origin/destination zone.

3.5.4 Combinations of modes

The main mode concept in Sampers means that it will be difficult to evaluate measures where combinations of modes are needed to reach the destination. This is illustrated by the following example:

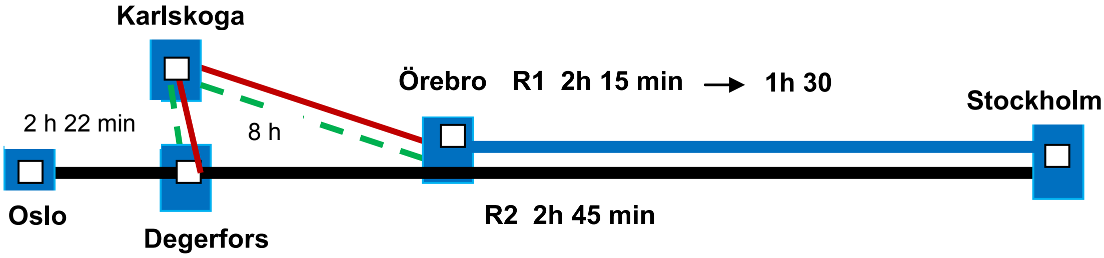
Between A and B there are airlines. Between B and C there are rail lines. Assume that we want to evaluate the effects of introducing high-speed rail between B and C for travellers going from A to C via B. The change in consumer surplus according to Vips is quite large but zero according to Samkalk. The reason is that Sampers disregards the possibility to go by air between A and B and then by rail between B and C. Since air is disregarded when rail is subject to change, Samkalk cannot calculate the effect.
The reason why the current Sampers implementation does not regard main mode combinations is that they were not found to be very frequent according to an analysis based on the national travel survey 1994-97. (For an empirical discussion see section 4.3)

3.5.6 Access/egress links

There is also another important aspect related to the Sampers main mode concept, which has to do with access/egress links.

To/from the main modes Sampers only uses access/egress links, each with a specified time, i.e., no public transport lines. The model assumes very low speed for these egress/access links in order to prevent too long access/egress links. If for example one can go by a train line to a city for which there are two possible alighting stations, Sampers tends to accept only the egress link with the shortest distance.

Assume that we want to evaluate the benefit to travellers from Stockholm to Örebro of reducing the ride time of rail line R1 (blue) Stockholm-Örebro from 2h 15 minutes to 1h 30 minutes while the speed on a more or less parallel line R2 (black) Stockholm-Oslo via Degerfors is unchanged. This black line does not stop in Örebro. See the figure below.



The red (dark) lines indicate regional buses according to Vips and the green (light) dotted lines egress links according to Sampers. The Sampers egress link Örebro –Karlskoga takes 8 hours and the Sampers egress link Degerfors-Karlskoga take 2h 22 minutes according to the actual Sampers coding of the network. Since Sampers does not take into account the real-world regional buses, only the egress links with excessively long times, nobody will choose to alight in Örebro and no travel time gain is calculated. Vips/Visum on the other hand takes the regional buses Örebro-Karlskoga and Degerfors-Karlskoga into consideration and calculates time gains for the travellers Stockholm-Karlskoga, via a transfer in Örebro. The question is of course to what extent the different approaches reflect actual travel behaviour.

3.5.7 Change of two modes

In the background report "Descriptive and theory report", there is an example of real calculations by both Sampers/Samkalk and Vips for business journeys between Sollentuna (15 km north of central Stockholm) and central Gothenburg. The reason for choosing this particular example was the – still ongoing – discussion about the likely effects of high speed trains on mode choice as well as whether high speed trains would be profitable or not in a social economic perspective. The intention here is not, however, to attempt to answer such general questions about high speed trains. Rather the idea is to illuminate differences in the response of the two model systems to changes of traffic supply similar to those that would occur with high speed trains, and to discuss – and possibly explain – how the different modelling approaches of the two model systems contribute to the differences of outcome between the runs with the two model implementations.

Let us first look at the reference situation before the assumed traffic supply changes. In the table below we give the initial distribution of business travel in the OD-relation between modes for the two model systems.

Table 3.5.7 Distribution of business travel demand between modes for the reference situation according to the Vips and Sampers model implementations. For the specific relation referred to in the table lines including air and rail totally dominated by one rail or air leg (>90 per cent of total trip length). The classification on mode for Vips is based on this dominating leg.

	Vips		**Sampers**	
Mode	**No of trips**	**per cent**	**No of trips**	**per cent**
Rail	384	66	213	37
Air	200	34	293	50
Car	0	0	77	13
Bus coach	0	0	1	0
Sum	**584**	**100**	**584**	**100**

The following two scenarios in terms of service changes were analysed and the corresponding outcomes of the model runs were observed:

a) That trains Stockholm-Gothenburg get half ride time (1.5 instead of 3 hours) in one step, called Rail time Down (RTD), and that airlines get 80 % lower frequency in a second step, called Air Frequency Down (AFD).

b) That trains Stockholm-Gothenburg get half ride time (1.5 instead of 3 hours) in one step, called Rail time Down (RTD) and that airlines get double ride time (or double price), called Air Time Up (ATU) in a second step.

- For scenario a) When rail ride time is reduced from 3 to 1.5 hours and air frequency is reduced by 80 % Sampers calculates that air still has a 20 % market share (down from 50 per cent) while it is only 1 % according to Vips.
- For scenario b) When rail ride time is reduced from 3 to 1.5 hours and air ride time is doubled Sampers calculates that air still has a 20 % market share (down from 50 per cent) while it is zero according to Vips.

We observe already in the reference situation that while Sampers distributes (some) demand over all modes, Vips only distributes demand over rail and air in this case due to the assumed fairly high inconvenience parameter for using car or coach for business travellers (See 3.3.5 above). This difference is primarily due to the fact that the estimated and calibrated logit model for business trips in Sampers in any OD-relation allocates demand over modes according to:
a) travellers' generalised cost including the mode constants that reflect the average travel conditions for business travellers,
b) with regard to the influence of random factors related to the traveller as well as the choice alternatives

In most cases as well as in this particular case the approach used by Sampers leads to some demand for all modes even if the differences in generalised costs are large
Vips, on the other hand, as we have discussed above, allocates demand over the feasible set of travel alternatives according to the RDT-formula (see 2.7.6). Neither car nor bus/coach belongs to the acceptable set in this case and therefore gets zero demand.

Looking only at the public transport modes (disregarding bus/coach) the market shares in the reference situation differ markedly between the two model systems in this case. In the reference situation the rail mode's share of total travel by public transport is 66 per cent for Vips and 42 per cent for Sampers. The corresponding shares for air are 34 viz. 58 per cent in this particular case. We cannot tell, however, whether one or the other of the outcomes are more or less in line with actual travel since we do not have access to observed travel data at this level.

With the changes described above, Vips predicts that the air mode share will change from 34 per cent to be close to zero while Sampers in both cases moves from 58 per cent to about 20 per cent. We note that the absolute change in percentage units is about the same for both systems. Thus the difference in outcome after the assumed changes seems to be partly influenced by the initial difference for the reference alternative with the two systems but elasticities look very different.[6]

We have also looked at the results for changes of consumer surplus (CS) according to Samkalk and Vips for the two changes of supply in this particular relation as assumed above.

For "Rail Time Down" calculation according to the principles assumed to be used by Samkalk and Vips gives nearly the same result. The similar results in this case for ride time change may depend on the fact that rule of the half for ride time changes is an approximation of the integral used by.

For the other assumed traffic supply change alternatives, Air Frequency Down (AFD) and Air Time Up (ATU) our calculations indicate that there may be significant differences between the outcomes using on one hand the Vips way of calculating CS changes and calculation according to the principles used by Samkalk on the other. We will discuss possible reasons for these differences in chapter 4. As mentioned above the most important thing to study in depth is to what extent these differences are due to different properties of the demand model viz. the algorithms for calculating consumer surplus for the two model implementations.

[6] Competitive changes of traffic supply as assumed in this case may lead to competitive responses e.g. adjustments of fares and fare structures for air, which may modify the impact on travel volumes. Neither model allows for this

4 Moving towards the ideal model for long distance travel

4.1 Introduction

In this chapter we will identify some key issues for further consideration and more in- depth analysis in modelling long distance travel. The choice of key issues to deal with is based on the general considerations about requirements on long distance travel models for the specific purpose of infrastructure and traffic planning that we have discussed in chapter 2. It is also based on our scrutinizing the current model implementations that are in practical use in Sweden in the light of these requirements, as well as a critical comparative analysis of these models in chapter 3. Further we have benefited from recent years' research and development experience.

We need for obvious reasons to model real world travel behaviour. Models require assumptions on rules for this behaviour. It should be obvious that we will never be able to exhaustively define rules that will allow us to predict travel exactly, even if we had perfect data. Therefore no model will be "correct" in the sense that all assumptions are valid. It will always be possible to reject any model because assumptions are not fully valid. That kind of discussion is not helpful.

Knowing that we cannot hope to establish the "best" model, i.e. a model that is based on "true" assumptions, we need to search for a second best model. In this process, we need to identify the strengths and weaknesses associated with different approaches, and we need to identify strengths and weaknesses associated with actual implementations of different approaches ("third best" solutions).

We therefore try to pursue our discussion in a way which we think will be helpful in terms of searching for better approaches, as well as in terms of understanding how present applications are affected (is Sampers likely to underestimate effects and is the Vips implementation likely to overestimate effects or vice versa?). In broad terms it seems from what we have observed in our model runs that Sampers/Samkalk always gives a higher estimate than Vips of the CS effects when headway is changed and a lower estimate of the CS effects when there are combinations of modes where one mode is changed in terms of ride time or price.

4.2 Key issues identified for more in-depth analysis

We have identified five main areas that given the characteristics and properties of the current models seem to require more in-depth analysis, as we briefly set out below.

1) How important for travellers' decision is heterogeneity of traveller preferences, taste, specific context etc. There is evidence that the importance of such heterogeneity may vary considerably depending on a number of factors. However, the weight that is put on the heterogeneity issue will have rather profound implications on how the travel decisions could reasonably be modelled. There is considerable experience with using RUM models as a means to deal with heterogeneity, but the validity of such models for travel decisions seems to require careful consideration of the model specification.

