Airport and Station Accessibility as A Determinant of Mode Choice

by

Reinhard Clever

MBA (University of Miami) 1977
MS (University of California, Berkeley) 1996

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering – Civil Engineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in Charge:

Professor Mark M. Hansen, Chair
Professor Samer Madanat
Professor Paul A. Ruud
Professor Phil Spector

Fall 2006

Airport and Station Accessibility as A Determinant of Mode Choice

by

Reinhard Clever

Abstract

Airport and Station Accessibility as A Determinant of Mode Choice

by

Reinhard Clever

Doctor of Philosophy in Engineering – Civil Engineering

University of California, Berkeley

Professor Mark M. Hansen, Chair

This dissertation examines whether the positive experience with High Speed Rail (HSR) in Europe and Japan is likely to be transferable to North America without significant modifications. It studies how important the access advantage was to the success of the Japanese system.

A utility impact analysis revealed that access and egress related variables account for 26 - 43% of the total absolute utility of typical air and rail travel. Terminal fixed effects were found to be about four times more important for HSR than for air. They capture not only accessibility attributes, but also the attractiveness of the terminal's location.

If the utility of HSR travel in the United States would depend on the magnetism of large downtown stations as much as it appears to in Japan, innovative solutions must be found to replicate this effect. Two solutions are outlined in the first and last chapters. They require a paradigm shift to convergence, seeing urban, regional, and high speed ground transportation as one coherent system.

Almost 50 000 air and rail observations of the 1995 Intercity Travel Survey could be used to estimate separate airport pair and HSR station pair choice models. The HSR choice set consisted of 1 260 station pairs and had to be randomized. Weights compensate for heteroskedasticity and extend the sample to a typical Fall day with daily expansion factors.

The small air and large HSR coefficients estimated with nested logit suggested that a new airport would mostly draw passengers away from other airport pairs, while adding HSR stations would have a high impact on other modes.

Methodologically, the research found that a log plus linear specification for access and egress distance is often called for. Linear, quadratic, and cubic functions underestimated the disutility of feeder distance at the median and mean by up to two thirds.

When randomizing choice sets the researcher should only draw conclusions from a *set* of models, each using a different number of alternatives, to detect converging values and standard errors likely to be incorrect. Using this approach, strong evidence of a threshold effect for access and egress started to emerge.

Dedication

To my parents, Margret and Peter Clever

Acknowledgments

The research on which this dissertation is based was financed in part by a scholarship from the University of California Transportation Center (UCTC).

First of all, I would like to thank my advisor, Professor Mark Hansen, for the invaluable guidance he has given me throughout my research. This did not only include advice of a statistical or econometric nature, but also valuable insights into the competitive interplay between air and high speed rail. I also want to thank my other committee members: Professor Samer Madanat, who, because of his excellent relationship with Professor Taka Morikawa of the University of Nagoya, was able to provide the dataset for this research. Having spent the greater part of one year looking for appropriate survey data, I can truly say, that without the high quality 1995 Japanese Intercity Travel Survey, I would not have been able to write a dissertation on even a similar topic; Professor Paul Ruud, for the help at a very critical and difficult juncture of this research; and Professor Phil Spector, who has probably taught me everything I know about statistical computing.

And, I do want to thank Dr. Andreas Oetting and Dr. Klaus Walther for having instilled in me a healthy suspicion of linear specifications.

Many people outside of the academic arena have provided great support for my dissertation research. I am thinking of all the help I received from systems specialists Greg Merritt, Thompson Won, and especially Jon Forrest. I want to thank my ex-wife, Jeanine, for all her financial support, and two professional counselors from Disabled Student Services at the University of California, Berkeley: Dr. Helen Beck and Connie Chiba. Your support was essential to my academic success!

Last but not least, I would like to thank my two spiritual advisors, Father Brian Joyce, the pastor of Christ the King Parish in Pleasant Hill, and Pastor David Baker of New Life Ministries in Hayward, who I also consider to be my best friend. Thank you for not only helping me to stay close to God during the ups and downs of this endeavor, but managing to bring me even much closer to Him. Without my religious and spiritual beliefs this dissertation would not have been completed.

Contents at a Glance

Detailed Table of Contents

List of Figures

List of Tables

List of Maps

List of Definitions

Abbreviations, Notation and Japanese Terminology

Abbreviations

ANA	All Nippon Airways
HSR	High Speed Rail
HSGT	High Speed Ground Transportation
BART	(San Francisco) Bay Area Rapid Transit
JNR	Japanese National Railways (until 1987)
JR East	East Japan Railway (since 1987)
JR Central	Central Japan Railway (since 1987)
JR West	West Japan Railway (since 1987)
GLM	Generalized Linear Models
MNL	Multinomial Logit
NL	Nested Logit
IIA	Independence of Irrelevant Alternatives
DEC	Daily Extension Coefficient
O&D	Origin and Destination
LOS	Level of Service
OVT	Out of Vehicle Time

Mathematical Notation

V_{in}	Systematic utility of alternative i for individual n
P_{in} or π_{in}	Choice probability of alternative i for individual n
x_{ink}	Attribute k of alternative i for individual n
$\mathbf{X}_{in}$	Attributes of the K explanatory variables of alternative i for individual n
$\boldsymbol{\beta}_i'$ or $\boldsymbol{\beta}_i$	Parameter vector of alternative i (column vector)
$\mathbf{S}_n$	Characteristics of individual n
$\mathbf{Z}_{in}$	Attributes of alternative i for individual n

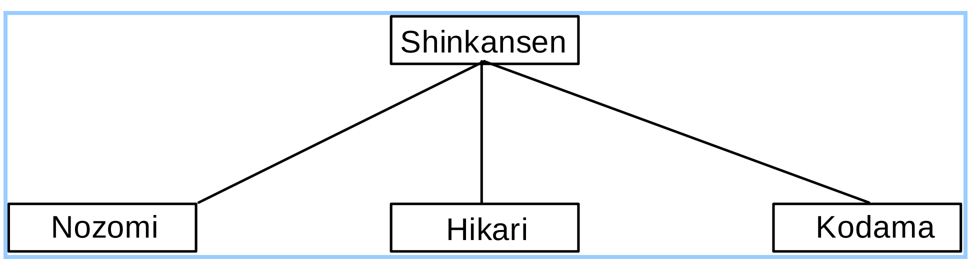
Alternative i can take on the values 1 .. 82 for airport pair choice models, 1 .. 1260 for HSR station pair choice models, and A or R for the upper level mode choice model denoting air and HSR respectively.

Model Designation

The indicator for airport pair choice models is 'A' and for HSR station pair models 'R.' Airport pair choice models are numbered sequentially, while the number in HSR station pair models points to the size of the choice set, the number of alternatives used. The suffix "-2" in the model designation indicates that a correction for heteroskedasticity was made. The "-2" suffix is often implicit and therefore suppressed. Capital letters at the end of the model number as in R 300A or R 300A-2 denote minor variations in the model specification.

Japanese Terminology

High Speed Rail *Shinkansen* service gives travelers a choice of three train categories. Local and regional service with stops every 30-40 km is called *Kodama* or "Echo" (mnemonic: ko-damned slow). There are two long-distance services: *Hikari* ("Lightning") and the extra fast, premium service *Nozomi* or "Hope" (mnemonic: zoom ahead).



Tokyo – Osaka = "*Tokaido*" corridor

Osaka – Fukuoka = *Sanyo* ("Mountain Sunshine") corridor (Map 1 on page 64)

1 Motivation and Introduction

1.1 The Internal Distribution Advantage of HSR

Every trip by air or rail has three components: 1) access to the airport or rail station, 2) the trip from airport to airport or station to station called "linehaul", and 3) egress from the airport or the rail station. This research examines how access and egress affect linehaul market share.

Our study will shed more light on a particular phenomenon of intercity markets in the range of 300 – 700 km that feature direct competition between air and rail: surprisingly high market shares for rail in spite of a considerably longer linehaul time (Table 1).

Table 1 - The Phenomenon

Market	Linehaul Time		HSR Market Share
	Air	HSR	
London - Paris	1h	2:45 h	67%
New York - Washington	1h	2:45 h	> 40%
Tokyo - Osaka	1h	2:30 h	80%

Sources: (Hughes, 2004; Jackson, 2004a)

A price advantage is not the reason for this phenomenon. In fact, for the Tokyo – Osaka line the opposite is true. The profitable Tokaido Shinkansen has to subsidize the construction cost of newer Shinkansen lines, built for political reasons into less populated areas, which increases its capital cost by over 30%. This forces JR Central to set fare levels higher than the inherent cost of providing the service. In the face of cut-throat airline competition, Yoshiyuki Kasai, president of JR Central, calls this an "economic distortion" (Kasai, 2000).

Research conducted on the mode choice behavior of business travelers between London and Paris indicated that many business travelers prefer the Eurostar over much faster air service because of the long period of uninterrupted time they can use productively. Air travel with the necessary trips to and from the airport is too fragmented and therefore less productive. This illustrates that access and egress play a central role in markets where air competes head on with High Speed Rail (HSR).

The investigation has far reaching implications for the design of High Speed Rail systems in the US. HSR has been able to capture substantial market share in Japan and Europe, where downtown to downtown connections offer a great advantage over the competing air mode. All European and Japanese high speed rail systems had to do was to serve the already existing central stations, and their access advantage over air was handed to them "on a golden platter". Downtown-to-downtown connections do not offer a great access advantage in decentralized widely dispersed US metropolitan areas. In North America, outside of the Northeast Corridor, most business trips do not originate or terminate in the central business district and a much more sophisticated approach is likely to be necessary for HSR to be competitive with air. A simple transfer of the European or Japanese experience to North America will probably not be successful (Kanafani & Youssef, 1993). The "internal distribution" advantage of HSR needs to be carefully studied. Before defining internal distribution advantage, we will first attempt to answer the question of what high speed ground transportation can do that air cannot, or only with great difficulty. These are the competitive advantages that high speed rail may have vis-à-vis air (Clever, 1994):

#1-speed:	Between metropolitan areas up to 200 km, high speed rail is faster or equally as fast as jet service. Examples are: New York – Philadelphia, Paris – Lille, or Cologne – Frankfurt. The fastest connection between the airports of Frankfurt and Cologne will soon be the train. Between metropolitan areas up to 400 km apart, high speed rail is about as fast as commuter aircraft service.
#2-subway:	Inside metropolitan areas, trains can go through a metamorphosis and turn into subway trains, even going underneath city streets, directly serving high density areas with as many stops as may be required by the traveling public.
#3-split up:	Unlike planes, trains can also split up into many pieces once they reach a metropolitan area, each piece serving a different commuter rail corridor.
#4-quick stop:	High speed trains can serve rural areas and mid-size cities more efficiently than airplanes. The penalty for an additional stop on a high speed segment is only a little more than 5 min. In contrast, the equivalent penalty for air is about half an hour.
#5-fully automatic:	Rail is also the easiest mode to fully automate. Train controls have to essentially concern themselves only with one-dimensional operation. On the other hand automobiles need 2-dimensional and planes 3-dimensional controls. Fully automatic operation enables service to be extremely frequent with very short trains the way it has already been realized with people movers.

Advantages #2-subway and #4-quick stop illustrate the adaptability of this mode to frequent stops. Competitive advantages #1-speed and #4-quick stop are very well understood and taken into consideration in every high speed rail study.

- We define the **Internal Distribution Advantage of HSR** as competitive advantages # 2-subway and #3-split-up.

The internal distribution advantage of HSR is generally not considered in present HSR studies.

A metaphor will illustrate the main point of this chapter and of this dissertation: dogs are both stronger and faster than cats. On a wide open field dogs could dominate cats completely. However in reality dogs do not seem to be able to catch cats very frequently. That is because cats can do things that dogs cannot: climb trees, squeeze underneath fences, jump onto rooftops, etc. Most HSR planning in the United States proposes to have dogs and cats compete on a wide open field.

In order to be competitive with air, HSR may have to serve several strategic points in each metro area. Due to its by and large longer linehaul times HSR needs to get much closer to travelers' origins and destinations than air. HSR stations have to be located almost in the center of those activity zones within a metropolitan area that generate the most traffic in a particular corridor. This dissertation will contribute to the understanding of the internal distribution advantage of high speed ground transportation.

Research on travelers' mode choice behavior almost exclusively concentrates on urban trip making. Only a very small portion looks at intercity trips. Within intercity mode choice models, investigation of access parameters is usually restricted to estimating coefficients of access time and access cost. In addition, there seems to be little agreement on the magnitude of access coefficients in intercity travel (please see Section 2.1 below). One reason may be that access parameters have been of little interest in almost all of the modern studies. The sole focus of this dissertation is an in-depth analysis of access related variables. Only after having gained a clear understanding on how access and egress affects mode choice behavior can we proceed to the next step proposed as recommended future research: deduce more information on the optimal location of HSR terminals. This is not possible with present mode choice studies (Schneider, 1993). After being able to give decision makers more information on the best routing of a high speed ground transportation system within a metropolitan region can we say that we are beginning to understand the internal distribution advantage of HSR.

A good example of the present state of the practice is the passenger forecast for the proposed high speed rail system in California (California Intercity High Speed Rail Commission & Charles River Associates, 1996). While very sophisticated survey and forecasting techniques are being employed for the analysis, little attention is being given to the optimal location of high speed rail terminals within the region. Suburban stations will be located just where the high speed rail corridor happens to enter the urban area. If the Altamount Pass route is chosen for access into the San Francisco Bay Area, there will be a high speed rail terminal in Pleasanton. If another pass is chosen, only the suburban locations that happen to lie on that path will be served directly by high speed rail.

As of this writing, an alignment via the Pacheco Pass to San Jose is the preferred corridor into the San Francisco Bay Area. That means there will be suburban stations in

Gilroy and along the Peninsular Corridor. In case a branch line to Oakland is built, some cities in western Alameda County could also be served. However, the fastest growing areas in the region are Eastern Alameda County, Contra Costa County, and the I-80 corridor. They are also poorly served by the present airport system and would therefore show some of the best ridership potential for a HSR link between Northern and Southern California. Unfortunately, this area will be completely bypassed by the proposed route alignment.