2) A standard approach used in transport demand modelling when RUM models are used for certain parts of the total demand modelling problem is to use a so-called assignment model which takes a (possibly mode-specific) matrix of origin-destination demands, together with transport networks based on connected links each with "level of service" information, and attempts to allocate the demand to appropriate lines through the network. In the course of doing this, it also produces matrices of (minimum) cost between each O-D pair for the purpose of generating cost input data to the RUM model which is used to produce demand matrices. In a second step the assignment model is used to allocate demand to the network(s). In the course of our project the properties of this assignment model – particularly in the context of public transport – has emerged as crucial for modelling of long distance demand. It seems important that the assignment algorithms are reasonably consistent with the nature of travellers' line choice decisions as identified in chapter 2. This is all the more important since the assignment model influences the final results at two stages as explained above.

3) The overall demand model for modes and routes/lines should be specified to be sufficiently consistent with travellers' choice structure and behaviour. In this case we have identified the decision behaviour relating to travellers' choice of mode and lines/services as a potential problem area that needs to be addressed. This relates mainly to the so-called main mode issue.

4) An important user requirement in the context of infrastructure and traffic planning is that it is possible to estimate with reasonable precision the change of consumer surplus resulting from various policy measures. The demand model has to give important input to the estimation of CS change. The underlying detailed components of CS appear also as drivers of traveller behaviour in the form of changes of travellers' generalised cost of journeys. Thus the GC-components of CS change will also determine travel demand for modes, lines etc. These relationships between the demand model and the CS calculation make it important to ensure reasonable consistency between the demand model and the evaluation of CS.

5) The model that generates demand matrices should be consistent with the model(s) that deal with mode and line choice.

We observe that the key differences between the two approaches reside in a) the different treatment of mode choice (with Sampers applying RUM to a "main mode" definition, and Vips treating it as a "route choice" problem), and b) the different treatment of headway in the assignment of competing lines. This then provides a structured sequence to the following points.

Based on the general considerations above we have drawn *the following tentative conclusions:*

- An approach where travellers' decisions are assumed to be between main modes and possibly sub-modes is based on the fact that there is empirical evidence that the modes per se play a role in travellers' decision making about journeys. If this assumption is correct it would be natural also to take advantage of the knowledge base around discrete choice models in developing the demand model. However, in the course of the project it has been questioned whether the actual decision structure for travellers is consistent with the main mode approach as well as whether the main mode approach is capable of dealing with all long distance journeys. The conclusion is that it is necessary to further consider the issue of main modes/sub modes in order to ensure that the behavioural model is specified and structured in a correct way. We have identified alternative ways to accomplish this such as applying a partial main mode approach, with PT and the car mode, applying a nested logit model structure or integrating heterogeneity with respect to travellers' mode preferences with the RDT model. There are also significant issues as to how the variable "headway", which enters into the calculation both of 'waiting time' (Emme/2) and 'mean schedule delay' (RDT), should be specified for the purposes of an RUM mode choice model. Further research will be required.

- An issue that has bearing on the issue of how to model modes and sub-modes is how important combinations of modes to accomplish journeys are in reality. This was discussed above in 2.6.5 as a largely empirical question. We provide additional and more detailed evidence below in a separate section. Though there could be alternative interpretations of this empirical evidence, we consider it important in principle to allow for combination of modes in an ideal model for long distance travel. Also, the decomposition of the mode and line choice problem into independent main mode and access/egress mode sub models should probably be reviewed. Of course, the extent to which effort is made to accomplish this – as usual – has to be traded off against user value, cost etc. The best way to accommodate long distance trips that involve combinations of two or more modes in the model requires further study.

- The so-called "Optimal Strategy" assignment algorithm that is implemented in Emme/2 seems better suited for urban transport than for long distance travel, because of the way in which the behavioural assumption that travellers use the timetable is dealt with. The RDT-algorithm has a more appropriate assignment property in the context of long distance travel, since this approach moves us a little bit closer to the full timetable based approach, which is the ultimately rigorous approach. We have arrived at this conclusion mainly based on theoretical considerations from the descriptions of the two algorithms.

- Using the more desirable assignment property of the RDT algorithm unfortunately comes at a price. Two problems exist – the first one is to establish a rigorous estimation/calibration method based on the existence of a desired calibration outcome. The second problem is related to the fact that the RDT algorithm models the expected route choice over all possible timetables, and not route choice conditioned on a specific timetable. As we assume that travellers will be sensitive to the time table, this means that the modelled route choice may be different from the observed route choice, the latter being conditioned on a specific timetable. It may therefore be difficult to find the proper calibration target. Further research is needed to find if and how these problems can be solved.

- Calculation of change of consumer surplus: The consumer surplus calculation will never be better than the underlying demand models. As long as schedule delay is not explicitly modelled, such effects will always be neglected using whatever measure. As long as traveller heterogeneity is not included in RDT applications, changes in GC will also give wrong estimates of CS changes. An additional question is the ability of different software to accommodate political constraints on the values used.

- There should be consistency between the different levels of an overall demand model. In this study we have concentrated only on the model level that deals with mode/line choice. We have observed, however, that further development of this model level in accordance with our conclusions will also make it necessary to investigate the implications for other model levels relating to overall demand. A specific issue that needs to be addressed is what input data on composite GC for an OD-relation should be used for estimating OD-flows.

4.3 Empirical evidence on combinations of modes for journeys

4.3.1 Data sources

For long distance travel, the main planning issues are related to public transport modes designed to cover longer trips such as air, long distance trains and long distance bus. We call these modes long distance travel modes. The use of such modes may be combined with uses of modes primarily designed for shorter trips, such as local and regional buses, metro, tram etc. These modes are usually used for terminal access and egress. Combinations of long distance modes may also be used. In addition to combinations of public transport modes, there may be combination involving different types of car use such as private car, hired car and taxi. Disregarding these combination possibilities implies that effects of facilitating public transport access and egress to the long distance modes are neglected, and that possibilities of long distance mode cooperative policies cannot be analysed. The issue of combining different modes to accomplish a journey also attracts considerable political interest. Therefore, as argued above, combinations must be regarded as an important aspect to take into consideration in the future development of models for long distance travel. In this section, however, we concentrate on giving an up to date picture of the use of mode combinations for long distance journeys based on available empirical evidence.

We are interested in finding out to what extent different modes are used on the same trip leg (in the sense of the outward or return journey). This is important for the definition of the main mode approach, building on the assumption that combinations of long distance modes are rare enough to be ignored, and for assessing the performance of the Vips application allowing combination of all modes without constraint. For both approaches, we are interested in to what extent several main modes are used on a trip leg, and also to what extent access and egress modes are used. This assessment requires information on all used modes for a certain trip leg and the associated trip lengths.

There are several Swedish survey sources for information on combinations between modes. One is the national travel surveys, of which RES05/06 is the most recent part, based on surveys from September 2005 to August 2006. Another one is the Tourist Data. A third data source is the Swedish Public Transport Coordinator (Samtrafiken i Sverige AB), from which information is available based on actual ticket sales. A forth data source is air and train terminal surveys. A fifth data source is individual operator statistics.

The RES05/06 source records all trips made on a pre-specified day (including long distance trips), and contains information on all modes used on the trip leg, including the distance travelled for each mode. In addition RES05/06 also records long distance trips (>100 km single trip) made during the month preceding the pre-specified measurement day, and longer distance trips (>300 km single trip) made in the month before. For the first and second month trips, information is included on the main mode (defined as being used on the longest distance) as well as on the mode used to access/egress the main mode terminal.

The TDB source records long distance trips (>100 km single trip) made in the last month, and contains information on what modes are chosen, but not the length travelled by each mode. The definition of the bus mode does however not permit a subdivision into local/regional and long distance buses (neither RES nor TDB permits a subdivision of local/regional bus use into local and regional bus use as respondents have not been assumed to be able to tell the difference between local and regional bus lines).

The Samtrafiken statistics includes information on a part of the tickets sold, which will make it possible to recognise public transport mode combinations for journeys where tickets have been purchased through Samtrafiken.

Here, we will try to give an idea of how these sources may contribute to give a picture of the extent of mode combinations. In this project, our resources do however not permit a full analysis of these sources.

4.3.2 RES05/06

We begin with the RES05/06 measurement day data. For this day, only 280 public transport trip legs were observed (which is a reason for the longer retrospective part of the survey). This is unfortunately the only data source that contains information on all modes and associated trip lengths.

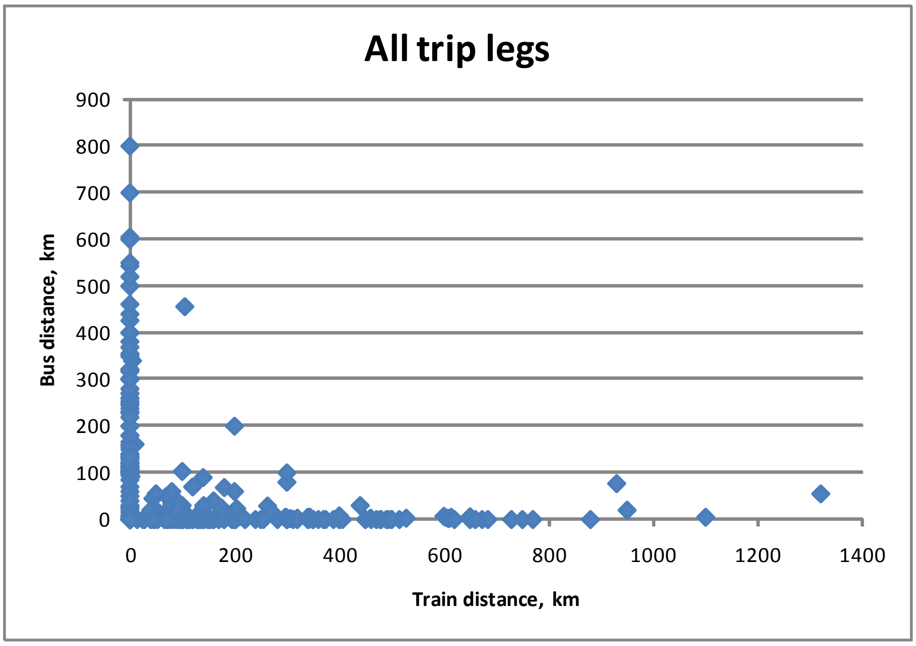
In the table below, , we classify trip legs into the long distance modes Air, Long distance trains and Long distance bus according to the longest distance travelled. Then, in the first column, we have calculated the share of trip legs for which some distance has been covered also by another long distance mode. In the second column, we have calculated the share of trip legs including local/regional bus in addition to other long distance modes. Finally, in the third column, we have included the share of trip legs including any other public transport mode.

Table 4.3.2.1 Share of mode combinations for main modes

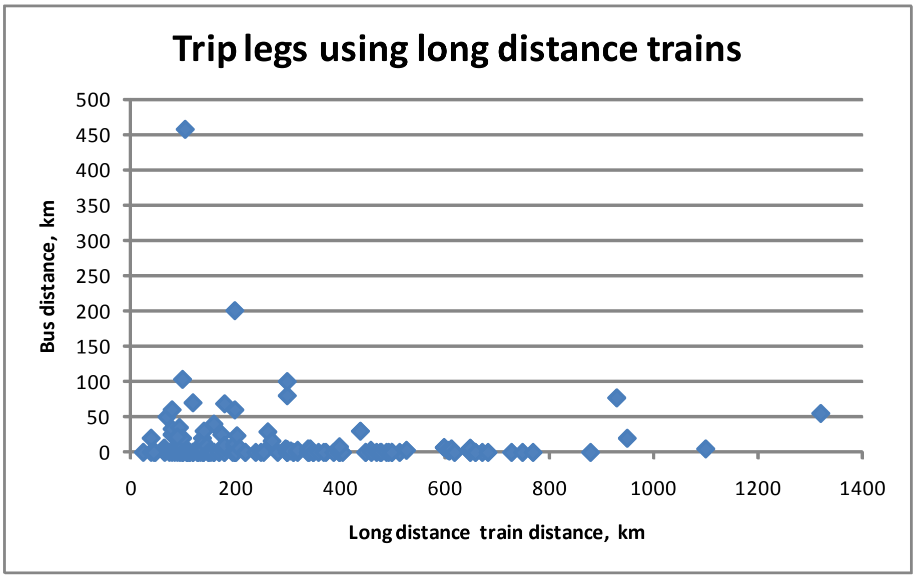
Long distance mode	Share of long distanced mode (left column) combined with other long distance modes	ditto incl local/regional bus	ditto incl all local public transport
Air	0,27	0,41	0,79
Train	0,02	0,34	0,63
Long distance bus	0,00	0,13	0,27
All modes	0,06	0,31	0,58

All local and regional public transport (buses and commuter trains) are included in the access/egress mode choice model in the Sampers system. In the Vips implementation, on the other hand, all regional bus lines, all airport bus lines but no local bus lines are integrated into the overall network that is used for long distance travel. The main mode concept is only used in Sampers, however, in a way that confines travellers to one and only one main mode. For the Sampers approach, therefore, the relevant column for main mode combinations is therefore the first one, showing an average share of main mode combinations of 6 percent. It is mainly in context with air trips that there is a substantial share of other main modes used. The major additional mode is long distance train, but the trip length for this mode is mostly below 45 km. In the case of long distance trains, the maximum trip length for the additional long distance bus is 30 km. In the second column local and regional buses have been included, to illustrate the effect of that inclusion. In the third column, all public transport modes are included, i.e. also subway and commuter trains.

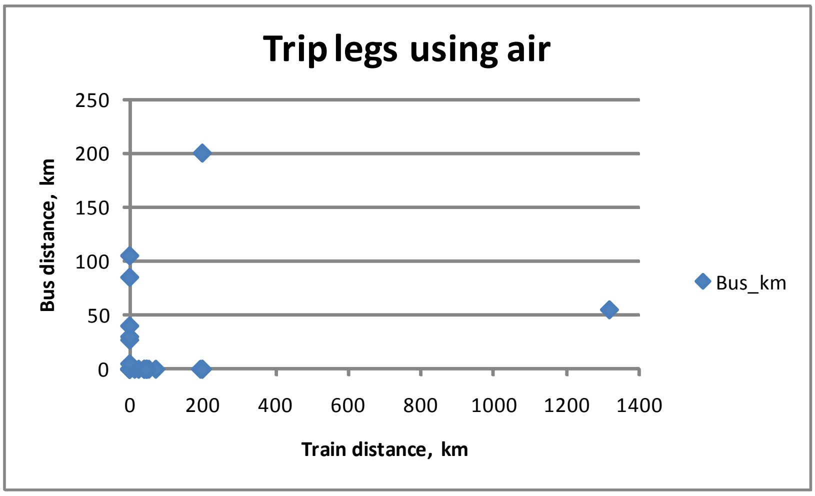
Not only the shares of mode combinations are important, but also the way of combining the modes – are modes combined in a “main mode fashion”, i.e. one mode is used for a minor distance share and the other mode for the rest, or are modes combined in a wider range of shares? In the figures below, distances covered by train and bus are plotted in a scatter gram, having train distance on the x-axis and bus distance on the y-axis (regardless of type of train or type of bus). If modes are combined in the “main mode fashion” the observations will be close to the x- and y-axis, and in the case of a wider range of combinations, observations will more scattered. In the first figure, all public transport trip legs have been plotted, regardless of other modes chosen (including car).



As can be seen, the main picture is that most trip legs are close to the x- and y-axis.
If we look at trips legs where long distance trains have been used, the following picture emerges:

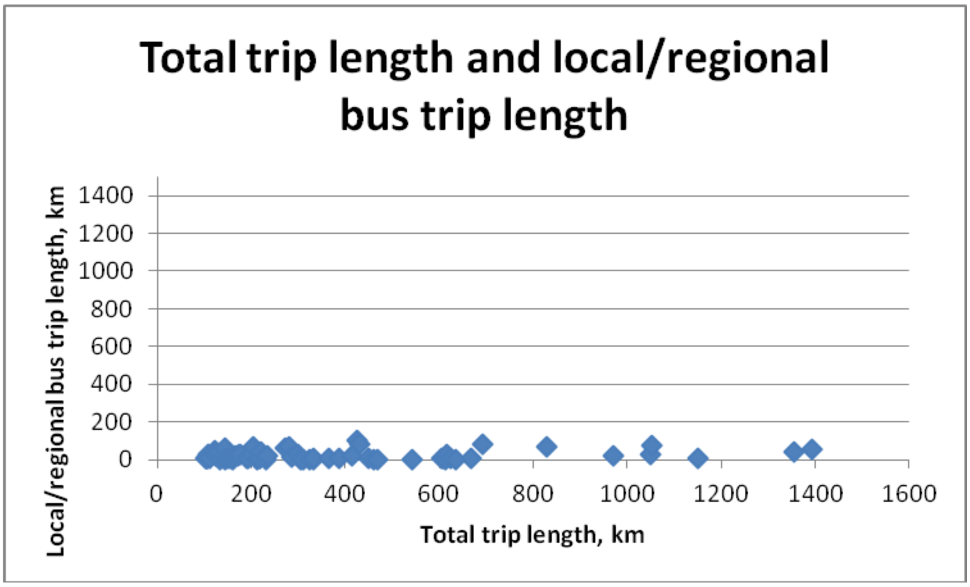


Most trip legs where long distance trains are used contain no or very short bus trips, although examples exist where trips are longer.



In the case where air is used on the trip leg, train and bus are mostly used in the normal access/egress range for airports.

To justify the current main mode approach in Sampers, there should be a low share of main mode combinations, and other modes (access/egress modes) should only constitute a minor part in distance terms of the total trip. In the figure below, the total trip length and the local/regional bus trip length distributions are shown for long distance trips including local/regional bus (access plus egress) based on the measurement day data set from RES 05/06



As can be seen for the figure, the local/regional bus part is quite small irrespective of total trip length. The shares for the local/regional bus part for different total trip lengths are shown in the table below:

Total trip length	100-200 km	201-400 km	401- km	All
Local/regional bus share	0,13	0,07	0,05	0,09

The figures reported from the RES 05/06 travel survey seem reasonably consistent with the analysis of the TDB data where long distance and access/egress modes are not separated. As trip lengths for the different modes used on a trip leg is not reported, no distance based main mode definition can be made.

4.3.3 TDB

The analysis of TDB data source shown below is made by Gunnar Lundgren, consultant for SJ (the Swedish State Railways).

The table below shows the percentage of combinations between various public transport modes of all journeys with public transport, according to TDB.

Table 4.3.3 Combinations between public transport modes of public transport journeys according to TDB. All types of public transport are included, i.e. also local and regional lines

Combinations	%
Bus/train to/from air	83
Air/train to/from bus	14
Air/bus to/from train	40
Average all transfers	**40**

Table 4.3.3 shows that of those who use public transport for the whole journey, 40 per cent use more than one mode. We can also see that the number of transfers is related to the number of stops per mode, since there are more bus stops than train stations and more train stations than airports. Therefore, evidently the largest need for combinations concerns air. There is **no** explicit “main mode” classification in this data – we are merely looking at the connections “on either side” of each identified mode

4.3.4 Samtrafiken[7]

Samtrafiken i Sverige AB (English official name: the Swedish Public Transport Coordinator) is the Swedish organisation in charge of integration of Swedish transport lines and modes in order to facilitate travelling and ticket purchase where several modes or operators are involved. It is owned by 32 transport companies or authorities, among them all 21 County Transport Authorities (Länstrafikbolag) who operate regional buses, all Swedish Rail operators, ferry operators and Express bus operators. Most combination tickets are train to train. Thereafter bus to train or vice versa, while a few are bus to air or bus to express bus. Note that combination tickets local bus to train or vice versa are negligible, which means that bus here is read regional bus.

These combination tickets are called Resplus (approximately Travel plus in English). The number of registered combination tickets bus/air will grow since such tickets were launched only in 2010.

The Resplus information system has 5 000 stops mentioned by name, but travellers can buy Resplus tickets from/to about 200 000 bus stops. The number of bus stops to/from train stations where combination tickets are sold is almost 5 000 per year.

In one year, July 1 2008 to June 30 2009, the number of Resplus bus/train combination tickets sold were about 1 200 000, i.e., pre-purchased tickets with at least two modes, of which one is

[7] This section is based on information from Robert Enskog, Samtrafiken, and from Anders Hansson, former general manager of Samtrafiken

bus. Of these 793 000 were sold in the 150 most sold bus-train combinations, bus stops to/from train stations.

About the same number of journeys with at least two modes, of which is bus is one are sold outside the Resplus system. One reason is that many people already have County travel cards valid for regional buses. Another reason is that some travellers buy tickets mode by mode along the journey. This means that altogether about 2.4 million travellers use combinations bus/train.