Airport access surveys conducted by the Metropolitan Transportation Commission reveal that there is a very high concentration of air travelers in high income areas like Walnut Creek. It is simply not believable that air travelers to Southern California would choose rail, when the air trip takes one hour and the rail trip takes three hours and the access time to Oakland International Airport is about the same as the travel time to the high speed rail station in downtown Oakland.

Since high speed rail will probably not get a second chance in the United States if the first all newly constructed system does not meet expectation, this is reason for grave concern.

Korea may serve as a warning. The initial patronage of Korea Trains eXpress (KTX), which started operating in April 2004, reached 50% of forecast. Forty (40) percent of KTX riders considered "Poor accessibility of the KTX stations" the main disadvantage versus competing modes. Other reasons trailed far behind: expensive price (19%), scheduling problems (15%), and other (26%) (Lee & Chang, 2006).

1.2 Reasons for Scant Research on This Topic

A combination of two powerful impediments to serious research, lack of interest and paucity of useful data is most likely responsible for the lack of analysis on the impact of access related variables on mode choice behavior. Lack of interest originated from "lack of choice" on the part of European railways. Central stations were built in the 19th century close to the cities' urban core, and their locations cannot be changed, except under very special circumstances. Since the entire urban transportation system is focused on these central stations, there is also little need to relocate them. Consequently, access parameters were looked upon as constants.

It may be due to the lack of interest in and the resulting lack of knowledge of access and egress related issues that we can see examples where HSR systems do not take travelers to the center of urban activity even though the infrastructure is already available. Paris to Brussels Thalys trains terminate at Brussels South Station, and do not continue on to Brussels Central and Brussels North. When Berlin's cross-city-line was rebuilt after the fall of The Wall, Berlin – Alexanderplatz station was not upgraded to handle long-distance trains, even though it is at a more central location and far more accessible by public transportation than the traditional East Terminal (Ostbahnhof).

Lack of interest leads to paucity of useful data. However, besides lack of interest, there are other formidable impediments to the availability of useful data, not the least of which is a concern for the protection of privacy. In order to be able to draw conclusions on how quickly rail market share drops with increasing distance to the station one needs detailed geographically precise data on trip origins and destinations. This data tends to be either government-owned and confidential (non-public use files of the 1995 American

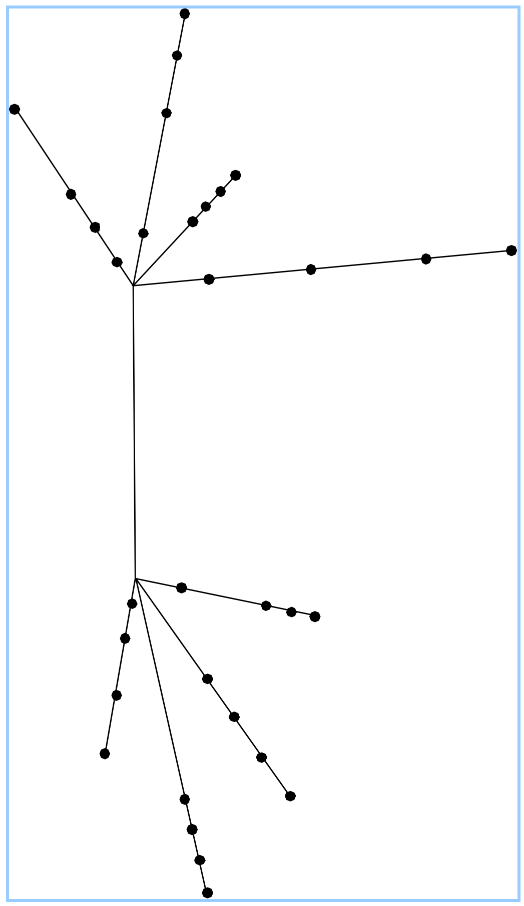
Travel Survey, Enquête Française, 1995 Japanese Intercity Travel Survey) or proprietary (Amtrak's KPMG Northeast Corridor telephone survey data, German Rail's Mobility 97, Fernverkehrspanel, Level of Service Data File for the Japanese Intercity Travel Survey, etc.). Even if some government-owned and confidential files were publicly available, they may still not be useful because they do not contain enough observations per cell (Enquête Française, German Rail's Mobility 97). In order to do this research one needs access to proprietary, in addition to government-owned and confidential databases.

To complicate matters even further, in order to gauge how access/egress variables and linehaul factors jointly affect intercity travel mode choice, one needs a data set with meaningful variation in these factors. There are only two countries with extensive enough high speed rail systems that would allow us to compare many different corridors: France and Japan. As mentioned earlier, the Enquête Française does not contain enough observations per cell to draw meaningful conclusions, leaving only Japan. For this research both the government-owned and confidential 1995 Japanese Intercity Travel Survey and the proprietary Level of Service Data File for the Japanese Intercity Travel Survey have been graciously made available to us.

1.3 Impact on City and Regional Planning

This research has potentially important implications for land-use planning. Consider for example a high speed rail line linking Northern and Southern California including several access branches at both ends (Figure 1). (The last chapter of this dissertation looks at the technical feasibility of such a system.)

Figure 1 - California Corridor HSR System with Access Branches



If such a system were to be built (sharing right-of-way or tracks with existing commuter rail systems on the access branches wherever possible), opportunities for complimentary land-use development could be substantial. The 16 access points depicted in **Error! Reference source not found.** for both Northern and Southern California (randomly chosen just for illustration purposes) would in a sense become mini-airports without the enormous space requirements of an airport. If transportation planning were to be properly coordinated with land use planning, these access centers could become the seed for a more auto-independent urban form in the Western United States.

Research has shown that closeness to a transit station positively impacts office rental prices (Cervero & Landis, 1993), making the real estate more valuable. Commercial land prices near LRT stations are typically valued 24% and near commuter rail stations 120% higher than comparable parcels in other locations (Cervero & Duncan, 2002). It has also been shown that the closeness of an office building to a transit station, together with employment density and two other variables explained 92% of the variation in rail modal split for commuters (Cervero, 1994). If these transit stations were also high speed rail intercity stations, the synergistic effects between transportation and land-use could reasonably be expected to increase.

Transit villages, which have been suggested some time ago (Bernick & Cervero, 1997) are presently being studied for a number of Northern California BART stations (DeFao, 2000). Combining some of the BART stations with high speed rail access stations would make the transit village concept considerably more attractive to developers and give it an additional boost.

1.4 Dissertation Outline

This outline should prove helpful to the reader who would like to scan the thesis in order to quickly find information relevant to his or her particular interest and research.

1. Motivation and Introduction
2. Literature Review
3. Contributors to Mode Choice
4. Model Definitions
5. Data Sources and Sampling Issues
6. Model Results
7. Conclusions and Recommended Future Research

Chapters 1 and 7 bracket the dissertation like bookends. Both deal with the over-arching vision motivating this project.

The first section of the next chapter briefly summarizes in tabular form the main results of the literature review and places our investigation in the proper context. The remainder of the chapter examines how specific research questions relating to access and egress were answered in the last 37 years, and offers the reader a brief overview of the evolution of discrete choice models as applied to intercity transportation.

Referring back to the literature review where appropriate, chapter 3 examines all major contributors to mode choice. Particular emphasis is given to access and egress and a simple model illustrates how terminal access can be the deciding factor when choosing between air and HSR.

Chapters 4 and 5 describe the research methodology. While chapter 3 analyzes the single components of the utility function, chapter 4 combines them to define the different models used in this dissertation.

Chapter 5 describes how the data set was assembled from many different sources. It elucidates the various approaches to different sampling issues so that all underlying model assumptions are met. Which weights are used in the analysis and why is discussed in detail in this chapter.

The focus of chapter 6 is on the practical results obtained from this investigation. Its methodological contribution is the discussion in Section 6.1 on the correct functional form for access and egress distance.

Chapter 7 gives a brief summary of the main conclusions and suggests the next steps to attain a better understanding of the internal distribution advantage of HSR.

2 Literature Review

2.1 Summary

The focus of this chapter is on archival literature, which reflect the state of the art, not the state of the practice.

Studying the summary table on the following pages, two points might become evident to the reader:

- The last in-depth analysis on how access affects linehaul market share was done almost 40 years ago, when none of the modern demand analysis tools was yet available (Beimborn, 1968).
- There is little agreement on the relationship between access and egress time on one hand and linehaul time on the other.

Most authors agree that access time is *more* onerous than linehaul time. Their estimates range from 1.5 times as onerous to 10 times as onerous. (Hensher, 1997) assumed access and egress time to be approximately valued at 1.5 times in-vehicle time and considered this “a generally accepted ratio.” (Stopher et al., 1999) stated that this value ranges in the literature between a low of 2 and a high of 10, and (Walther, 1991) found the disutility of out-of-vehicle time rising exponentially with increasing length of out-of-vehicle time.

There is also very little research done on how high speed rail station locations affect ridership and market share (Schneider, 1993).

This literature review shows, that while the methodology applied to the analysis of intercity travel demand became more and more sophisticated, the attention to access and egress issues ranged from nonexistent to barely considered for reasons of completeness. No recent study of intercity mode choice focused exclusively on access/egress issues.

The literature review indicates that a detailed study on how accessibility affects market share would be timely.

The remainder of this chapter is organized as follows: The first section after the summary tables (Section 2.2) discusses the last dissertation with a topic very similar to this one. It is followed by a review of publications evaluating determinants of terminal accessibility (Section 2.3). Papers on airport choice and airport access models are included here. Section 2.4 deals with transfer penalties, and Section 2.5 presents a detailed analysis of the treatment of access issues in major intercity mode choice studies. It also provides a nice overview of the evolution of discrete mode choice models relating to intercity passenger transportation.

Methodology	Study	Considerations Given to Access/Egress Issues	Important Results Regarding Access and Egress
	(Beimborn, 1968)	Focus of the Dissertation.	Increase in rail mode share from adding a station at Capitol Beltway in Maryland is equivalent to the rail mode share increase that could be obtained from reducing intercity travel time from 133 to 40 minutes.
Nested Logit	(Morrison & Winston, 1985)	Dummy variable for availability of bus and rail transit at destination.	none
	(Koppelman & Hirsh, 1986)	none	none
	(Koppelman, 1989)	none	none
	(Forinash & Koppelman, 1993)	Estimated ratio of OVT to IVT.	Out-of-vehicle time (OVT) is four times as onerous as in-vehicle time (IVT).
	(Wen & Koppelman, 2001)	Estimated ratio of OVT to IVT.	The ratio was closer to 3:1 for the generalized nested logit (GNL) model.
	(Lyles & Mallick, 1990)	none	none
	(Vrtic & Axhausen, 2003)	Estimated ratio of access time to in-vehicle time.	Ratio was 2.7 in mode choice and 2.9 in joint mode and route choice model.
Stated Preference	(Morikawa, 1989)	Estimated ratio of terminal to linehaul time.	Ratio of terminal time to linehaul time a stable 5:1 in all RP/SP combined models.
	(Morikawa, Ben-Akiva & Yamada, 1991)	Estimated one combined coefficient for access and egress.	RP model: access/egress coefficient double linehaul coefficient. Combined RP/SP model: ratio reversed.
	(Hensher, Brotchie & Gunn, 1989)	Access/egress times and cost held constant for each mode and respondent.	none

Methodology	Study	Considerations Given to Access/Egress Issues	Important Results Regarding Access and Egress
Stated Preference (continued)	(Hensher, 1998)	none	none
	(Widlert, 1998)	none	none
	(Vrtic & Axhausen, 2003)	none	none
Box-Cox Transformations	(Mandel, Gaudry & Rothengatter, 1993)	none	none
Logit Captivity	(McCarthy, 1997)	none	none
Random Effects Logit	(Hensher, 1997)	Access/egress parameters not varied in stated choice design.	To compare his value of time estimates with other studies an assumption was made: "access and egress time are approximately valued at 1.5 times the main mode in-vehicle time – a generally accepted ratio."
Neural Networks	(Xie, 2000)	none	none
Other	(Stopher et al., 1999)	Did literature review to obtain ratio of in-vehicle to access time.	Found ratio in his literature review to be between a low of 2 and a high of 10.
	(Walther, 1991)	Studied the relationship of different components of out-of-vehicle time with respect to in-vehicle time.	Normalizing the disutility of in-vehicle time with a time valuation factor of 1, the disutility of out-of-vehicle time rises exponentially with increasing length of out-of-vehicle time.

Table 2 - Treatment of Access/Egress in Major Intercity Choice Studies

2.2 The Last Dissertation on the Impact of Access upon Intercity Mode Choice

In **1968** a dissertation was published by Edward **Beimborn** at Northwestern University with a title very similar to the one of this dissertation: "An Examination of the Effects of the Access Portion of a Trip upon an Intercity Mode Choice" (Beimborn, 1968). Its major findings were published by the American Society of Civil Engineers (Beimborn, 1969).

Beimborn developed a general model to investigate the degree to which local access or terminal locations affect the choice of an intercity mode. He exercised his model to analyze travel between Washington, DC and Philadelphia.

His model was based on the disutility of travel; travelers were assigned to the mode which minimized the impedance (time and cost) of each trip. Beimborn developed two models, a mode choice model and a travel demand model. His mathematical model for mode choice could be geometrically represented by a family of indifference curves representing the tradeoff a trip maker is willing to make between travel cost and travel time. To determine the demand of travel between the local zones of the two cities, census data was used to find the proportion of high, medium and low income families in each zone. He could then apply the appropriate travel intensity factors to each income group. A new terminal could be added to his model by finding the coordinates of its location, and by assuming access time and access cost equations, in addition to linehaul time and linehaul cost characteristics.

2.3 Determinants of Terminal Accessibility

Noting that it had already been shown by Beimborn that accessibility to intercity linehaul termini has a significant influence on modal choice behavior, **Leake and Underwood (1977)** proposed an intercity terminal access modal choice model. This was undertaken as part of a detailed analysis of the bi-modal choice between air and rail in Great Britain.

The mathematical formulation for their analysis was based on the entropy maximization model of A.G. Wilson. It was used to develop a theoretical trip distribution model similar to the gravity model. It is interesting to note that when the model is applied to modal split analysis it is almost identical to the now very familiar logit formulation. The only differences are that instead of using utility (a positive number), generalized cost (a negative number) was exponentiated, and instead of calibrating a vector β, only a scalar α (a constant term) and a scalar β (coefficient of the generalized cost difference) was solved for.

Access time and cost arrays were developed for each city region using 5 km grid squares. To calculate the generalized cost, a value of time coefficient was adopted for each city region and each trip purpose.

The value of β measured the variation of modal split with increasing generalized cost difference. The higher the value of β, the higher the sensitivity to changes in generalized costs would be. Leake and Underwood found that the values of β were greater for travel to rail termini than to airports and concluded: “This suggests that the values of modal split to and from rail termini are likely to be more susceptible to changing costs and / or times than those for airports."

Leak and Underwood developed an intercity *terminal access mode* choice model, not an airport pair or HSR station pair choice model as we have done in this research. So their findings are not directly comparable with ours. But the similarity is nevertheless noteworthy.

Sobieniak et al. (1979) proposed a disaggregate demand model for the choice of access mode to intercity transportation terminals in Canada's national capital region (Ottawa-Hull and vicinity). For their methodology they referred to a paper by Rassam et al. first published in 1970, then republished in the Highway Research Record in 1971 entitled "The n-Dimensional Logit Model: Development and Application" (Rassam, Ellis & Bennett, 1970, 1971). Note that Rassam et al. published their research on logit methodology even before McFadden (1972), (1974). The paper by Sobieniak et al. clearly reflected the state of the art and not the state of the practice.

Separate models were developed for access to air, rail, and intercity bus terminals, stratified by personal and business trip purposes.

The findings, as the authors pointed out, were neither surprising nor controversial. But they had definite implication on public policy regarding the improvement of public transit access to intercity terminals. "Capital investments to improve transit linehaul times to intercity terminals are unlikely to attract a significant number of new passengers." Access at trip origin, service frequency, and baggage handling were found to be much greater determinants of mode choice. They concluded that shared-ride taxis offered a compromise between the lower cost of transit services and convenience and speed of private automobile and taxi and could become the predominant mode. As we know today, their prediction turned out to be accurate.

These findings are very important, because they emphasize that access time is not a very good metric of accessibility to an intercity terminal. It also means that if several high speed rail terminals in a region competed with one single airport, the airport may have better “accessibility” even though it was further away, if it was better served by shared-ride taxis. Airport shuttles can handle many-to-one origin and destination patterns very well, but they might never be able to reach critical mass with a multitude of high speed rail terminals in a region. This directly contradicts the vision set forth in the last chapter of this dissertation. The most effective way to overcome the disadvantage of worse accessibility by shared-ride taxis would be for high speed rail terminals to offer ample free parking, which is not available at airports.

Lunsford and Gosling (1994) gave a review of the most important airport choice and airport ground access mode choice models to that date. They also noted that travel time does not provide a comprehensive representation of ground access quality at an airport. In his airport choice models **Harvey (1988)** included ground access quality with a variable that represented the expected utility from his ground access mode choice model.

Harvey (1986a) was able to quantify the deterrent to using public transit as an access mode created by the amount of luggage. He found that the implicit price of each additional piece of baggage per person, which is the fare reduction required to offset the disutility of having to handle the additional bag, was \$ 11.17 for non-business travelers, about 2.4 times the implicit price for business travelers.

Harvey (1987) emphasized the importance of stratification between business and non-business travelers. He noted that airport choice models for business travelers had more explanatory power than the models for non-business travelers. A cause for this, he suggested, may be that non-business travelers do not behave as rationally as business travelers. It did not seem likely to him that the same variables could be used to adequately explain the behavior of all types of travelers. Variables regarding the gender of the traveler, the amount of luggage, as well as the effect of income on cost sensitivity are much more important in the non-business model than in the business model. Generally it appeared that variables related to travel time are most important to consider for business travelers, while variables related to cost are more important to non-business travelers.

Two studies found that female travelers show a strong preference for modes that provide escorted door to door service, like being dropped off at the airport, or more secure modes like taxis and limousines (Harvey, 1988; Sobieniak et al., 1979). A third study (Harvey, 1986b) found that considering gender was only important for non-business travelers. These findings suggest that non-business females evaluate the concept of "accessibility" to an intercity terminal completely differently than males. It emphasizes the need to include sex as an important socio-economic variable when modeling the effect of accessibility on linehaul market share.

Lunsford and Gosling (1994) noted that there was very little consistency in the parameter values estimated across the different models. There may be many explanations for this. In different parts of their working paper the authors noted that the nested logit model had been used in relatively few studies to date. Since the MNL assumption of Independence from "Irrelevant" Alternatives cannot be maintained across all individual and public transportation airport access modes, we would expect biased parameter estimates from many of the models for this reason alone. Important explanatory variables like fare levels were also omitted in airport choice models, since information on that is often not included in air passenger surveys. Harvey (1988) also identifies improper coding of the survey results into a computer file as a problem with all of the surveys he has worked with to develop choice models.

The lack of consistency in the parameter values across different models indicate how difficult it would be to develop access related models that are transferable across regions.

Lunsford and Gosling suggested the following improvements for future airport choice and airport ground access mode choice models:

Access/Egress Differentiation. Since survey data is usually obtained from enplaning passengers, little information is known about the trip *from* the airport. **Ellis (1974)** developed the only model that took trips from the airport into account (Ellis, Bennett & Rassam, 1974). He found that the difference in parameters between trips to and from the airport is comparable to the difference in parameters between business and non-business trips. That means the distinction between access and egress is very important.

Passenger Awareness of Ground Access Alternatives. All of the models reviewed by Lunsford and Gosling used the actual observable attributes of the competing modes. No consideration was given to travelers' perceptions of the values of these service attributes. **Spear (1984)** recommended that models should be calibrated with data on passenger awareness of ground access modes. This would allow researchers to determine how better marketing could increase ridership.

Perceived Variance in Travel Times. Lunsford and Gosling suggested that lower variance in travel time to the airport may be more important to many air travelers than shorter travel times. This is due to the relative low frequency of departures in many markets and the resulting great inconvenience of missing the flight. The authors also predicted that the perceived variance for travel time parameters would differ greatly between business versus non-business and short haul versus long haul flight stratifications.

Automobile Operating Cost. **Gosling (1984)** wanted to capture the difference between perceived and actual automobile operating cost, which is substantial. He suggested the cost to be expressed in terms of travel time.

Pels, Nijkamp, and Rietveld (2003) looked at access to and competition between airports in a case study of the San Francisco Bay Area. They found access time elasticities to be considerably larger than access cost elasticities for both business and leisure travelers. Their recommendation was that airport planners "that want to increase their share in the local market should invest in relatively fast access modes." The authors felt that the relatively low fare elasticities indicated that an increase in fares to pay for the faster access modes "will have little effect on demand and that an investment in the quality (accessibility) of an airport will be beneficial to both airports and airlines."

The only study specifically considering *distance to the terminal* could be found in the urban transportation literature. It was also the only investigation using gradients:

Cervero (1994) has calculated what he called a "ridership gradient" to determine how fast rail transit's market share falls off with increasing walking distance to the station. He found:

$$\text{Percent Rail} = 1105(\text{Distance in feet})^{-0.795} \tag{1}$$

To arrive at this result, Cervero controlled for 9 other factors determining mode choice for commuters in a binomial logit model (private automobile versus rail).

Cervero's research focused on the question whether transit-focused development could lure significant numbers of Californians out of their cars. He examined the ridership impacts of existing large-scale office projects near stations of five of California's urban rail systems. The selected rail systems had been in operation the longest and therefore offered mature station environment to study the ridership impact of transit-focused office development.

2.4 *Transfer Penalties*

2.4.1 Transfers in the Middle of a Long Distance Trip

It may seem like a digression to discuss the penalty associated with a change of trains or planes in the *middle* of a long distance trip, but understanding traveler behavior in this instance does help explain the high market share of high speed rail in spite of its comparatively long line haul times.

Before 1990, many papers had been written on the tradeoff between fare and frequency in the airline industry (e.g. Morrison & Winston, 1985). **Hansen (1990)** developed an airline hub competition model and for the first time looked at to what extent frequency could compensate for the inconvenience of a connecting flight through a hub. Assuming just one direct flight a day, he found that with no extra distance involved, “15 flights per day on each leg would be required for the hubbed service to gain a 50% share.”

This shows how little travelers like to transfer enroute and is completely consistent with the empirical evidence collected by European railways on the same topic.

Vrtic and Axhausen (2003) estimated a combined revealed and stated preference route choice model of regional and long distance public transportation trips in Switzerland. They concluded that as long as transfer time is included as a separate variable, the transfer itself is penalized with 19 minutes in-vehicle time. (This compares to 13 minutes in-vehicle time in the *urban* transportation literature discussed in Section 2.4.3 below. Again, transfer time was included as a separate variable).

After the introduction of the TGV Atlantique, some conventional trains from Paris to Atlantic seaboard cities were replaced with a combination of high speed TGV trains and diesel trains. Conventional trains have the advantage that at the end of the electrified line, the electric engine at the front of the train is switched with a diesel engine. Passengers can stay on the train. On the other hand, if a TGV high speed train set is used for the electrified portion, passengers have to change from the TGV across the platform to a diesel train at the end of electrified line. **Vilmart and Paix (1994)** found the transfer penalty to be equivalent to between 40 minutes and one hour of in-vehicle time. In other words, the TGV diesel train combination had to be at least 40 minutes faster than the conventional train in order for people to find it worthwhile to switch to the TGV.

Before the opening of the first TGV line, initial plans were to have TGV's only run on the new high speed lines and make passengers transfer to conventional trains even when connecting lines were electrified. These plans were immediately abandoned after research showed how many potential passengers would be lost due to the forced transfer. The initial plan was never implemented and TGV's now connect Paris with every major city in the country as long as at least a small portion of the trip can be run on high speed lines.

Because of the polycentric structure and the resulting dispersed nature of travel patterns in Germany, it is not practically possible to design a long distance train system where no transfers are necessary. "Integraler Taktfahrplan" or Integrated Timed Transfer (ITT), also known as Pulsed Hub System (Maxwell, 2003) became one of the main buzzwords in transportation planning in Germany, Switzerland, and Austria. The goal behind this type of scheduling approach is to "blanket" the country with hubs, where all trains, buses, trams, etc. arrive and depart at approximately the same time in order to minimize transfer times at as many places as possible (Clever, 1997). It is important to note that while an ITT system is designed to minimize transfer *times*, it is also being designed to minimize the *number* of transfers. Great care is taken to do exactly that. The German long distance train timetable for the year 2000 was tested against an origin and destination matrix with 185 124 cells. The simulation showed that 61.6% of long distance travelers never have to switch trains, while 79% of those having to change only need to do it once. This was made possible because of line switching. Even though each line is served at least once every 60 minutes, InterCity trains, e.g., only follow the same route every 4 hours (Krista & Oltrogge, undated).

Experience in Germany on an anecdotal level reveals that travelers from Mainz to Hamburg often prefer to take the conventional InterCity train on the circuitous route via Cologne in order to avoid having to change trains in Frankfurt, half an hour away from Mainz. Traveling via Frankfurt on the ICE is almost 2 hours faster.

In every instance passengers like to enjoy as long a portion of their journey as possible to be completely uninterrupted. And they are willing to pay considerable penalties for that privilege.

A good summary on how passengers feel about switching trains enroute is given in a report to German Rail by **Infratest (1992)** (Infratest Sozialforschung, 1992):

> "Die ganze Gemütlichkeit wird unterbrochen. Erst breitet man sich aus, verteilt seine Sachen überall, packt die Brotzeit aus und dann muß man alles wieder zusammenpacken, Taschen schleppen und dann wieder alles auspacken. Die Zeit, in der man es sich gemütlich machen kann, wird erheblich verkürzt."

> "The whole coziness is totally interrupted. First you get comfortable, spread out your stuff all over the place, get out the sandwiches, and then you have put everything back together again, lug your baggage around everywhere, and then take everything out again. The time in which you can make yourself comfortable is significantly shortened."

Infratest also gives the result of its survey of intercity rail passengers in Germany in the same report. It found that if a transfer was unavoidable, 55% of the respondents would prefer it at the beginning of the journey, 22% in the middle, and 22% at the end. This indicates that a transfer at the end of a trip would be considered as onerous as a transfer in the middle of a long distance trip.

Traveling by plane from London to Paris with 1 hour access time and 1 hour egress time is having the trip cut into three equal pieces, none of them long enough to get comfortable. The disutility of such a trip can be better understood with the concept of "lost time, " time that cannot be used productively. For a 3 hour Eurostar trip with an access time of 30 min and an egress time of 15 min, we have a total lost time of 45 min. For a one hour plane trip with an access time of one hour (including waiting in check-in lines) and an egress time of one hour (including waiting at the baggage carousel) we have a total lost time of two hours. Even though the total door-to-door time is shorter for the plane trip, the total lost time is shorter by rail. Plus, spending one hour in constrictive economy seating may not be that productive either. So many travelers may actually perceive the 3:45 h train trip to be more pleasurable than the 3:00 h total travel time by plane.

2.4.2 Transfers at the Beginning or End of a Long Distance Trip

Unfortunately, to the author's knowledge, no research has ever been done that assesses the penalty of a transfer between an intercity mode and a local access mode, like a subway or a taxi. All studies assumed that a transfer between an intercity and a local access mode takes place.

However, it is entirely conceivable that many final destinations of business trips would be within walking distance of a HSR station located right next to the financial district. No information is available on how such close proximity to an intercity terminal at the end or beginning of a trip affects mode choice.

A simple example will help to illustrate why it is important to know the transfer penalty at the beginning or end of a long distance trip: Connecting air passengers have to walk as far as 1500 m between arrival and departure gates (Detroit Midfield Terminal before the opening of the people mover line). If train passengers were willing to walk as far at the beginning or end of a rail journey, and if a future California high speed rail system were able to use the four downtown San Francisco Market Street stations, practically all of downtown San Francisco, including the financial district and the South of Market area would be within walking distance of a high speed rail station. This could have a noticeable impact on mode shares, since, as pointed out just above, people absolutely do not like to transfer at the end of a trip. If one of the competing common carrier modes required no transfers to reach an area with a very high trip end density, the impact on mode share could be significant.

2.4.3 Transfer Penalties in the Urban Transportation Literature

Since no information at all is available on the penalties for transfers at the beginning or end of a long-distance train trip, it is necessary to take a brief look at the urban transportation literature.