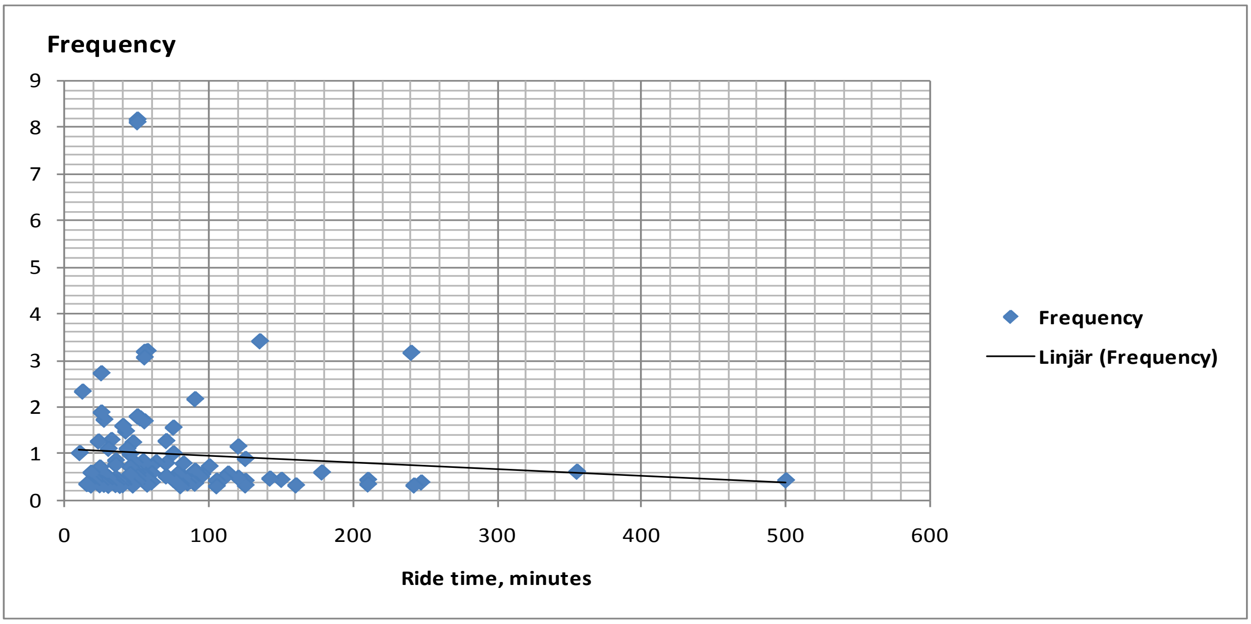
According to the Sampers matrices used in the project there are about 19.5 million private public transport journeys and 7.3 million public transport business journeys per year, altogether 26.8 million. Of all journeys then about 9 per cent use the combination bus/train. If we assume that only private travellers use the combination regional bus/train the share of private travellers who use combinations bus/train is around 12 per cent. Maybe more realistically, if half of the business travellers and all private travellers use combinations bus/train the combination share of all journeys is about 10 per cent. In addition there are combinations airport bus/air express bus/air and rail/air, see next section.

Of the 100 most sold stop-station pairs including bus in the year July 1 2008 to June 30 2009, 46 had a bus ride time over 1 hour and 16 had a bus ride time over 2 hours. The average ride time, weighted with number of trips, was 70 minutes.

The diagram below shows the frequency of each ride time and a linear regression line.

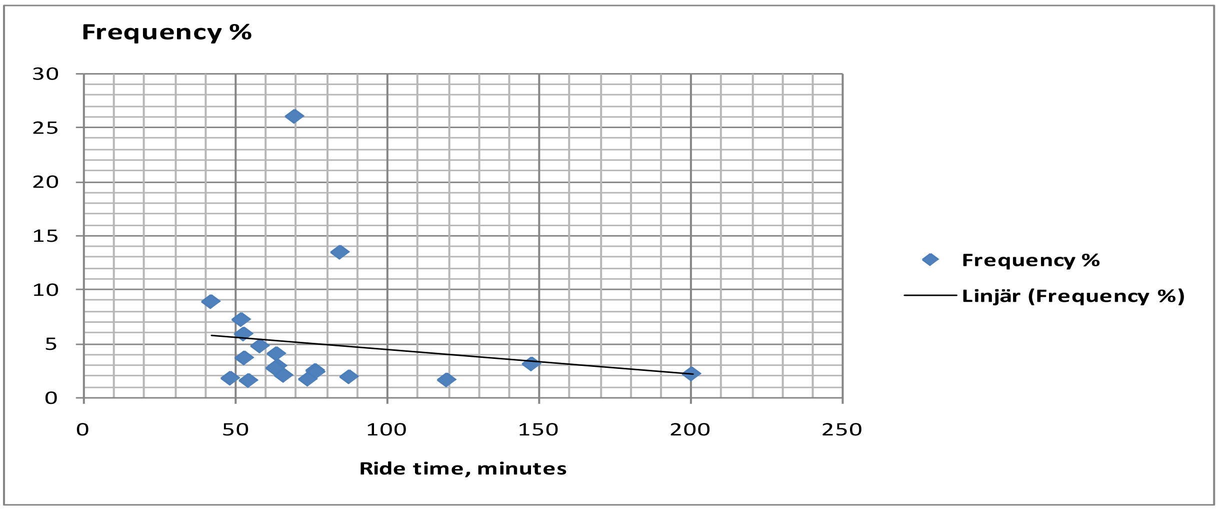
Note that some bus ride times to or from a rail service take very long, 250 and 350 minutes and one even 500 minutes.

Diagram 4.3.4.1



In the next diagram the 100 bus stops to train stations are aggregated. The 100 are separated into 20 classes to show the results for the average ride times of these 20 classes.

Diagram 4.3.4.2



Concerning the concept "main mode", one may now wonder whether some of these bus rides to/from train are main modes and not access/egress ride to/from train. In such cases the Sampers model would classify such bus lines as main modes. According to information from Samtrafiken such cases are negligible, but with one exception. The 500 minutes bus ride is between Luleå and Sundsvall, from where the travellers take the train. For the travellers who go southbound to Stockholm, the train distance is about 380 km, while the bus distance Luleå-Sundsvall is about 540 km. For those who go further south by train, the train distance is typically longer than the bus distance. However, it may be that Sampers regards that Luleå-Stockholm has main mode bus.

If we remove Luleå-Sundsvall as being an access bus ride to train, the average bus ride time is 68 instead of 70 minutes. The diagrams below show the results when Luleå-Sundsvall is removed.

Diagram 4.3.4.3

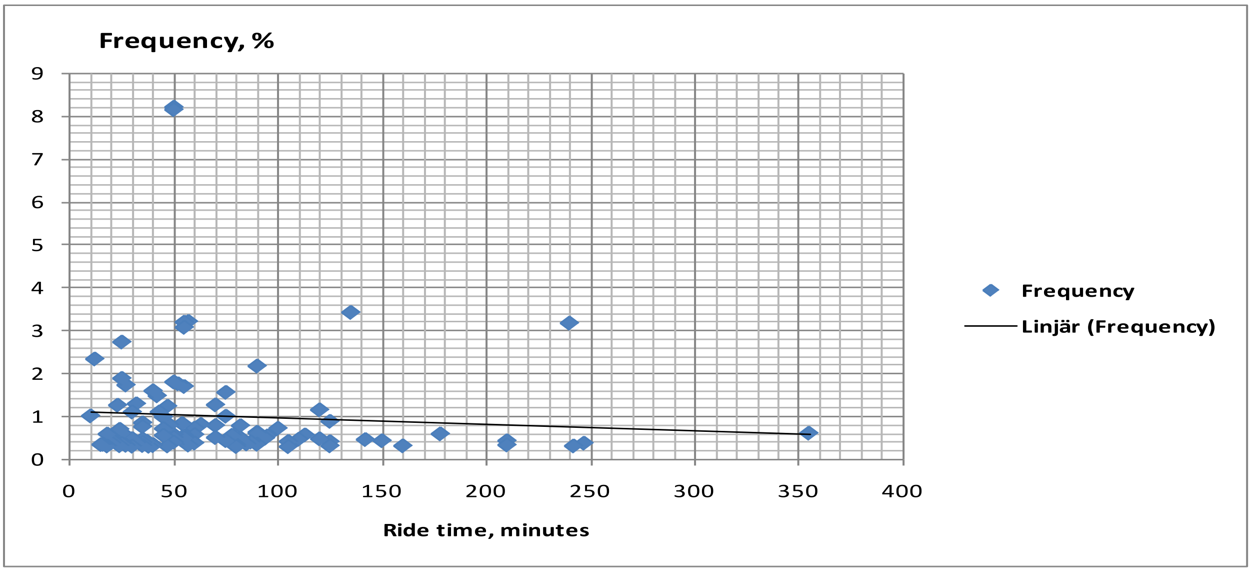
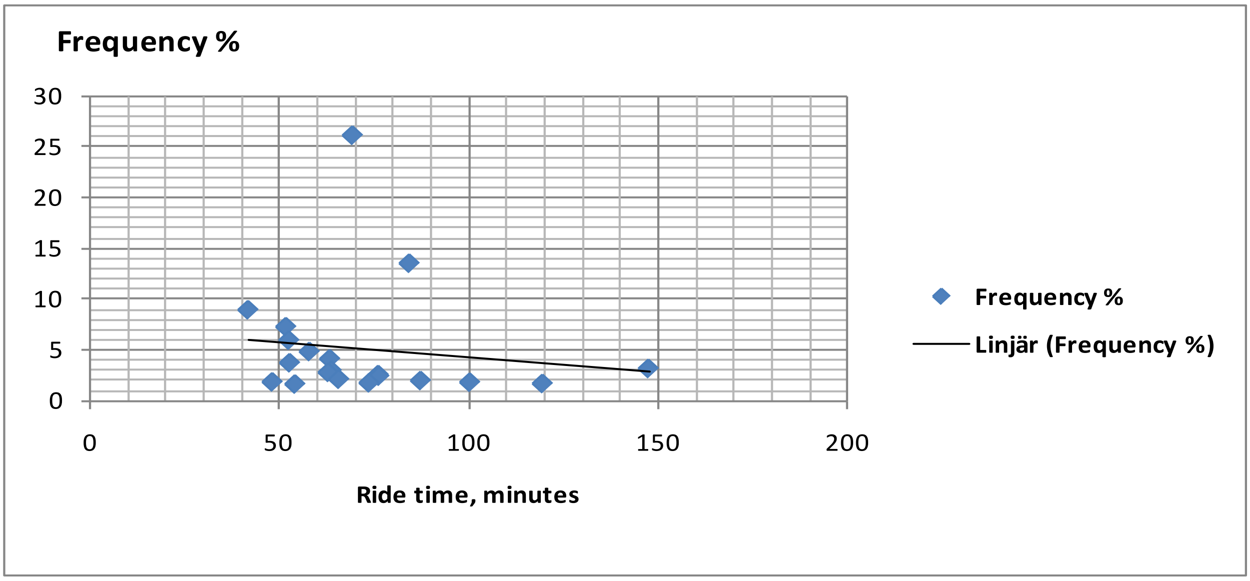


Diagram 4.3.4.4



4.3.5 Access to Arlanda airport

The information in this section is delivered by the Swedish airport authority Swedavia.