Walther (1991) quantified travelers' valuation of distinct travel time components. A minute waiting time is always perceived to be more unpleasant than a minute spent in a moving vehicle. He defines the time valuation factor while inside a moving transit vehicle as 1.

His empirical research, conducted in Germany, found that the time valuation factor for waiting time increases exponentially with the necessary wait. For a waiting time of 2 minutes the time valuation factor is approximately 2, for 6 min it is 4, and for 10 min about 8. That means he would calculate the disutility of a 10 min wait at a transfer station to be 10*8 = 80.

Walther's research was done in the context of urban transportation planning, but his models were extended to intercity travel (Oetting, 1995).

However, North American research did not find the time valuation factor to increase exponentially with the waiting time. "…there is no a priori reason to expect that time valuation would change markedly in the course of a wait." (Horowitz, 1981) Therefore, most models use a constant time valuation factor for waiting time. In the course of the development of a consensus paper on how transit transfers affect ridership during the evaluation of the Houston METRO project, Don Pickrell wrote in a memo to Chairman Lanier:

> "Somewhat surprisingly, the results of these models imply that transit passengers view time spent transferring from one vehicle to another – including from buses to relatively high-frequency rail service in the controlled environment of modern rail stations, or from one rail vehicle to another, as from <u>three and one-half to nearly four times as valuable as time spent riding aboard transit vehicles</u>." (*26 June 89, unpublished*).

Sometimes researchers use a time valuation factor of 4 as a general estimate: "Waiting time is evaluated by the passenger as four times riding time, …" (Bakker, Calkin & Sylvester, 1988, p. 7). A comprehensive Central Transportation Planning Staff study in Boston clearly distinguishes between a time-independent transfer penalty of 13 minutes in-vehicle time, and an additional waiting time penalty of approximately 2½ times the in-vehicle time (Central Transportation Planning Staff, 1997). The Boston study states confidently that "a transfer penalty for broken trips of 13 min could be established at the 95% confidence level."

It is this time-independent transfer penalty measured in minutes of in-vehicle time, that would be so helpful to know for the transfers between intercity and local access modes. The time-independent portion of the transfer penalty would be an indication for the benefit of a railway station being within walking distance from the point of trip origination or trip end.

2.5 The Treatment of Access Issues in Major Intercity Mode Choice Studies

2.5.1 From Abstract Mode Choice Models to Logit

Most intercity travel models developed prior to the mid 1970's were based on deterministic and aggregate methods of analysis. That meant that the models were more descriptive than causal, had questionable transferability properties, and limited applicability to policy analysis (Stephanedes, Kumar & Padmanabhan, 1984). The key element in the deficiencies of those approaches was the lack of a behavioral basis. That was an inevitable result of the analysis being done at the zonal level (city, region), whereas the behavioral unit is the individual or household (Koppelman & Hirsh, 1986). The emphasis from the mid 60's to the mid 70's was on the development of aggregate models mostly in conjunction with the Northeast Corridor project. A summary of passenger demand modeling from the mid 60's to the mid 80's is given by Koppelman (1984).

Quandt and Baumol introduced the abstract mode choice model in **1966**. It received considerable attention for obvious reasons. Entirely new modes were being contemplated at the time, e.g. maglev service in the Northeast Corridor. Travel models used at that time were entirely inadequate to model demand for as yet non-existent modes. Quandt and Baumol's approach looks at modes purely with respect to their performance characteristics, like travel time, trip cost, and number of departures. It completely ignores the perception a traveler may have of a particular mode. It assumes a business traveler would be indifferent between using air or taking a Greyhound bus, provided all the attributes like travel time, trip cost, and number of departures are the same. This approach has obvious disadvantages. Later research found that automobile travel is consistently believed to have better attributes than it actually has, while public transit is consistently believed to have worse attributes than it actually does. Part of the reason may lie in the disparity of advertising dollars spent to convince people to buy automobiles versus to ride public transit. Please see the discussion on Brög, 1982 later in this chapter (page 36). Mode specific constants are essential parts of today's mode choice models. Among other reasons, mode specific constants are necessary to ensure that the location parameter η of the Gumble distribution in case of the logit model can reasonably be assumed to be zero (M. E. Ben-Akiva & Lerman, 1985, p. 104).

Brown (1968) developed in his dissertation an intercity passenger transportation demand model that improved upon most models used at that time by not simply extrapolating trends, but instead developing a structural model that described the interrelationship between socioeconomic and transportation system characteristics related to trip making. He also overcame the inability of many models of that time to predict demand when current modes are improved or new modes are introduced. Brown described transportation modes abstractly. His model was entirely theoretical without use of empirical data. It did not specifically address access and egress issues.

By the mid 1970's the theoretical basis of the now ubiquitous logit analysis was developed (Domencich & McFadden, 1975; McFadden, 1972, 1974). However it was not until the latter part of the 1980's that these models did not only become state of the art, but also state of the practice.

In the field of intercity travel forecasting, the first major study using disaggregate analysis was by Peter **Watson (1974)**, (see also Stephanedes, Kumar & Padmanabhan, 1984). By the mid-1970's many claims were being made that aggregate models contain significant errors. However, there was no evidence available until Peter Watson was able to show with a dataset that was suitable for the estimation of both aggregate and disaggregate models that this claim was correct. He used data from the Edinburgh-Glasgow Area modal split study. In his model he did use a separate variable to capture the walking and waiting time for train journeys.

Kraft and Kraft (1976) estimated one of the first intercity mode choice models where the demand for each mode was derived from an individual utility function with a stochastic disturbance term. Individual travel mode characteristics were derived from aggregate data. Parameter estimates were given for cost, comfort, convenience and substitution. It was assumed that a person is more comfortable the shorter the travel time. Travel time was adjusted for the time spent traveling to and from mode departure points.

Peers and Bevilacqua (1976) contributed to the discussion of aggregate versus disaggregate models by proposing an alternative approach, developing a set of direct-demand models for estimating intercity transit travel for the Sacramento-Stockton-San Francisco Bay area corridor study. They described in their paper why structural rather than sequential models were chosen, and why direct demand models rather than probabilistic choice models were used. The authors only used a single variable to describe total transit time and total transit cost respectively. No distinctions were made between the access/egress and the linehaul portions of each trip.

Cohen et al. (1978) used a 2-stage modeling process to forecast the 1975-1980 rail patronage in the Empire State Corridor (New York City – Buffalo). The authors used gravity formulations to relate annual volume to city size, government employment, and hotel and motel sales receipts. To estimate mode choice, binary logit models were developed where rail competed differentially with air, auto, and bus, in order to circumvent the Independence of Irrelevant Alternatives (IIA) assumption of logit. To estimate the logit model, only an aggregate database of 31 city pairs was used. No socio-economic variables were built into the model. Rail service and terminal quality variables were included together with linehaul time, cost and frequency. No consideration was given to access/egress issues.

Werner **Brög** introduced the situational approach into the English literature in **1982**. It is an almost entirely qualitative approach relying on long and detailed interviews with relatively few people. This method does not involve any model calibration. Werner Brög's approach is contrarian inasmuch as it does not make use of complicated mathematical models. He has nevertheless been able to achieve considerable forecasting accuracy with his alternative methodology, which has recently also been employed by US researchers, (e.g. Goulias, Brög & Erl, 1998).

Brög has always emphasized the important role of perception in his travel demand analyses. Basic research in Germany during the second half of the 1970's showed that demand models need to take perception into account to have better predictive value. This applies especially to changes in transportation supply characteristics like travel time and travel cost. Surveys indicated that people *over*estimated public transit travel times by 30% and *under*estimated automobile travel times by 15%. Therefore public transit would have had to be about 50% faster to be *perceived* as being equally as fast by the general public (Brög, 1988). That means using only level of service indicators from secondary sources like travel time could introduce systematic errors and thereby cause the assumption of independent identically distributed error terms to be violated. In the last chapter of "Discrete Choice Analysis" entitled "Models of Travel Demand: Future Directions," Ben-Akiva and Lerman note (pp. 363 – 364):

> "The statistical demand models in use typically treat utility as a function of a vector of observed, physically measured attributes. This approach does not explicitly represent the process by which physically measured attributes are perceived and acted on by individual decision makers. It has been proposed that individuals assess alternatives by first constructing some intermediate variables and then evaluate their alternatives based on these intermediate constructs."

This problem has been addressed in recent years with latent variable models.

Brög has also tried to deal with the issue of captivity in his situational approach. As a good illustration of this problem he points out that in Germany in 1995 12% of all local person-trips were made using public transit. Analyzing the other 88%, 24% of the trips could not use public transit due to a particular situation, like traveling with a lot of baggage, small children, etc., in 35% of the cases the area was not adequately enough served by public transit to be seen as a viable alternative, 20% of the people did have no information at all about public transit service in their area and no intention of ever getting it, because it was simply not seen as an option, and 4% were so negatively presupposed towards public transit that they would not consider it under any circumstances, leaving only 5% of the 88% that had used non-transit modes "free" to make a choice. Of the 12% that had used public transit, 4% were objectively required to use public transit. These were captives because they lacked the availability of an automobile. Another 4% needed to subjectively use public transit (they did not want to use a car). That left 4% of the 12% that had used public transit free to make a choice. Brög believes that standard logit models could only be applied to this 5% + 4% = 9% of all person-trips. The issue of captivity has recently been addressed with logit-captivity models. Please see the discussion on McCarthy, 1997 later in this chapter (page 48).

Stephanedes et al. published two major intercity mode choice studies based on fully disaggregated models in **1984**. The background was the recently passed Bus Regulatory Reform Act of 1982, which eased the regulatory burden on the bus industry. The authors found that bus companies could expect a 58% increase in business fare revenue by moderately reducing travel time with the introduction of more nonstop services and fares. The first paper (Stephanedes, Kumar & Padmanabhan, 1984) uses data from a transportation survey of business travelers conducted by the authors in the Minneapolis/St. Paul – Duluth corridor, the second paper (Stephanedes & Kumar, 1984) employs travel data of business trips from the Minneapolis – Chicago city pair. The authors specifically distinguished between out-of-vehicle and in-vehicle-time and found that that distinction significantly improved model performance. Comparisons were made to earlier studies by Grayson (Grayson, 1981) and Stopher (Stopher & Prashker, 1976) which used partially disaggregated data. The authors concluded that National Travel Survey data employed by Grayson and Stopher did not contain accurate information regarding travel time and travel cost. The authors also pointed out inconsistencies in Stopher's model casting "doubt upon that entire specification."

Lyles and Mallick (1990) employed a multinomial logit model to study the Detroit – Chicago corridor using passenger survey data. The authors found that high speed rail would not be able to effectively compete with other modes in that city pair. Out of vehicle time was the only variable related to access and egress and it was dropped in the final analysis because the unexpected sign of its coefficient.

Peter **Stopher et al. (1999)** developed a model for high speed rail patronage in Thailand in the absence of comprehensive local data and any intercity forecasting models for the nation. The method adopted was to create synthetic demand and mode choice models, using parameter estimates from other recent studies. These models then were adjusted by using some of the data that was available in Thailand. Several studies were reviewed to obtain a ratio of in-vehicle to access time. Stopher looked at both urban transportation and intercity mode choice models. He found the ratio to vary between a low of 2 and a high of 10, with most values around 3 to 7. He selected 5 as the mean of the values reported.

2.5.2 Nested Logit Models

Steven A. **Morrison** and Clifford **Winston (1985)** estimated one of the first *nested logit models* for intercity travel. The authors developed a multidimensional mode, destination, and rental car choice model for both intercity business and vacation travelers. The only access/egress related parameter used in their models was a dummy variable for the availability of bus and rail transit at the destination in the rent a car choice model. The estimation of the rental car mode choice model for vacation travelers was successful with all the parameters having the expected signs and in most cases exhibiting a reasonable amount of statistical significance. However, Morrison and Winston were unable to obtain reasonable parameter estimates for the rental car choice of business travelers.

Their major conclusion were that both rail and bus could be successful in diverting both vacation and business patronage from other modes by significantly reducing travel time and increasing frequency, even if this resulted in a small fare increase. Those measures, however, would not have any major effect on increasing the overall size of the market. On the contrary, air's most successful strategy would be to increase market size by reducing fares for vacation travelers, even though this approach would not likely divert patronage from other modes. Morrison and Winston emphasized that any specification of an intercity demand model should attempt to capture the effect of the number of household members on a trip, trip distance, and household income, since these variables had the potential to contribute significantly to the predictive power of the model.

By the time Frank **Koppelman (1986)** introduced his conceptual structure for models of intercity passenger decision making, most studies of intercity travel behavior were still based on the analysis of aggregate data and lacked a behavioral basis. Koppelman postulated that because of the lack of an appropriate behavioral framework for intercity travel, none of the travel surveys conducted so far had collected all the relevant data needed to test important hypothesis. In his follow-up paper in **1989** Koppelman proposed a 4-stage model hierarchy: 1) trip frequency choice, 2) trip destination choice, 3) mode choice, and 4) service class choice. He suggested that this interrelated choice structure could be represented by the nested multinomial logit model. The nested logit model had been introduced into the literature by McFadden (1978a) a decade earlier. In both of his papers Koppelman took a very macro view of the conceptual structure of intercity passenger decision making models. Access and egress issues were not considered.

Christopher **Forinash** wrote his Masters Thesis in **1992** on the application of nested logit models to intercity mode choice. Major research findings were published by **Forinash and Koppelman** in **1993**. At the Third International Conference on Behavioural Travel in 1977, the nested logit model was recommended for "immediate implementation" (Westin & Manski, 1979). Nevertheless, Forinash and Koppelman stated in their paper that the multinomial logit model had been used almost exclusively "until recently." The authors noted Morrison and Winston (1985) as one of the few exceptions. The nested logit model had been limited due, in part, to the lack of more flexible software. Forinash and Koppelman compared the results of several multinomial and nested logit specifications on passenger survey data for the Windsor, Ontario – Quebec City corridor. Their paper demonstrated "a statistically significant rejection of the multinomial logit model in favor of three alternative nested logit models." The adoption of either of the nested logit models resulted in substantially higher rail probabilities at the individual level and greater rail shares on the aggregate level. Thus, the exclusive utilization of multinomial logit models would have significantly underestimated the effect of improvements in rail service. At the average distance traveled of 231 km, it was found that the value of in-vehicle travel time was 22 Can$/h for high income and 16 Can$/h for low income travelers. The values for out-of-vehicle time were 92 Can$/h and 83 Can$/h for high and low income travelers respectively. This means that travelers at that average distance considered out-of-vehicle time about four times as onerous as in-vehicle time.

The restriction of the multinomial logit model (MNL) that the distribution of the random error terms is independent and identical over alternatives leads to the independence of irrelevant alternatives (IIA) property, best illustrated by the famous red bus / blue bus problem (Train, 1986, p. 19). **Wen and Koppelman (2001)** note that while the nested logit model allows the error term of *pairs or groups* of alternatives to be correlated, "the remaining restrictions … may be unrealistic in important cases." Many other models relaxing the assumptions of MNL and all based on McFadden's generalized extreme value (GEV) model have been developed. Among them are the paired combinatorial logit model, the cross nested logit model, the ordered generalized extreme value model and the product differentiation model. The generalized nested logit model (GNL) introduced by Wen and Koppelman (2001) includes all these, plus the MNL and the 2 level nested logit models as special cases. It also closely approximates the nested logit model with more than 2 levels. The authors note that the GNL model "provides a unifying structure for previously reported GEV models, with the exception of the nested logit model, and provides a framework for understanding the properties of these models." To illustrate their model, Wen and Koppelman used again the 1989 VIA Rail dataset that was collected among other reasons to estimate demand for high speed rail service in the Toronto-Montreal corridor. This dataset was also used by Forinash and Koppelman (1993), discussed in the previous paragraph. The hypotheses were tested that the restrictions imposed by the MNL, nested logit, cross nested logit, and combinatorial logit models were realistic in this case. These hypotheses were "rejected at very high levels of significance, in excess of 0.001."

The basic ratio of value of time for out-of-vehicle versus in-vehicle time was 4:1 for most models. However that ratio was closer to 3:1 for the generalized nested logit (GNL) model.

2.5.3 Stated Preference Analysis

The method of *stated preference analysis* for the estimation of travel demand models had been introduced into the literature as early as 1981 by Louviere et al. It had previously already been used in consumer research (Green & Srinivasan, 1978). However, the use of SP data had often been and still is sometimes rejected due to their unknown reliability.

A major milestone in the acceptance of SP data for use in travel demand models was the dissertation by Taka **Morikawa** in **1989**. Its major results were summarized in a paper published by **Ben-Akiva and Morikawa** in **1990**. SP responses cannot be used alone for forecasting actual behavior because of their unknown bias and error properties, but they contain very useful information on trade-offs among attributes. They also add critically important information on the acceptability of yet non-existing new services. And, as this study showed, they can solve an identification problem, if the RP data does not exhibit enough variability. Therefore, the best forecasts combine RP with SP data. Ben-Akiva had presented a general framework for statistical estimators which combine data from multiple sources. This framework was the basis of Ben-Akiva and Morikawa's data fusion method that allow the combination of two or more complementary data sources into a single data base. The authors stated that their framework is general enough to be applicable in a wide variety of contexts (e.g.: combining disaggregate and aggregate data for the estimation of discrete choice models, combining survey and census data to estimate trip tables, or even combining data from different regions to spatially transfer a travel demand model). The methodology was illustrated with survey data of intercity travelers in The Netherlands. The survey included RP data and 2 SP experiments. Ben-

Akiva and Morikawa found that the RP model alone was not able to determine the coefficient for the important linehaul travel time variable (it had an incorrect sign and was not statistically significant due to the limited variation of linehaul travel time in the RP data). In the SP model the coefficients of linehaul time were negative and significantly different from zero. These results were seen to be encouraging for practitioners to employ the joint estimation method. The ratio of terminal time to linehaul time was a stable 5:1 in all of the RP/SP combined models.

The first major study applying this methodology to intercity mode choice analysis was published in **1991** by **Morikawa, Ben-Akiva, and Yamada**. A rail operator in Japan was considering upgrading additional trains in an unnamed corridor and wanted to know how many passengers could be expected from other modes. Three models were estimated in this study, a revealed preference model, a stated preference model, and a combined SP/RP model. In the RP model the coefficient for access and egress travel was double the magnitude of the coefficient for linehaul time, which is what we would expect. In the combined RP/SP model however, that ratio was reversed. Unfortunately, this issue was not discussed in the paper.

In **1989**, the year of Taka Morikawa's dissertation, **Hensher, Brotchie, and Gunn** outlined their method developed to identify the demand for a very fast train (VFT) between Sidney, Canberra, and Melbourne. The integration of revealed and stated preference data was a central feature of their approach. The authors described their intensive data collection effort in detail. Their primary raw data was an intercept survey of present travelers in the corridor (29,982 trip records), a telephone survey of a random sample of 2,116 households in 12 population centers within the catchment area of the new train service in order to be able to relate the data from the intercept survey to the population at large, a detailed face to face interview with its sample drawn from both the intercept and the telephone survey in order to delve in more detail into the underlying reasons for a particular travel behavior, and an international tourist survey. Secondary raw data included level of service characteristics of the travel network, population growth rates, and socio-economic and demographic zonal profiles. The stated choice design focused on the trade-off between linehaul travel time and linehaul travel cost. Access and egress times and cost were held constant for each mode and each respondent, so no conclusions could be drawn on how changes in access and egress parameters would affect mode choice. The trip distances covered in the stated choice design ranged from 120 km to 1700 km. The results suggested that VFT was being viewed as an appealing mode for interregional travel, with its highest popularity being displayed in the medium distance market around 700 km, which also happens to be the approximate distance between the Bay Area and Southern California.

David **Hensher (1998)** developed a nested logit stated choice analysis of pricing options for the demand of sleeper services between Sidney and Brisbane. All intercity mode choice models up to that time had only considered a simple weighted average fare without contemplation of the potential for switching between classes of travel within the train market. Hensher was able to establish a matrix of switching elasticities sensitive to the class of rail fare, thereby enabling rail planners to predict more accurately the impact of differentiated fares on patronage and revenue. He also noted that all things being equal, the current modal use increases the probability of continuing with that mode. This was an important variable to include in a model because it recognizes the role of inertia “and other dimensions of prior experience” that are not picked up by other explanatory variables. The data revealed that the actual door to door travel time for car, coach and train were very similar within each observation, resulting in the coefficient for travel time only being significant for the plane mode. No consideration was given to access/egress issues.

A large number of stated preference studies had been carried out, but relatively little was known about how the design of stated preference studies affected the results until Staffan **Widlert** first presented his research at the 7^{th} International Conference on Travel Behavior in Santiago, Chile in **1994**. His work was later republished in **1998**. For the project 25 different types of interviews were carried out during the same time period of the year on long distance trains in Sweden. 5 700 travelers completed the interviews. Passengers were asked to evaluate different train alternatives, and the questions between the 25 different designs were as similar as possible. The disparity of results between the distinct types of interviews was very large, the smallest and highest value of time differed by a factor of four (!). The coefficients from which the values of time were calculated were all highly significant. It was not possible to conclude which results – if any – were correct by simply analyzing the goodness of fit measures. Widlert found that the most important bias was caused by the simplifications the interviewees made when completing the SP games. The second most important bias arose when the games where not customized to the actual trip the interviewee was making.

Vrtic and Axhausen (2003) estimated a combined revealed and stated preference route choice model of regional and long distance public transportation trips in Switzerland. This study was already referred to on page 26 in Section 2.4.1 on transfer penalties.

The attributes in the SP design were in-vehicle time, number of transfers, transfer time, headway, and price. No access or egress parameters were estimated. However, results from another estimation based on different SP data were briefly reported in a summary table at the end of the same paper. Access time was included in that dataset. The estimated ratio of access time to in-vehicle time was 2.7 in a mode choice model (MNL) and 2.9 in a joint mode and route choice model (Nested Logit).

2.5.4 Box-Cox Transformations

A major research focus for Marc **Gaudry** has been the estimation of the functional form of travel demand models (Gaudry & Wills, 1978). He introduced the Box-Cox Transformation for logit models, named after the two statisticians Box and Cox (Box & Cox, 1964). In his **1991** paper he described three families of mode choice models applicable to intercity travel demand, the second family being the Box-Cox logit family (Gaudry, 1995). Box-Cox transformation allow the relaxation of the assumptions that the familiar S-Curve of the logit probability be symmetric and that the thickness of the tails (indicating captivity) cannot vary. The methodology also makes it possible to do in a continuous fashion what researchers often do on a trial and error basis to obtain a better fit, e.g. exploring quadratic or logarithmic transformations of a variable.

A disaggregate Box-Cox logit mode choice model of intercity passenger travel in Germany by **Mandel, Gaudry, and Rothengatter**, in **1993**, did not contain access or egress parameters at all. “Breaking up the travel time between some of its components (in-vehicle and out-of vehicle time), adding elements (number of driving pauses), …, either did not increase the log likelihood value, or introduced multicollinearity."

2.5.5 Logit Captivity Models

While linear logit models assume that there are no captive market shares for any of the modal alternatives, the *logit captivity model* specifically relaxes this assumption. It is sometimes referred to as dogit, because it also "dodges" the IIA property. According to Patrick McCarthy (see below) the model was developed separately by McFadden in 1976 in an unpublished memorandum of 30 September, by Ben-Akiva (1977) in a working paper, and by Gaudry and Dagenais (1979). Since the logit captivity model is nested in the linear logit model, one can easily use the likelihood ratio statistic to test the hypothesis that the captivity parameters are zero.

Patrick **McCarthy (1997)** used an aggregate time series model estimated on annual intercity travel data for the 32-year period 1960 to 1991, in order to, among other things, identify the effects of regulatory reforms of the 1970's and 1980's. McCarthy could strongly reject the null hypothesis that the captivity parameters are zero. The likelihood ratio test statistic was 35.7, considerably greater than the 13.2 critical value for four degrees of freedom. While the rail and intercity bus captivity coefficients were significantly different from 0, the likelihood of captivity was less than one percent. On the other hand, intercity car travel showed a likelihood of captivity of 48% (!). That means that there are no viable alternatives for a large portion of intercity car trips. It also means that since standard logit models assume intercity car trips to be non-captive, the resulting parameter estimates and elasticity measures are biased and the predicted effects of transportation policy on intercity travel are incorrect. McCarthy estimated both linear logit and the more general logit captivity model and concluded that the latter produced significantly different elasticity measures for important variables. He found that car shares are virtually price-inelastic with respect to travel cost changes.

Concerning the effect of regulatory reform McCarthy concluded that airline deregulation had a strong positive effect on the market share of air travel, while bus deregulation had no identifiable effect on the likelihood of using this mode.

McCarthy's only modal attributes used in his model were travel cost, travel time, fatality rate and mileage. No consideration was given to access and egress issues. He employed a Box-Cox transformation and found that a linear specification of travel cost (Box-Cox parameter not significantly different from one) and logarithmic transformation of travel time (Box-Cox parameter equal to zero) produced the best fit. The models were estimated using TRIO, a software program written by Marc Gaudry and Associates at the Center for Research on Transportation of the University of Montreal. An earlier version of his paper was presented at the Transportation Research Forum Conference in Chicago in October 1995.

2.5.6 Random Effects Logit Models

Standard logit models assume that all individuals exhibit the same preference e.g. towards certain modes, because in a standard logit model only one mode specific parameter is estimated for each mode and it is assumed to apply equally to all individuals. It would be impossible to estimate one mode specific parameter for each individual (fixed effect), but we do not need to, we only need to know the *distribution* of that mode specific parameter across all individuals (random effect). If we assume the distribution of a certain parameter across individuals to be Gaussian, all we need to estimate now is two parameters (the mean and the standard deviation) instead of one parameter (the mean) in a standard logit model. The advantage of this approach is that it relaxes the independence of irrelevant alternatives (IIA) assumption of the MNL model without forcing the researcher to *a priori* define a partitioned tree structure as required by the nested logit model.

Hensher (1997) studied the switching behavior and the calculation of induced demand in the presence of a new mode (high speed rail) between Sidney and Canberra. He employed a stated choice heteroskedastic (random effects) logit model, which had recently been implemented by Bhat (1995) and others. Hensher found relatively high fare share direct elasticities in the business market for air travel, which was not surprising given that the HSR fares were considerably lower than equivalent air fares with door to door travel times for HSR being shorter for most O-D pairs. As in his 1989 study, access and egress times were not varied in the stated choice design, making it impossible to obtain information on the sensitivity of linehaul market share to changes in access and egress. The value of time for the current business air market ranged from 36 Aus$/h to 74 Aus$/h, but included in-vehicle, wait, and transfer times. To be able to compare these with other estimates, Hensher assumed "that access and egress time are approximately valued at 1.5 times the main mode in-vehicle time – a generally accepted ratio." Hensher also pointed out that the appeal of car travel is a single cost independent of party size. He therefore controlled for the number of people in the traveling party, an important issue ignored in the literature "with the exception of Wardman et al. (1994)." However, Morrison and Winston had already emphasized this point in 1985 (see page 39).

2.5.7 Neural Networks

Machine learning techniques have expanded significantly just over the past few years. Neural Network (NN) models do not require any distributional assumptions, can model non-linear systems and have a better ability than classic statistical models to deal with noisy data (Mohammadian & Miller, 2002).

Xie (2000) used a city to city aggregate database of Amtrak passenger travel to model and predict intercity passenger flows. He applied both neural network and standard linear regression models. The focus of his dissertation was to evaluate the predictive ability of NN models employing very large databases with many zero values (Amtrak passenger flows between 97 stations). The passenger flows predicted by the model followed the same seasonal pattern the data exhibited. The root mean square errors improved 38 – 51% over those of the linear regression model.

Abolfazl **Mohammadian and** Eric **Miller (2002)** compared the performance of artificial neural networks to that of nested logit for the prediction of household automobile choices. They found that while both methods gave satisfactory results, the "artificial neural network yielded a better predictive potential."

2.5.8 Activity Based Models

Activity based models play an ever increasing role in the urban transportation literature. Their obvious theoretical attractiveness is that since travel is an induced demand one could get much better results modeling the underlying demand for activities which then result in travel choices. However "the inherent complexities of an approach based on … a whole collection of activities and their associated travel, rather than on characteristics of a single isolated trip," have been a formidable barrier in the development of the approach "beyond either qualitative or rudimentary statistical analyses" (Recker, 2001). Recker offers an analytical framework unifying the complex interactions, thereby "bridging" travel demand modeling and activity based travel analysis.