Arlanda airport 45 km north of Stockholm is by far the largest airport in Sweden. 60 % of the air travellers depart/ arrive here. According to Swedavia there were in 2009 about 16 million travellers departing from/arriving at Arlanda airport, of which 4 million are domestic journeys. Of the domestic air journeys Arlanda airport has 30 % of the travellers.

The table below shows the last mode used for arriving to the airport, based on about 100 000 interviews per year. According to Swedavia, it is most reasonable to assume that travellers use the same mode both to and from the airport.

Table 4.3.5.1

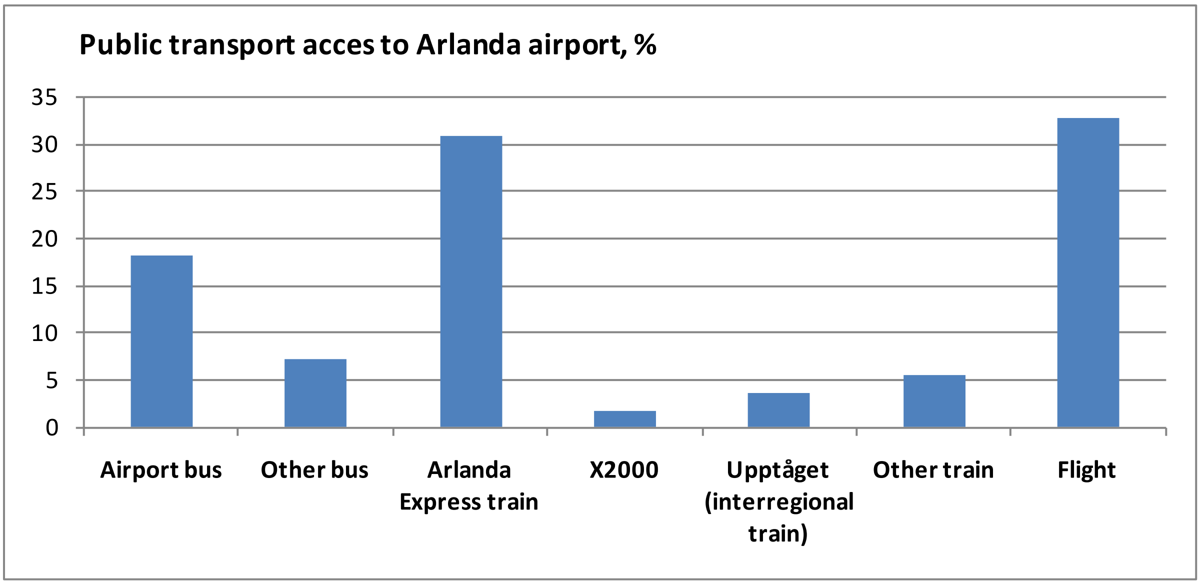
Transport to/from Arlanda airport		
All modes	**%**	**Number of journeys**
Car	28	1,1
Taxi	15	0,6
Airport bus	10	0,4
Other bus	4	0,2
Arlanda Express train	18	0,7
X2000	1	0,0
Upptåget (interregional train)	2	0,1
Other train	3	0,1
Flight	19	0,7
Sum	100	4,0

Table 4.3.5.2

Public transport to/from Arlanda airport		
Mode	**%**	**Number of journeys**
Airport bus	18	0,5
Other bus	7	0,2
Arlanda Express train	31	0,9
X2000	2	0,1
Upptåget (interregional train)	4	0,1
Other train	5	0,2
Flight	33	0,9
Sum	100	2,9

The diagram below shows access by public transport.

Diagram 4.3.5.1



Of access by public transport 2/3 use other public transport modes than flight.

Airport buses depart from Stockholm city terminal and Uppsala. Arlanda Express train departs from Stockholm city terminal. The speed train X2000 and other trains are long-distance rail services departing from various cities south and north of Arlanda.

The tables below show the itineraries of various SJ products that stop at Arlanda airport, and the number of travellers in 2009, according to SJ sales statistics.

Table 4.3.5.3

No. of journeys, in 2009 Itinerary, county or city	Product	To ARLANDA No. 2009
County of Dalarna-Stockholm	InterCity	64 204
	Regional	7 199
	X2000	5 866
Not allocated	InterCity	1 417
Eskilstuna-Stockholm	Regional	3 580
Gävle-Göteborg	X2000	3 720
Gävle-Stockholm-Linköping	Regional	73 274
Göteborg-Stockholm	X2000	55
County of Jämtland-south Sweden	Night	5 302
Oslo-Värmland-Stockholm	InterCity	1
Sundsvall-Stockholm	X2000	57 205
Uppsala-Stockholm	Regional	10
Östersund-Stockholm	InterCity	4 287
	X2000	3 421
Östersund-Sundsvall-Stockholm	X2000	6 609
North Norrland-south Sweden	Night	10
Total		236 160

Table 4.3.5.4

No. of journeys, in 2009 Itinerary, county or city	Product	From ARLANDA No. 2009
County of Dalarna-Stockholm	InterCity	49 906
	Regional	5 440
	X2000	4 624
Not allocated	InterCity	582
Eskilstuna-Stockholm	Regional	81
Gävle-Göteborg	X2000	1 911
Gävle-Stockholm-Linköping	Regional	52 062
County of Jämtland-south Sweden	Night	3 824
Sundsvall-Stockholm	X2000	47 830
Östersund-Stockholm	InterCity	4 340
	X2000	2 625
Östersund-Sundsvall-Stockholm	X2000	5 120
Total		178 345

All in all there are over 400 000 journeys with SJ-trains to/from Arlanda airport in one year.

Of access by train as the last mode used, 11 % of the travellers use other trains than Arlanda Express. Also, of those who use Arlanda Express trains of airport buses from the Stockholm city terminal, some use other trains or buses to Stockholm city terminal before they board Arlanda Express or the airport buses to Arlanda airport.

An Upptåget interregional train service, managed by Upplands Lokaltrafik (Upplands County Transport Authority), is going Gävle-Tierp-Uppsala-Arlanda airport-Upplands Väsby.

From Gävle there are 10 departures per day to Arlanda airport. From Tierp and Uppsala there are 32 departures per day to Arlanda airport.

The table below shows the number of stations and distances.

Table 4.3.5.5

	Gävle to Tierp	Tierp to Uppsala	Uppsala to Arlanda	Sum Gävle to Arlanda
No of stations between	5	5	1	11
Kms between	51	61	30	142

The table below shows all bus lines that go to/from Arlanda airport.

Table 4.3.5.6

Destination	Line	Operator
Arboga		Arlandalinjen
Enköping via Sigtuna-Bålsta	803	UL
Enköping	866	Flygturen (Swebus Express och)Västmanlands lokaltrafik
Gävle		Y-buss
Karlskoga		Arlandalinjen
Karlstad		Arlandalinjen
Kista		Flygbussarna
Knutby via Långhundra och Almunge	806	UL
Kristinehamn		Arlandalinjen
Köping		Arlandalinjen
Märsta station, pendeltågsstation	583	SL
Norrtälje via Rimbo	806, byte Åby vägskäl till buss 677	SL
Rotebro via Rosersberg, Östra Upplands Väsby	538	SL
Stockholm City (Cityterminalen) via Ulriksdal/Järva krog, Frösunda, Frösundavik, Hagaterminalen och S:t Eriksplan		Flygbussarna
Stockholm City (Cityterminalen)		Flygbussarna
Stockholm City (Cityterminalen)		Swebus
Nattbuss: Stockholm City via Märsta station, Upplands-Väsby, Rotebro station, Norrviken station.	Nattbuss: 592	SL
Nattbuss: Stockholm City	Nattbuss: 593	SL
Stockholmsmässan /Älvsjö		Flygbussarna
Umeå		Y-bussen
Uppsala station och centrum	801	Flygbussarna (Upplands Lokaltrafik)
Västerås	866	Flygturen (Swebus Express och Västmanlands lokaltrafik)
Västerås		Arlandalinjen
Örebro	866	Flygturen (Swebus Express och Västmanlands lokaltrafik)
Örebro		Arlandalinjen
Östra Steninge via Märsta Station, Valsta	589	SL

Of the bus lines to/from various places in the table above the following cities are more than 70 km from Arlanda airport.

Table 4.3.5.7

City	Kilometres
Arboga	166
Enköping	79
Gävle	138
Karlskoga	253
Karlstad	318
Kristinehamn	276
Köping	151
Umeå	609
Västerås	112
Örebro	205

4.3.6 Summary of results from Samtrafiken and data about access to Arlanda airport

Altogether there are in one year about 2.4 million travellers use combinations regional bus to/from train.

According to table 4.3.2 above the number of public transport journeys to/from Arlanda airport is as follows.

0.3 million are using SJ- trains to/from Arlanda airport.
0.1 million are using the Upptåget to/from Arlanda airport.
0.2 million are using regional bus or express bus to/from Arlanda airport.

Altogether 0.6 million travellers are using these public transport modes to/from Arlanda airport.

If we then add the 2.4 million travellers who use combinations regional bus to/from train there are about 3.0 million travellers who use regional bus to/from train stations and regional bus, express bus or train to/from Arlanda airport.

Of all public transport private and business journeys per year, 26.8 million, and these 3.0 million journeys constitute a mode combination share of 12 %.

In this combination share we have not included airport buses and Arlanda express trains to Arlanda airport. The number of travellers with these modes to/from Arlanda airport is 1.4 million. If we add the number with these modes the number of combinations is 4.4 million of 26.8 million journeys in total, which means that the combination share of all journeys is 16 %.

In addition there are regional bus and express bus services to other airports in Sweden, for which we have no information. Note that the combinations at these other airports may be large since they have 70 % of all domestic air journeys.

Altogether we conclude that the combination share of all domestic journeys is around 15 per cent.

If also international journeys to/from Arlanda were considered and analysed, the number of combinations that should be taken into account at Arlanda is 2.0 million instead of 0.6 million, and consequently the average combination share would be larger than 15 per cent.

5 Ways forward

5.1 New user requirements call for development of existing model tools for long distance travel

The aim of the project, as given in section 1.5 above, is to give recommendations as to the way ahead for the further development of long distance travel demand models in Sweden. These recommendations should be based on general considerations of user requirements for such models and the findings in the present project, as well as on recent research and development within the field. The time has now come to speak of these things.