John **Bowman** developed a complex model system in his **1998** dissertation which was subsequently used by the Metropolitan Planning Organization (MPO) for Portland, Oregon to forecast future travel demand in the region. This practical implementation of an activity based approach for the modeling of urban travel by an MPO represented a big milestone.

2.5.9 Logit Kernel or Continuous Mixed Logit

The most commonly used discrete choice models, multinomial and nested logit, rely on simplifying assumptions, whose validity has been subject to much debate. In her dissertation Joan **Walker (2001)** noted that typically extensions to the basic models were examined and applied in isolation. She presented a generalized methodological framework that integrated these advances. In order to avoid complex multidimensional integrals she employed a logit kernel formulation leading to a probability simulator in MNL form that can be used in maximum simulated likelihood estimation. The advantage is that the resulting discrete choice model has both probit-like and additive i.i.d. Gumbel disturbances. It therefore "combines the flexibility of probit with the tractability of logit." The disadvantage is that "seemingly obvious specification and estimation practices can have unintended consequences." This can lead to serious identification problems, the subject of a special paper (Walker, 2002).

A popular textbook teaching discrete choice methods with simulation (Train, 2003) is available online at http://elsa.berkeley.edu/books/choice2.html.

3 The Nature of Competition between Air and Rail

The reader cannot be expected to be intimately familiar with either Japanese geography or the peculiarities of the Japanese transportation system. Without this short introduction later sections and chapters would be very hard to understand. Because of the frequent use of Japanese words and the exclusive use of SI in this dissertation, it might at times become difficult for the reader to follow the argument, had we not presented some mnemonic and visualization aids with the material. SI is the international abbreviation for **S**ystème **I**nternational d'Unités, or International System of Units. For more information please see www.bipm.org/en/si/.

The first section focuses on the state of competition of passenger transportation in Japan at the beginning of the 21st century. We will see that the Japanese transportation system differs in important respects from those in other industrialized countries. The following sections then will explain which combination of factors enabled the country to achieve such a balanced passenger transportation system. In a balanced system all modes, each one having its unique set of advantages and disadvantages, are able to do what they do best. Frequent short haul feeder flights within a metropolitan area, as e.g. between Burbank and Los Angeles International Airport, made necessary by the slow and unreliable uni-modal ground transportation system, would indicate the opposite – an unbalanced transportation system. Airplanes are required to fulfill a transportation function that is handled more efficiently elsewhere by rail. The focus in this chapter and in this dissertation is on air and HSR, and those two modes do not *complement* each other as in Europe, but *compete* head-on in Japan. Moreover, we do not have to contend with market distorting public subsidies, since both companies which operate HSR service in the corridor of interest, are highly profitable, and mostly so because of their HSR service.

The Japanese transportation system then provides us with an excellent environment to study the nature of competition between air and HSR (we will also touch on conventional rail where appropriate). The individual contributors to mode choice discussed in different subsections below are intended to be a checklist of factors that need consideration in any intercity mode choice analysis which includes air and/or rail. Particular emphasis is given to accessibility simply because it is the most “underappreciated” contributor to mode choice in HSR studies to date.

This chapter also highlights the unique convergence of factors that makes it possible for Japan to develop such an outstanding transportation system, which contributes largely to the country enjoying the lowest per capita green house gas emission of any industrialized nation.

3.1 The State of Competition in Japan

Figure 2 illustrates Japan's unusual high mode share of both transit rail and inter-city passenger rail compared to all other industrial nations. The United States and Canada have the lowest mode share of passenger rail. It is interesting to note that for freight transportation the reverse is true. Railroads in the United States have one of the highest shares of the total freight market (40.6% of tonne-kilometers (tkm) in 1995) – but in Japan rail cargo is insignificant (4.5% of tkm, (Mayrzedt, 2001)). In North America rail is mostly used for freight transportation, while the Japanese rail system is almost exclusively employed for passenger transport. As the map on page 64 shows, most major cities in Japan are seaports, which explains both the importance of coastal shipping for freight and ferries for passenger transportation. Coastal shipping accounts for about half of all domestic "tonnes lifted" (Jackson, 2005).

Air also plays a significant role in passenger travel in Japan (Figure 3). However, we notice in Figure 4, that personal vehicles are used much less than in other countries. There are two reasons for this:

1) Japanese cities are extremely densely populated making them ideal for rail but not conducive for automobile transportation. Parking is very limited.

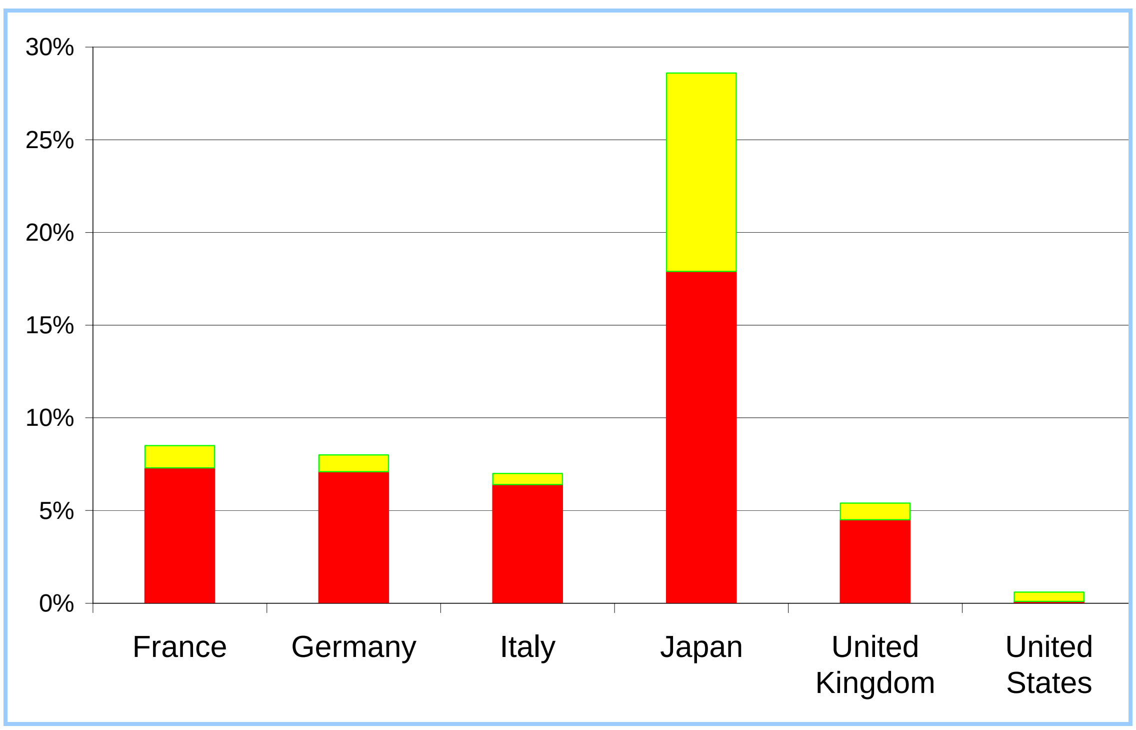
2) Japan has 9000 km of private toll ways including most of the important inter-urban roads (Mayrzedt, 2001). This makes intercity travel by car very expensive.

The bus share of total domestic passenger-km (Pkm) is 6.7%, which is average for most industrial countries. This leaves 11.2% of the domestic Pkm. (*G-7 Countries: Transportation Highlights*, 1999). Ferries dominate this "other modes" category.

Competition between air and high speed rail was boosted by the breakup *and privatization* of the former Japanese National Railroad (JNR) into six vertically integrated regional passenger railways and one national freight carrier in 1987, followed by domestic airline deregulation introduced in stages in the 1990's. The main island, Honshu, is served by three passenger railways: JR East, headquartered in Tokyo, controls the region to the east and north of Tokyo. .JR Central operates out of Nagoya and runs the nation's busiest intercity corridor between Tokyo and Osaka. JR West is based in Osaka and is responsible for the Sanyo Corridor between Osaka and Fukuoka (Map 1 on page 64). The other three passenger railways are JR Kyushu, JR Shikoku, and JR Hokkaido.

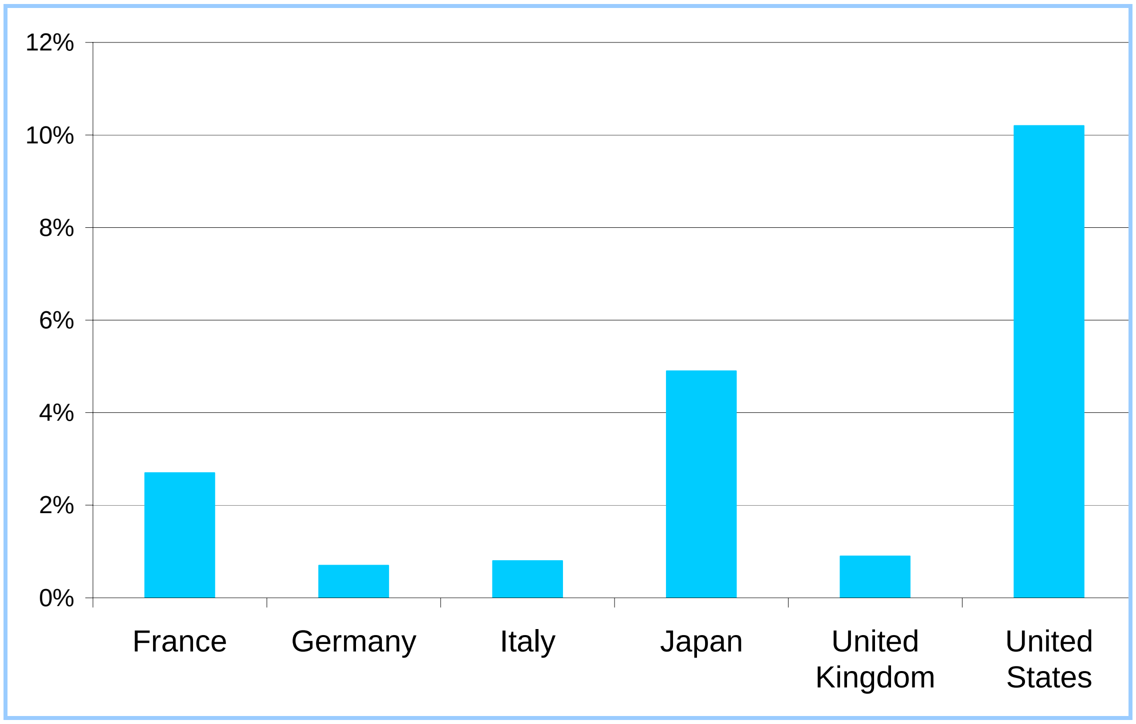
The reader of the electronic version will notice the color coding maintained throughout this dissertation: air is sky-blue and intercity rail fire-engine-red.

Figure 2 - Share of Total Domestic Pkm – Intercity Rail and Transit Rail



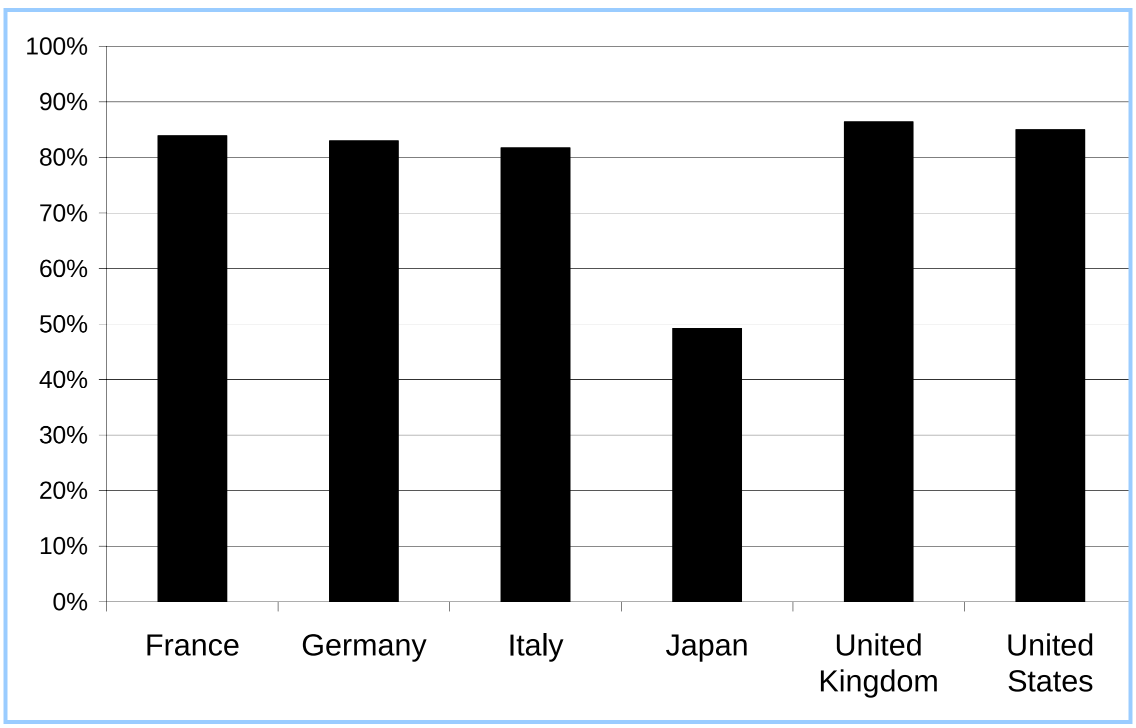
Source: (*G-7 Countries: Transportation Highlights*, 1999)

Figure 3 - Share of Total Domestic Pkm – Air



Source: (*G-7 Countries: Transportation Highlights*, 1999)

Figure 4 - Share of Total Domestic Pkm – Personal Vehicles



Source: (*G-7 Countries: Transportation Highlights*, 1999)

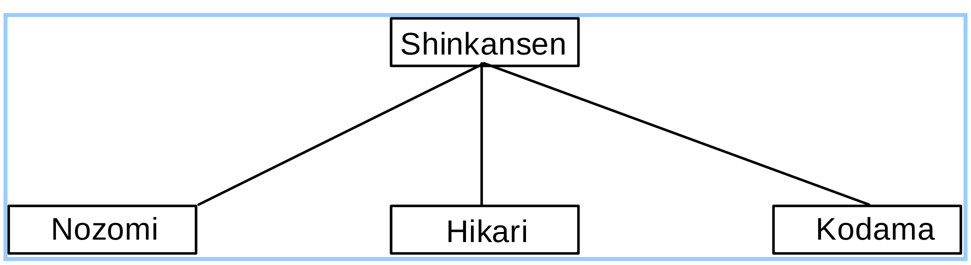
Japan's conventional railways are narrow gauge (1067 mm). Their advantage is the smaller footprint, which is the reason why meter gauges are often used for street cars and special mountain railroads elsewhere. However the small gauge limits both transport capacity and speed. On October 1st, 1964 the first standard gauge "new trunk line" (in Japanese: "shin kan sen") was opened in the heavily populated Tokyo – Osaka = "Tokaido" corridor. By 1975 the line was extended through the Sanyo ("Mountain Sunshine") area to Hakata/Fukuoka on the island of Kyushu. The Tokaido and Sanyo Shinkansen lines serve two thirds of the Japanese population. Three quarters of the Japanese economy (the second largest in the world) is concentrated in the Tokaido and Sanyo areas (Jänsch & Rump, 1990).