Of course there is a need for continuous and gradual development of all model tools that are used in transport planning. However, regarding demand modelling for long distance travel, there are particular reasons to consider more substantial reviews of existing model tools, i.e. the current standard Sampers implementation, Vips, and Visum. These reasons are to a certain extent related to some weaknesses of the long distance Sampers model that have been observed by the users and were briefly touched upon in section 1.3 above.

In addition to this, and possibly also related to the criticism that has been voiced, rather fundamental changes to long distance transport supply are now envisaged. High speed train lines are being seriously considered, regular train traffic continues to attract considerable political interest, the role of air transport may have to be reassessed, and car transport, which is still the dominating single mode for long distance travel as well, is under pressure for various reasons. The national public transport agencies as well as regional and local authorities will have to be well informed in order to be able to respond in a rational way to the challenges that will come up in the course of these developments. Against this background there is an urgent need to reassess and also improve the available model tools for long distance demand modelling.

The revised planning system, where the four year planning cycle has been replaced by a more continuous evaluation and planning approach, will require all model tools to be more flexible as well as versatile, since there will be less time and resources available for large scale up-dating of comprehensive set of input data. The new organisation of the Swedish national transport agencies, with more emphasis on the intermodal issues, will also provide a role for model use and design. It will be even more important than before to put emphasis on journeys, movements, and interconnections rather than the transport modes themselves.

5.2 Development should be planned in different time perspectives

As usual the users' need to have access to better models is urgent. However, it is also well known that there is a need for stability in the planning situation, which clearly limits the pace at which changes could be implemented with the users. Moreover, users, researchers, and developers are all aware of the fact that model development takes time and that quality assurance of the development products will add considerable time before new/revised models are ready for implementation.

We therefore suggest looking at further development of long distance transport models in three time perspectives: short run (0-2 years), medium run (2-4 years), and long run (4-10 years).

Typically, short run development will use existing algorithms and also largely rely on existing data: the basic effort will be to do the programming, testing etc. that is required to implement such improvements.

In the medium run it is realistic to envisage development of new algorithms and programs that could be based on new but limited sets of data, as well as combining elements, e.g. algorithms, from different model implementations and introducing new or amended model features within existing model implementations e.g. Sampers, Visum, and Vips.

In the long run it is realistic to consider more sweeping development strategies such as developing totally new model systems with new algorithms, newly estimated behavioural models based on new, possibly large, data sets. It would also be possible in the long run to re-consider the overall approach to the management and coordination of models, data, and scenarios.

5.3 Suggestions for short and medium term developments

Based on the conclusions we have drawn in chapter 4, we consider it wise in the short to medium term to pursue three main parallel development lines. One is related to the RUM/Logit approach. The second deals with development of the RDT algorithm in various respects. The third line tackles the possibilities for taking advantage of the strengths of the logit approach combined with the RDT assignment algorithm.

Refine an abstract "ideal model" as a benchmark in the development process
The development activities discussed below are all more or less directed towards implementation in various time perspectives. However, we think that all such development activities could benefit from the existence of a common general benchmark, which we denote as the "ideal model". Ideas about and outlines for such an ideal model have been discussed in the project. We suggest that these outlines be further synthesised and developed to refine/define the "ideal model" as an appropriate benchmark: There is no proposal to implement such a model, but merely to establish, in mathematical form, the key influences, so that the potential "loss" from any practical proposal can be assessed against the ideal.

Development of logit models
The project has cast some doubt about the suitability of the specification of the MNL that is implemented at the mode choice level in the standard Sampers model. There are many alternative logit model specifications and model structures that may lead to a better model specification than the current one, as has been demonstrated in research projects related to the development of long distance logit models (see for example Beser Hugosson 2003). This development may have to draw on supporting studies of travellers' decision structure as well as a thorough assessment of the available evidence on how various combinations of modes and sub-modes have been and possibly will be used in the future.

Development relating to the RDT-algorithm
Though the RDT algorithm operates with the assumption that travellers know the timetable it has some weaknesses, namely that it does not take traveller heterogeneity into account and that it assumes uniform distributions of travellers desired departure time viz. departure time of services. One way to improve and expand the properties of the RDT-algorithm in the context of modelling long distance travel that we have identified and taken first steps to develop in this project is to modify it with a mechanism that allows various aspects of traveller

heterogeneity directly to influence the algorithm. The proposal is therefore to promote development of theory, algorithms, and practice for a random term relating to “mode” that could be combined with the RDT-algorithm according to the ideas presented in appendix 1. This methodology could be used to improve the present RDT-approach to include some influence from various random factors relating, e.g., to modal “tastes”. It could also be incorporated in both options a) and b) below.

Another potential improvement of the RDT-algorithm is related to the basic assumptions about uniform distributions of traveller desired departure time and the uniform distribution of supply of public transport departures. The question that has to be studied is whether it would be possible to use alternative assumptions about these distributions based on empirical evidence, and if and how such distributions could be implemented in the algorithm. A special case is to do this only for the desired departure time for travellers. Development efforts along these lines should be supported by empirical studies giving further evidence about travellers´ actual departure times, possibly in separate time slices over the day.

Combinations of RDT and logit
An interesting option that we suggest should be investigated within the existing structure of models is to combine cost calculations by the RDT algorithm with the logit model approach in various ways, as explained below.

a) Apply RDT to all modes; construct “composite” generalized cost to use in Sampers destination choice

b) Apply RDT to pt modes only (with car costs generated by standard network models), construct composite pt GC for Sampers mode choice (restricted to car vs. pt), and then destination choice

A preparatory step for option a) would be to make a systematic comparison between relevant logsums and composite GC as estimated by the RDT-algorithm. The aim of this exercise would be to assess the magnitude and the nature of possible systematic differences. Given that the cost data will be different in each of these options, the Sampers choice models will have to be re-estimated as well as re-calibrated.

One could also consider a similar approach to that in points a) and b) above to calculate GC with the RDT algorithm separately for each of the Sampers-defined “main modes” and pass the thus calculated GC to Sampers’ mode choice and then destination choice. It is understood that this option is potentially tedious for the RDT model, but it should be possible to proceed by “discouraging” use of other modes for the long-haul section of the journey. Although possible improvements of GC-estimates may also improve the outcome of subsequent model steps in Sampers (mode and destination choice) in this case, it would be a useful test as to whether the main mode approach is worth pursuing in this way. The gain would be to continue to allow for the RUM random terms within main mode choice.

The issue of integrating the access/egress modes still remains. Even though the vast majority of access/egress is undertaken by car, regional and local modes are used to reach railway stations or airports. Some O-D pairs need the combination of air and rail. Many O-D pairs need combination between IC-trains and X2000 (or high-speed trains).

As of today there is no systematic and scientifically sound way to calibrate the parameters of the Vips model with the RDT-algorithm. We have pointed out that there may be difficult issues to deal with here. We suggest that it is investigated whether one can find a sound scientific method to estimate parameters with the RDT approach, as is done for discrete choice models. However, both for discrete choice models and models using RDT, there is a potential problem insofar as neither approach is sufficiently close to the "true" model. Whether we estimate or calibrate we will have observations that are generated by a different mechanism than is assumed by either model: this could cause the estimation/calibration to produce biased parameters.

Timetable based options

In the medium term we think that timetable-based assignment options should be investigated carefully. However, there are two options at hand here that could already be addressed in a pre-study in the short term, namely the Sampers timetable model and the Visum timetable based assignment. With a timetable one can abolish one of the two assumptions behind the RDT principle, i.e., uniformly distributed departure times of competing lines.

The development of the Sampers departure time model is based on supporting studies of the asymmetry between early and late arrivals. As is expected, late arrivals are valued more negatively than early arrivals. These results are available if one would decide to go ahead with the timetable based approaches.

As mentioned earlier the Visum timetabled based version takes into consideration this asymmetry between early and late arrivals.

There have earlier been worries that timetable-based models require too much information and are costly to run. But some developments have appeared over the last years. In the near future it may be that also for the whole of Sweden all timetables for all lines and modes can be exported to Visum, Emme, or some other timetable-based software.

Calculation of consumer surplus change

Though an important role of long distance travel demand models is to provide projections/estimates about various aspects of demand such as the distribution of long distance travel over modes and lines, the models have an important additional role which is to provide the basis for reliable estimates of change of consumer surplus. The latter is a primary input to the social cost benefit analyses that have to be prepared for all major policy measures relating to infrastructure, traffic supply, and other transport policy measures.

In the project we have argued, based on theoretical considerations as well as practical model runs, that the calculation of consumer surplus change should as far as possible be compatible with the demand model. Failure to do this may result in erroneous estimates of consumer surplus change. We have identified this type of compatibility problems for the two model systems we have studied in detail in this project, the standard Sampers implementation and the Vips implementation. Based on these findings it is clear that the short/medium term development also needs to address the issue of calculation of consumer surplus change.

Segmentation

One would wish for better empirical data relating to segmentation. In particular it seems urgent further to accommodate recent findings about time values for various categories and situations.

5.4 Some thoughts about long run development

We have not considered development in the long term perspective in any detail. However, the general framework for modelling of long distance travel that we have outlined in chapter 2 gives a basis for some wider thinking about the potential issues. Of course the models that are now operational are actually based on theoretical and methodological developments that took place many years ago, while the implementations of the current models are influenced by the planning issues and concepts that were topical when these implementations were developed. Many things have changed since then, both on the methodological side – not least regarding numerical methods and computing power – and on the user side, where new requirements have developed gradually.

A long term development effort could and should therefore start out from a situation that in many respects is very different from the earlier situation. In order to take advantage of new development opportunities it is necessary to take an unconstrained position in relation to current model tradition, without of course forgetting lessons learnt from earlier development and practice.

One example of issues that should be subject to further considerations is the **definition** of long distance journeys. In Sweden and Norway the distinction is conventionally put at 100 km. The reason is mainly because national travel surveys contain retrospective questions on trips exceeding 100 km, since they are poorly represented in a trip diary for one day. There is no methodological reason for a limit at 100 km, however.

One of the main problems discussed in the project is that of mode and line choice when the frequency of public transport services is low and hence people use timetables when they decide on what mode and line to use for a journey. This is a general problem that in fact depends more on the quality of public transport services than on distance. Even in generally dense urban areas these problems emerge in the outskirts, and are amplified in week-ends and late evenings when frequency of services tend to drop. The implication is that the traditional discrete choice models may get into some problems long before we reach 100 km of trip length.

On the other hand, when distances become long enough, the headway of public transport services nearly becomes irrelevant for mode choice, except in very special circumstances. For a trip of -say- 1000 km, car is out of the question except when you are staying for a longer time at the destination and need the car to get around, or if no other acceptable alternatives are available. Scheduled bus services are also out of the question, as the trip may take two days. The only modes that can compete for regular trips involving a relatively short stay at the destination are air and high speed rail (if available). In cases like this, neither estimated travel cost nor estimated travel time will tell the true story. The schedule delay involved for different modes may be a matter of a day or more and not a fraction of some headway.

In the long term perspective it may be useful to distinguish between an estimation/model developing phase and an implementation phase. It is a great advantage to have very accurate data in the estimation of discrete choice models. For example, we should try to get as close as possible to the attributes of the alternatives actually faced by the travellers in a sample. Different types of measurement errors at the level of individual observations will tend to give biased parameters for important variables. Average values for level-of-service variables "skimmed" from a network might not be sufficiently precise to get unbiased parameters.

On the other hand, when a model is implemented we will normally need to use "skims" from an assignment model. The consequences will then mainly be confined to some types of aggregation bias that might be far less severe than biased parameters from the estimation phase.

Against this background we think that instead of formulating detailed development tasks, the long term development in this field should be formulated as a comprehensive research and development program which can lead to implementation but which also systematically looks to the ways in which innovative developments can be brought to implementation in due course.

Such a development program should openly and without prejudice address the issues that we have discussed in chapter 2, which are summarized in the table below. The modelling issues to deal with in the research program are here defined by 11 different aspects and corresponding alternative approaches to be dealt with. Note that the listing of possible approaches is not assumed to be exhaustive, and should not be seen as a constraint on further work.

MODELLED OBJECT
a) Tour
b) Trip leg

TIMETABLES
a) Only headways
b) Exact timetables
c) Mix of exact timetables and headways

TRAVELLERS´ DEPARTURE/ARRIVAL TIMES
a) Uniform distribution per day
b) Uniform distribution per time slice of a day
c) Distribution profile over the day
d) Choice of departure time considering asymmetric penalty function

TRAVELLER (DECISION UNIT) DEFINITION
a) Party
b) Individual
c) a) and/or b) with additional influences e.g. from employer

SEGMENTATION AND/OR SEPARATION OF BEHAVIOURAL MODELS
Party size
Journey duration
Trip purpose
Socio-economic
Value of time

TRAVELLERS´ PREFERENCES
a) Deterministic Uniform within segment
b) Additional exogenously chosen dimensions of heterogeneity as basis for segmentation

PRICES
a) One price for each mode for each O-D pair per segment
b) Differentiation of prices between lines for each mode and segment
c) Time related price structures combined with b)
d) Differentiation of prices on the same departure

NETWORK AND CHOICE SET
a) Main modes (car, bus, train, air) at national zone level and access/egress modes (regional public transport, taxi, private car and kiss-and-ride) at regional level, independently treated.
b) Main modes (according to a) and sub modes)
c) No main mode concept, all modes (various trains, airlines, coaches and cars) integrated but modes are retained as attributes
d) Timetable specification

MODELLING APPROACH

a) Separate choice models for line choice (network model, optimal strategy approach) within a mode and between main modes (logit model)
b) Route choice model (e.g. Random Departure Times approach or other approaches) for choice of line and modes (Combinations of commuter trains, regional buses, coaches, various types of trains, various airlines, car)

DESTINATION CHOICE

a) Based on logsum from a logit model that is correctly specified and if necessary nested in a suitable way
b) Based on calculation of generalized cost based on RDT
c) Based on calculation of generalised cost based on other assignment algorithm that is compatible with the behavioural requirements for long distance travel

CONSUMER SURPLUS

a) Separate calculation per mode, for existing/remaining travellers and for new/lost travellers (rule-of-the half). Combinations of modes are disregarded
b) Regarding all lines and modes simultaneously with a composite measure for travel time components, price and GC
c) Logsum used as an estimate of CS

References

Algers, S., Mattsson, L-G., Rydergren, C., and Östlund, B. (2009) Sampers – experience and long and short term development possibilities (in Swedish), Linköping University and Royal Institute of Technology, 2009

Bates, J. (2004). Comments to our basic report. Unpublished working paper.

Bates, J. (2009). Notes on the RDT Model.

Ben-Akiva, M., and Lerman, S (1985). Discrete Choice Analysis, Theory and Application to Travel Demand. MIT Press, Cambridge, Massachusetts.

Beser Hugosson, M. (2003) Issues in estimation and Application of Long distance Travel Demand Models. Doctoral dissertation, ISBN 91-7323-0044-8, Royal Institute of Technology (KTH), Stockholm Sweden 2003.

Daly, A. (2004). Comments to earlier work. Unpublished working paper.

Daly, A. (2010 Summary of points on Public Transport Models. Unpublished working paper

Hasselström, D. (1981). Public transportation planning – a mathematical programming approach. Ph.D. thesis, Department of Business Administration, University of Gothenburg, Sweden.

Jansson, K., Ridderstolpe, B. (1992). A Method for the Route-Choice Problem in Public Transport Systems. Transportation Science, vol. 26, no. 3.

Jansson, K., Lang, H. and Mattsson, D. (2008). Optimal Economic Interventions in Scheduled Public Transport. Research in Transportation Economics, Transit Economics. Elsevier.

Jansson, K., Lang, H., Mattsson, D, Mortazavi, R. (2008). Models For Public Transport Demand and Benefit Assessments. In Hensher, D.A. and Button, K (eds.), Transport Modelling Handbooks in Transport, Second edition. Vol. 1, Elsevier,

Jansson, K et al (2010) Descriptive and theory report. Royal Institute of Technology (KTH), 2010.

Larsen, O. I. and Sunde, Ø. (2008). Waiting time and the role and value of information in scheduled transport. Research in Transportation Economics, Transit Economics. Elsevier

Larsen, O. I. (2009a). Estimating market shares and consumer benefits with random departure time, version 2. Unpublished working paper.

Larsen, O. I. (2009b). Assignment vs. discrete choice models - some unsolved issues. Unpublished working paper.

Louviere, J.J., Hensher, D.A. and Swait, J. (2000) Stated Choice Methods and Analysis. Cambridge UniversityPress, Cambridge.

Lundberg, A I. (2010). Konkurrens och samverkan mellan tåg och flyg (Competition and cooperation between rail and air), Royal Institute of Technology (KTH), Stockholm.

Preston, J. (2004). Comments to our basic report. Unpublished working paper.

Spiess, H. and Florian, M. (1989), Optimal Strategies: A New Assignment Model for Transit Networks, Transportation Research B, Vol. 23B, No.2, 83-102.

Train, K. (2008). EM Algorithms for Nonparametric Estimation of Mixing Distributions, Journal of Choice Modelling, 1(1).

Train, K. (2009), Discrete Choice Methods with Simulation, 2nd edition, Cambridge university press.

Transek (2004) Utveckling av Sampers 2.1. (Development of Sampers 2.1).

Transek (2004) Sampers Technical Documentation, Long Distance Trips,
Sampers version 2.1 – Kalibrering av modellen för långväga resor, Transek report 2005:12 (in Swedish).

Tseng Y-Y, Verhoef E T. (2008), Value of time by time of day: A stated-preference study, Transportation Research Part B 42 607–618.

Appendix 1 Alternative ways to deal with heterogeneity in the context of modelling long distance mode and route choice – combining RDT and discrete choice (DC) models

Basic principles

By Odd I Larsen; Möreforskning Molde AS

An RDT – model can handle the schedule delay aspect related to scheduled traffic in an approximately correct way. The concern with deviations from preferred departure and/or arrival time is also relevant for the choice of mode. Travellers might for various reasons also have mode-specific preference that follows some statistical distribution across individuals and modes. This is the concern of discrete choice (DC) - models, which - at least in traditional applications – are unable to handle the schedule delay issue in a correct way.

A more correct model might be an RDT-model with the discrete choice aspect included for modes. Another option is a DC-model with the schedule delay aspect of route choice included in a correct way. While the two approaches should give approximately the same results, at present it seems that the first option is the most flexible and is easiest to implement.

It can be proved rigorously that most continuous statistical density functions can be approximated very accurately by a discrete five-point distribution with appropriate weights for each point. We write the (observable) generalized cost of an alternative by:

$$G_{m,i} = G^*_{m,i} + H_{m,i}v_{m,i} \tag{1}$$

Where m denote mode and I an alternative within a mode, i.e. a route. $G^*_{m,i}$ includes costs in terms of monetary cost, in-vehicle time etc, all normalized to the mean value of minute deviation from preferred arrival and/or departure time. $H_{m,i}$ is the headway and $v_{m,i}$ is a random variable on the unit interval. The product is a fraction of the headway. An RDT-model will assign trips on modes and routes for the integral over all values of $v_{m,i}$ across all modes and routes.

A discrete choice assumes preferences for modes. Thus for each alternative within a mode we will have an additional "cost" ε_m. This cost which can be both positive and negative can be considered as a random variable with a continuous distribution over modes and travellers. If this distribution is approximated by a 5-point distribution we have the number of combinations of points and modes equal to 5^modes. We can make an assignment for each combination, which in the case of 3 modes will come to 125 assignments.

Let the 5 data points be $d_1,\ldots,d_5$ and the corresponding weights be $w_1,\ldots,w_5$.

For a particular assignment with points i,j,k respectively for the three modes respectively, we add the term $\sigma\cdot d_i$ to all fixed costs for lines belonging to mode 1, $\sigma\cdot d_j$ to lines belonging to mode 2 and $\sigma\cdot d_k$ to lines belonging to mode 3. σ is the standard deviation assumed for the distribution.

We then make an RDT-assignment with these adjusted costs and weigh the results by

$W_{ijk} = w_i \cdot w_j \cdot w_k$. $\sum_{i,j,k} W_{ijk} = 1$ and we can thus add up the sum of weighted assignments to get the results for a combined RDT/DC – model. In practice, the necessary number of assignments is (5^modes – 5) and for symmetric distributions even less.

Conceptually this is straightforward if we know the distribution and the standard deviation to apply. However, for a full implementation also involving multimodal trips, it becomes somewhat more involved and tedious to implement if the same mechanism shall be applied at every possible transfer point.

More a "main mode" approach with appropriate modelling of access egress for the main mode, it is difficult to see any big problems except for computational time and resources.