To help decentralization and regional development efforts Shinkansen lines from Tokyo into the eastern parts of the country began operation in the early 1980's. Since these lines were built into regions with low population densities, they did not necessarily have to be profitable. The southern section of the Kyushu extension, which will eventually connect Fukuoka with Kagoshima, was opened in March 2004. However to reach the high speed rail segment, travelers first have to transfer to a conventional train in Fukuoka. The analysis in this dissertation focuses on the heavily populated Tokaido and Sanyo corridors (Tokyo – Osaka – Fukuoka).

Standard gauge (1435 mm) high speed Shinkansen services give travelers a choice of three train categories (Figure 5). Local and regional service with stops every 30-40 km is called Kodama or "Echo" (mnemonic: ko-damned slow). There are two long-distance services: Hikari ("Lightning") and the extra fast, premium service Nozomi or "Hope" (mnemonic: zoom ahead). The fare for Kodama and Hikari service is the same, so no passengers would use a Kodama between two Hikari stations. Nozomi require a surcharge and Japanese rail passes are not valid on this train. Table 4 compares the Tokaido and Sanyo Shinkansen lines with North East Corridor service. Notice a 146 km distance for New York – Philadelphia and Tokyo – Shin-Fuji. If the Acela Express were to stop in between New York and Philadelphia not only in Newark, but also in Metropark, Princeton Junction, and Trenton and still made the trip in 1 h and 9 min, its service would be equivalent to the all-stop Kodama. The average speed including all stops is the same for both Kodama and Acela Express trains: 132 km/h or about 80 mph.

The distance from Tokyo to Nagoya is equivalent to New York – Washington. Due to the fast Nozomi and Hikari train service this market was abandoned by the airlines shortly after the opening of the Tokaido line. By contrast, New York – Washington is one on the largest air travel markets in the United States. Punctuality is one of the railways' strengths. "The average delay per train throughout the year is 0.4 minutes, including delays caused by typhoons, earthquakes, snowfall, heavy rain, and other natural disasters" (Kasai, 2000). The distance between Tokyo and Fukuoka is about the same as from New York to the border between South Carolina and Georgia. Once the Shinkansen line is extended to Nagasaki it would be equivalent to New York – Atlanta.

Figure 5 - Japanese High Speed Train Categories



By December 1st, 2002 the total length of high speed double track Shinkansen lines had reached 2049 km. However the original 515 km Tokaido line still carries well over half of all passenger kilometers (Hope, 2003). Air competition is significant in the Tokyo – Osaka market where the Shinkansen still carried 80% of all passengers before the time table change in 2003. However that means a decrease over 5 years by 6% (Table 3). Three airlines have gone on the offensive and introduced cheap fares and frequent flights. After the opening of the Shinagawa station 9 km southwest of Tokyo Central on October 1, 2003, Nozomi service was improved from 3 to 7 trains/h (one every 8.5 min). Three cars in each Nozomi are now available for passengers without reservations ("Tokaido Line to be Improved," 2003). The number of Hikari semi-fast trains was reduced, as was the fare differential between Nozomi and Hikari.

Shinagawa station, which shortened access time by 20 – 30 min for travelers in the southwestern suburbs (Knutton, 2004), cannot be taken into account in the present analysis which uses a 1995 data set. On the other hand Tokyo-Ueno, 4 km north of Tokyo Central, was connected to Tokyo Central by underground tunnel in 1990. Together with three other Northern Shinkansen suburban stations (not shown in Table 4) it is included in our research.

Table 3 - Air Rail Market Shares in the Tokaido and Sanyo Corridors

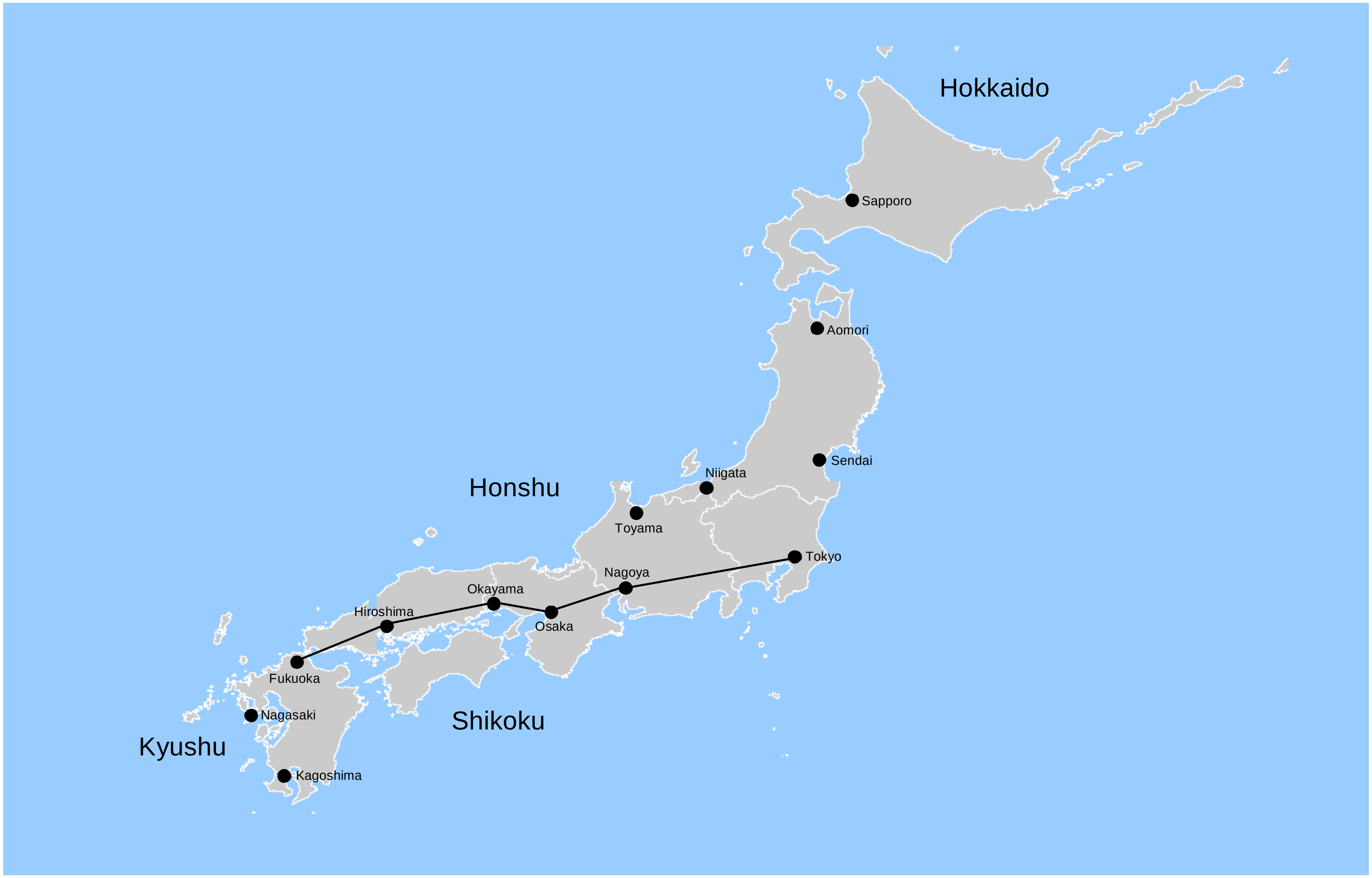
Combined Air - High Speed Rail Market (1999)				
		Market Share		
km	Market	Air	HSR	Passengers / day
366	Tokyo - Nagoya	0%	100%	54,000
553	Tokyo - Osaka	14%	86%	103,000
733	Tokyo - Okayama	18%	82%	6,000
894	Tokyo - Hiroshima	44%	56%	12,000
1180	Tokyo - Fukuoka	88%	12%	22,000

Source: (Hope, 2003)

Table 4 - Tokaido and Sanyo Shinkansen Compared to the Northeast Corridor

			Acela Express				Travel Time from Tokyo			Average Speed in km/h		
			Travel Time	*Average Speed*								
km			*from NY*	*in km/h*	**km**		**Nozomi**	**Hikari**	**Kodama**	**Nozomi**	**Hikari**	**Kodama**
0	**New York**	**NY**	0:00		0	**Tokyo**	0:00	0:00	0:00			
16	Newark	NJ	0:13	74.3	29	Shin-Yokohama		0:17	0:16		102.4	108.8
40	Metropark	NJ	0:26	92.3	84	Odawara			0:36			140.0
93	Trenton	NJ			105	Atami			0:46			137.0
					121	Mishima			0:55			132.0
146	Philadelphia	PA	1:09	127.3	146	Shin-Fuji			1:08			128.8
187	Wilmington	DE	1:29	125.9	180	Shizuoka			1:22			131.7
					229	Kakegawa			1:44			132.1
					257	Hamamatsu			1:57			131.8
298	Baltimore	MD	2:09	138.5	294	Toyohashi			2:17			128.8
318	BWI Airport	MD	2:21	135.3	336	Mikawa-Anjo			2:35			130.1
362	**Washington**	**DC**	2:44	132.5	366	**Nagoya**	1:38	1:56	2:49	224.1	189.3	129.9
375	Alexandria	VA			396	Gifuhashima			3:03			129.8
415	Manassas	VA			446	Maibara			3:22			132.5
472	Culpeper	VA			514	**Kyoto**	2:16	2:41	3:47	226.8	191.6	135.9
542	**Charlottesville**	**VA**			553	**Shin-Osaka**	2:32	2:59	4:03	218.3	185.4	136.5
					590	Shin-Kobe		3:14	4:18			137.2
					612	Nishi-Akashi			4:31			135.5
641	Lynchburg	VA			644	Himeji			4:45			135.6
					665	Aioi			4:59			133.4
742	Danville	VA			733	**Okayama**	3:15	4:01	5:19	225.5	182.5	137.9
					758	Shin-Kurashiki			5:35			135.8
					791	Fukuyama		4:22	5:49			136.0
					811	Shin-Onomichi			6:02			134.4
824	Greensboro	NC			823	Mihara			6:13			132.4
843	High Point	NC			862	Higashi-Hiroshima			6:27			133.6
900	Salisbury	NC			894	**Hiroshima**	3:52	4:54	6:53	231.2	182.4	129.9
					936	Shin-Iwakuni			7:10			130.6
967	Charlotte	NC			987	Tokuyama		5:34	7:30			131.6
1003	Gastonia	NC			1031	Okoori			7:47			132.5
1091	Spartanburg	SC			1093	Shin-Shimonoseki			8:25			129.9
1141	Greenville	SC			1112	Kokura	4:39	5:56	8:34	239.1	187.4	129.8
1189	Clemson	SC			1180	**Hakata (Fukuoka)**	4:57	6:17	8:55	238.4	187.8	132.3
1244	Toccoa	GA										
1304	Gainsville	GA										
1382	**Atlanta**	**GA**										

Source: (*Overseas Timetable*, 2000)



Map 1 – Tokaido and Sanyo Corridors in Relationship to Major Islands and Cities

3.2 Accessibility Related Mode Choice Determinants

3.2.1 Definitions

The etymological origin of the English word *access* can be found in the Latin verb *accedere*, to approach. That lead to *accessus*, the approach, and *access* to be understood as "permission, liberty, or ability to enter, approach, communicate with, …" (*Merriam Webster's Collegiate Dictionary 10th Edition*, 1993). *Accessible* then means "capable of being reached" both in a literal and figurative sense. *Accessibility* is often used in the context of physical disabilities, but for the purpose of this dissertation we would like to use the following definition:

- **Accessibility** is to be understood as the weighted sum of all attributes of every mode of transportation commonly used to travel from the trip origin to a linehaul terminal or from a linehaul terminal to the final destination. The weights are proportional to the perceived importance that each traveler assigns subjectively to each mode and mode attribute.

In everyday language the word *trip* can either have a uni-directional (trip back home) or a bi-directional (trip to Europe) meaning. For the purposes of our research, we would like to define this term precisely:

- **Trip** is defined to only refer to one-way travel.

- To denote bi-directional travel, we will use the term **Round Trip**.

Each **Observation** in our database describes one trip with one origin and one destination for a single traveler. That means in this dissertation the terms observation, trip, traveler, trip maker, and passenger can be used interchangeably.

Empirical research differentiates between access and egress based on two distinct effects:

1. Access and egress referring to the beginning or end of a trip.
2. Access and egress referring to the home or non-home part of a trip.

We prefer the second definition for two reasons:

- Those non-linehaul portions of the trip, which are perceived as most similar by the traveler are encapsulated under the same term. Access for example refers to non-linehaul travel in the home region, which the trip maker is likely to be intimately familiar with.. This makes for a more meaningful comparison between access and egress, than if travel in his or her home district is assigned to access and egress only on the basis of trip direction.
- High speed rail can get travelers close to big clusters of office buildings or hotels. However, it cannot do the same for residential areas – the home part of the trip.