Mathematics for approximation of distribution

By Harald Lang *Royal Institute of Technology (KTH), Stockholm*

Theorem:

Let X be a random variable whose distribution is symmetric about zero, (i.e., X and $-X$ have the same distribution) and for which $E[|X|^9] < \infty$. Then there are positive numbers w_0, w_1, w_2, a_1, a_2 such that for any polynomial $p(x)$ of degree at most 9, it holds that

$$E[p(X)] = w_2 p(-a_2) + w_1 p(-a_1) + w_0 p(0) + w_1 p(a_1) + w_2 p(a_2)$$

The numbers w_0, w_1, w_2, a_1, a_2 can be computed like this:

First solve the second degree equation in r

(1) $$r^2(c_2c_6 - c_4^2) + r(c_4c_6 - c_2c_8) + (c_4c_8 - c_6^2) = 0 \quad \text{where} \quad c_n = E[X^n].$$

This equation has two positive roots, r_1 and r_2. The numbers are now given by

(2) $$a_1 = \sqrt{r_1},\ a_2 = \sqrt{r_2},\ w_1 = \frac{c_2r_2 - c_4}{2r_1(r_2 - r_1)},\ w_2 = \frac{c_4 - r_1c_2}{2r_2(r_2 - r_1)},\ w_0 = 1 - 2w_1 - 2w_2.$$

Proof:

It is obviously sufficient that

$$w_2(-a_2)^n + w_1(-a_1)^n + 0^n + w_1a_1^n + w_2a_2^n = c_n = E[X^n] \quad \text{for} \quad n = 0, 1, \ldots, 9.$$

Because of the symmetry, this relation holds for all odd values of n, hence it is sufficient that

(3) $$w_0 + 2w_1 + 2w_2 = 1 \quad \text{and}$$

(4) $$2w_1a_1^n + 2w_2a_2^n = c_n \quad \text{for } n = 2, 4, 6, 8.$$

It is trivial to check that

$$2\text{w}_1r_1 + 2\text{w}_2r_2 = c_2 \quad \text{and} \quad 2\text{w}_1r_1^2 + 2\text{w}_2r_2^2 = c_4$$

hold for any numbers r_1 and r_2 when w_1 and w_2 are defined as in (2). It is also easy, albeit very tedious, to check that $2\text{w}_1r_1^3 + 2\text{w}_2r_2^3 = c_6$ and $2\text{w}_1r_1^4 + 2\text{w}_2r_2^4 = c_8$ if r_1 and r_2 are defined such that

(5) $$r_1 + r_2 = \frac{c_2c_8 - c_4c_6}{c_2c_6 - c_4^2} \text{ and } r_1r_2 = \frac{c_4c_8 - c_6^2}{c_2c_6 - c_4^2}.$$

Hence, if we find two positive real numbers r_1 and r_2 such that (5) holds, then (2) is a solution to (4). It is shown in the lemma 1 below, that the coefficient for r^2 and the constant term in (1) are positive, whereas the coefficient for r is negative. Hence, the two roots r_1 and r_2 of (1) satisfy (5) and they both have positive real parts. It is shown in the lemma 2 below that they are indeed real. This proves the theorem. *Q.E.D.*

Lemma 1:

For any probability distribution it holds that

$$E[X^4]^2 \le E[X^2]E[X^6]$$

$$E[X^6]^2 \le E[X^4]E[X^8]$$

$$E[X^4]E[X^6] \le E[X^2]E[X^8]$$

Proof:

The first two inequalities follow readily from Cauchy-Schwartz' inequality. The third follows from Hölder's inequality applied twice:

$$E[X^4] = E[X^{4/3}X^{8/3}] = E[X^2]^{2/3}E[X^8]^{1/3}$$

$$E[X^6] = E[X^{2/3}X^{16/3}] = E[X^2]^{1/3}E[X^8]^{2/3}$$

Multiplying these together gives the third inequality. *Q.E.D.*

Lemma 2:

The roots r_1 and r_2 are real.

Proof:

We need to prove that the expression under the square-root sign in the expressions for the roots r_1 and r_2 is positive. If we minimise that expression w r t c_8 we get

$$c_2^2c_8^2 - 6c_2c_4c_6c_8 + 4c_2c_6^3 + 4c_4^3c_8 - 3c_4^2c_6^2 \ge \frac{4}{c_2^2}(c_2c_6 - c_4^2)^3 \ge 0$$

The last inequality follows from lemma 1. *Q.E.D.*

Example: the normal distribution

For the normal (0,1)-distribution, we have $c_2 = 1, c_4 = 3, c_6 = 15, c_8 = 105$, which gives

$$a_1 = \sqrt{5 - \sqrt{10}},\ a_2 = \sqrt{5 + \sqrt{10}},\ w_1 = \frac{7}{60} + \frac{\sqrt{10}}{30},\ w_2 = \frac{7}{60} - \frac{\sqrt{10}}{30},\ w_0 = \frac{8}{15}$$

A non-symmetric case; the Gumbel distribution

The Gumbel (0,1)-distribution is not symmetric, so the above procedure does not work. One way to obtain a five-point discrete approximation to the Gumbel distribution is to note that if X is Gumbel (0,1), then

$$X = -\ln(-\ln(\Phi(Y)))$$

where Y is a normal (0,1)-random variable and $\Phi(y)$ the cumulative normal (0,1) probability function. Hence, we may use the weights of the normal (0,1)-distribution (see above) and the locations

$-\ln(-\ln(\Phi(-a_2)))$, $-\ln(-\ln(\Phi(-a_1)))$, $-\ln(-\ln(\Phi(0)))$, $-\ln(-\ln(\Phi(a_1)))$, $-\ln(-\ln(\Phi(a_2)))$

(with weights w_2, w_1, w_0, w_1, w_2). In this case, the equality of the Theorem holds not for polynomials, but for all functions

$$p(x) = q(\Phi^{-1}(\exp(-\exp(-x))))$$

where $q(y)$ is any polynomial of degree at most 9.

As a numerical example, let us compute $E[X]$, $E[X^2]$ and $E[e^{-X}]$ for the Gumbel distribution using the approximate formula and compare with the true exact values. We get

$$E[X] = 0.577215 \approx 0.577223, \quad E[X^2] = 1.9781 \approx 1.9778 \quad \text{and} \quad E[e^{-X}] = 1 \approx 0.9999966$$

where the first values are the true (rounded) values, and the second the values obtained with the approximate formula.

Appendix 2. Calculation of consumer surplus in association with using different algorithms in the demand model

By Harald Lang, Royal Institute of Technology (KTH), Stockholm

Assume that we have two goods, *A* and *B*, which are substitutes (two brands of tooth paste, for example) Let their prices be p and q, respectively. Now, if we lower the price q on *B*, then there will be a gain in consumer surplus. The gain in consumer surplus can be reasonably well calculated as follows: Let y_0 and y_1 be the demand for good *B* before and after the price change, respectively, and let q_0 and q_1 be the prices on *B* before and after the price change. Then the increase in consumer surplus due to the price cut is to a good approximation

$$\Delta CS = \frac{1}{2}(y_1 + y_0)(q_1 - q_0)$$

Note that since the two goods are substitutes, the price cut of the price for *B* will result in a decrease in demand for good *A*. However, this reduction in demand for good *A* does not enter the calculation of the change in consumer surplus.

Now consider the situation when *A* and *B* are two bus lines, serving the same origin–destination pair, but depart from different stations. Assume their headways are h_1 and h_2, respectively. These headways can be seen as "generalised prices". Assume now that we change the headway h_2 from h_2^0 to h_2^1. It might then be tempting to believe that, in analogy with the situation described above, for a proper assessment of the gain (or loss) in consumer surplus due to the change in headway for bus *B*, it suffices to consider the change in demand and generalised price for bus *B*. Whether this is true depends on the behavioural of the travellers. If we assume that travellers decide upon which bus to choose based on headways, but not on actual time schedules, then this is true.

The problem arises if travellers decide on which bus to choose based on time schedules. The reason can be explained as follows. If travellers decide upon which bus to choose based on headways, then, on average, their waiting time for the bus is half the headway, i.e., $h_1/2$ and $h_2/2$ respectively. These *schedule delays* may be seen as generalised prices. Note that if we change h_2, the generalised price $h_1/2$ is unaltered.

However, if travellers decide on which bus to choose based on time schedules, the mean waiting time for, say, travellers with bus *A* will depend not only on the headway h_1,but also on the headway h_2 of bus *B*. Indeed, let us say that travellers will choose whichever bus leaves first, and choose station accordingly. If h_2 is longer than h_1, then some travellers with bus *A* will have a schedule delay time close to h_1.But if we now reduce the headway h_2 below h_1,then nobody travelling with bus *A* will have a schedule delay time longer than h_2 (since in that case he would have chosen Bus *B*.) Hence, a reduction in headway h_2 may (and will, typically) reduce the "generalised price" for bus *A*. Note again: not only will a change in headway (i.e., "generalised price") h_2 for bus *B* affect the *demand* for bus *A*, but also the (generalised) *price* for bus *A*. This has to be taken into account when the change in consumer surplus is calculated.

The conclusion is that if travellers are assumed to consult time schedules, then a change in headway for one line will in general affect both the demand and mean schedule delay times for all alternative lines, and thus all these changes have to be taken into account when the effect on consumer surplus is calculated. The danger is that this is overlooked, and only the line for which the headway is changed is considered.
If one applies ROH for all alternative lines it will be approximately correct but on the other hand this is an unnecessarily complicated way (you move round in a circle) so in practice I see no use for this. But if one applies ROH for only the line that has been changed, the outcome will of course be completely wrong. And there is a risk that this is overlooked.

A more detailed discussion on consumer surplus calculation is found in section 9 in the Descriptive and theory report.

Appendix 3 Contents and authors of Descriptive and theory report

Contents

Contributions by

Kjell Jansson: Sections 2-6, 7.2- 7.4, 8.5, 9.2, 9.3, 10-13, appendices 2-6, 8

Staffan Algers: Section 7.1, appendix 7

Harald Lang: Section 9.1.

Odd Larsen and Reza Mortazavi: Sections 8.1-8.4

John Bates: Appendix 1

Andrew Daly: Appendix 9

Division of Traffic and Logistics

The Division of Traffic and Logistics conducts fundamental and applied research that advances the state of the art and state of practice in the planning, design, operations and management of transport systems and facilities. Our research is multidisciplinary by nature and always seen in a multimodal context. KTH

The Division work in four main areas: road traffic, rail transportation, public transport and logistics. Quantitative methodologies (including simulation, analytical approaches and modeling), qualitative analysis and advances in computing, information and sensor technologies are common approaches used in our research and taught in our courses.

www.kth.se/abe