After making the last distinction, we are still left with three different meanings of the English word *access*:

1. It usually means a single one-way trip to or from the linehaul terminal (uni-directional).

2. It may also be used to refer to the sum of access and egress (bi-directional).

3. When the writer deems the distinction to be of no importance in the context, access can refer to either access or egress (uni-directional).

We will only use the word access in its first meaning.

- We define **Access** as one trip *between* home or office and HSR station or airport.

- **Egress** is defined as one trip *between* the non-home destination and airport or HSR station.

There is no English term which unambiguously encompasses the sum of access and egress. To avoid these problems we propose a new “inclusive” term that covers both access and egress. It is already widely used in the field of transportation engineering: FEEDER.

Major arteries have feeder roads. Railroads have main lines and feeder lines. Air transportation distinguishes between trunk and feeder routes. All of these examples imply bi-directionality.

Access to the terminal and egress from the same terminal would be called the terminal's feeder service. When emphasizing its operational aspects we describe it as the terminal's feeder operation. Access operations would only deal with incoming feeder traffic. Feeder time as distinguished from linehaul time refers to the sum of access and egress time.

- The term **Feeder** denotes the sum of access plus egress.

This leads to an intuitive and easily memorized equation:

$$\text{Travel Distance} \equiv \text{Feeder Distance} + \text{Linehaul Distance} \tag{2}$$

… and to the precise definition:

- **Travel Distance** describes the sum of linehaul distance plus feeder distance.

Analogously we would define travel cost as the sum of linehaul cost plus feeder cost.

$$\text{Travel Cost} \equiv \text{Feeder Cost} + \text{Linehaul Cost} \tag{3}$$

With travel time we also have to take terminal time, which is the sum of origin terminal time plus destination terminal time, into account.

$$\text{Travel Time} \equiv \text{Feeder Time} + \text{Terminal Time} + \text{Linehaul Time} \tag{4}$$

- **Travel Time** expresses the sum of feeder time plus terminal time plus linehaul time.

The terms travel time, travel cost, etc. are used somewhat informally by some researchers. When the feeder portion of a trip is not explicitly considered in the analysis travel time or travel cost is often used when only referring to linehaul time or linehaul cost. Such imprecise terminology is best to be avoided.

Please note that the terms access, egress, feeder, and linehaul are relative terms, whose meanings depend on the research focus. Our inquiries concentrate on the competition between high speed rail and air, so any travel, no matter how long, on conventional rail to the high speed rail terminal is considered to be part of the access portion of the trip. Were we to analyze commuter rail operations, only the part from home to the commuter rail station would be understood as access.

The above definitions ensure that the word *access* has only one single meaning, a trip *between* home and HSR station or airport, and it is used only with this distinctive meaning for the remainder of the dissertation.

This only leaves one definition for this subchapter. The term *Competitive Triangle* is illustrated and defined in the following subsection.

3.2.2 The Competitive Triangle – The Relationship between Feeder and Linehaul Time

The first section illustrates with the help of a simple model how access can be the deciding factor to determine mode choice between air and HSR. This simple model makes a strong case that access and egress variables should never be left out of a mode choice model, which, unfortunately, is the case with many models up until today. This model has been kept as simple as possible with a generic specification for the explanatory variables and no mode specific constants. That means it is an abstract mode choice model (p. 33). For an illustrative model this is not a major restriction because we can surmise from Table 1 on page 2 that travelers do not appear to have a significant negative disposition towards using high speed rail.

For this simple example we also assume that terminal time is included in feeder time.

As briefly described earlier, research has shown that business travelers between London and Paris prefer the 2:45 h journey on the Eurostar featuring short access and egress times to the 1 h air travel which necessitates long travel to and from the airport in addition to long waiting times at security checks, ticket counters and in departure lounges. The distinction between X denoting access time and L representing linehaul time is the distinction between unproductive and productive time.

Formally, in the model:

$$V = \alpha_1 X + \alpha_2 L \tag{5}$$

where:

V denotes the systematic or observed utility of a representative traveler
X denotes the sum of access and egress, i.e. feeder time, and

L denotes linehaul time

we would expect:

$$\alpha_1 < 0, \ \ \alpha_2 < 0, \text{ and } |\alpha_1| > |\alpha_2| \tag{6}$$

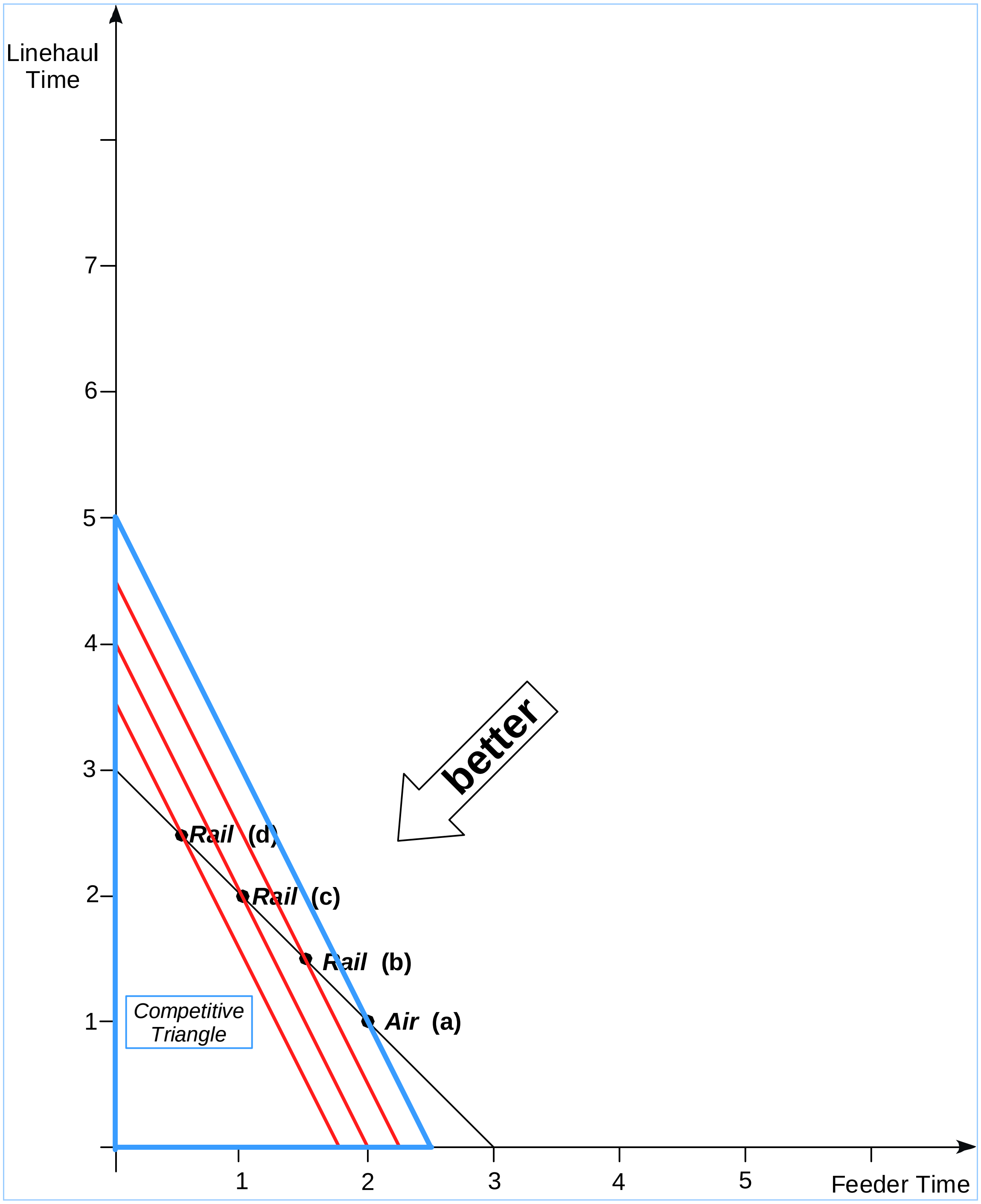
In a hypothetical situation:

- air (a) with a 1 h linehaul, 1 h access, and 1 h egress time (2 h feeder time) competes with
- rail (b) characterized by a 1.5 h linehaul time, but requiring only 1 ½ h of feeder time.

Total travel time for the two competing modes is the same, but it is split up differently between unproductive access and egress and more productive linehaul time. If we accept the hypothesis that feeder time is more onerous than linehaul time, then, *ceteris paribus*, we would expect rail to have a higher market share than air. As we look at two other hypothetical scenarios (c) and (d) we would expect rail to increase its market share as linehaul time increases while feeder time decrease.

Figure 6 plots the different scenarios on the X-L plane (X symbolizing feeder time, and L denoting linehaul time). While all 4 competitive configurations lie on the same line of a constant total travel time of 3 h, (a) represents a lower utility than (b), (b) represents a lower utility than (c), etc. So we know the indifference curve for scenario (a) is further from the origin than the indifference curve for scenario (b), etc. In this case we assumed access time to be doubly as onerous as linehaul time, or $\alpha_1 = -2$ and $\alpha_2 = -1$ in our simple model $V = \alpha_1 X + \alpha_2 L$.

Figure 6 - The X-L Plane with Indifference Curves

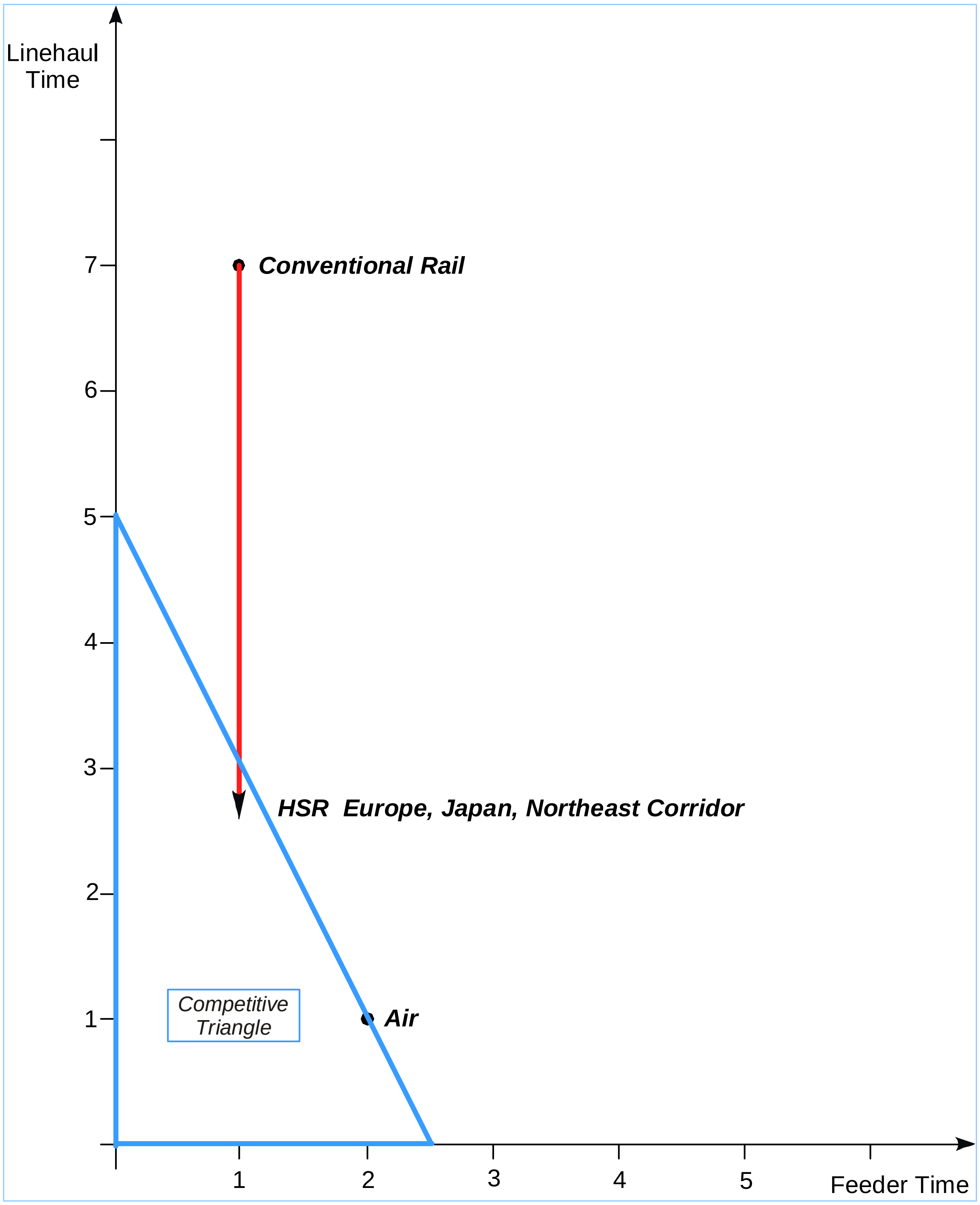


These assumptions for $\alpha_1 = -2$ and $\alpha_2 = -1$ appear reasonable, or even conservative, given the results of the literature review. All authors agree that access time is more onerous than linehaul time. Their estimates range from 1.5 times as onerous to 10 times as onerous (please see Section 2.1 starting on page 14). The thick blue indifference curve representing the utility of air, connecting 5 h linehaul time and 2.5 h access plus egress time (= feeder time) represents the outside border of the "competitive triangle" that high speed ground transportation systems have to get into in order to be competitive with air.

- **Competitive Triangle** describes a right triangle bounded by the non-negative x- and y-axes of a Cartesian coordinate system representing feeder time and linehaul time respectively. Its hypotenuse is the locus of all feeder time and linehaul time combinations with equal level of utility, namely the utility of the air service for that city pair.

The simple model developed so far is sufficient to illustrate how high speed rail in Europe and Japan could have been so successful while similar high speed rail projects in the United States could fail unless modified. Figure 7 describes a city pair with a distance of 700 km, e.g. Madrid – Barcelona. The California Corridor between Los Angeles and San Francisco is approximately equidistant. Conventional rail in both Europe and Japan would take about 7 h to connect the end points of that corridor. Given the land use patterns in Europe and Japan with very high population densities in the urban cores best served by rail, the assumption of a one hour combined access and egress time seems realistic. High speed rail for the most part serves the same terminals as conventional rail, so the feeder time would be the same.

Figure 7 - High Speed Rail in Europe and Japan



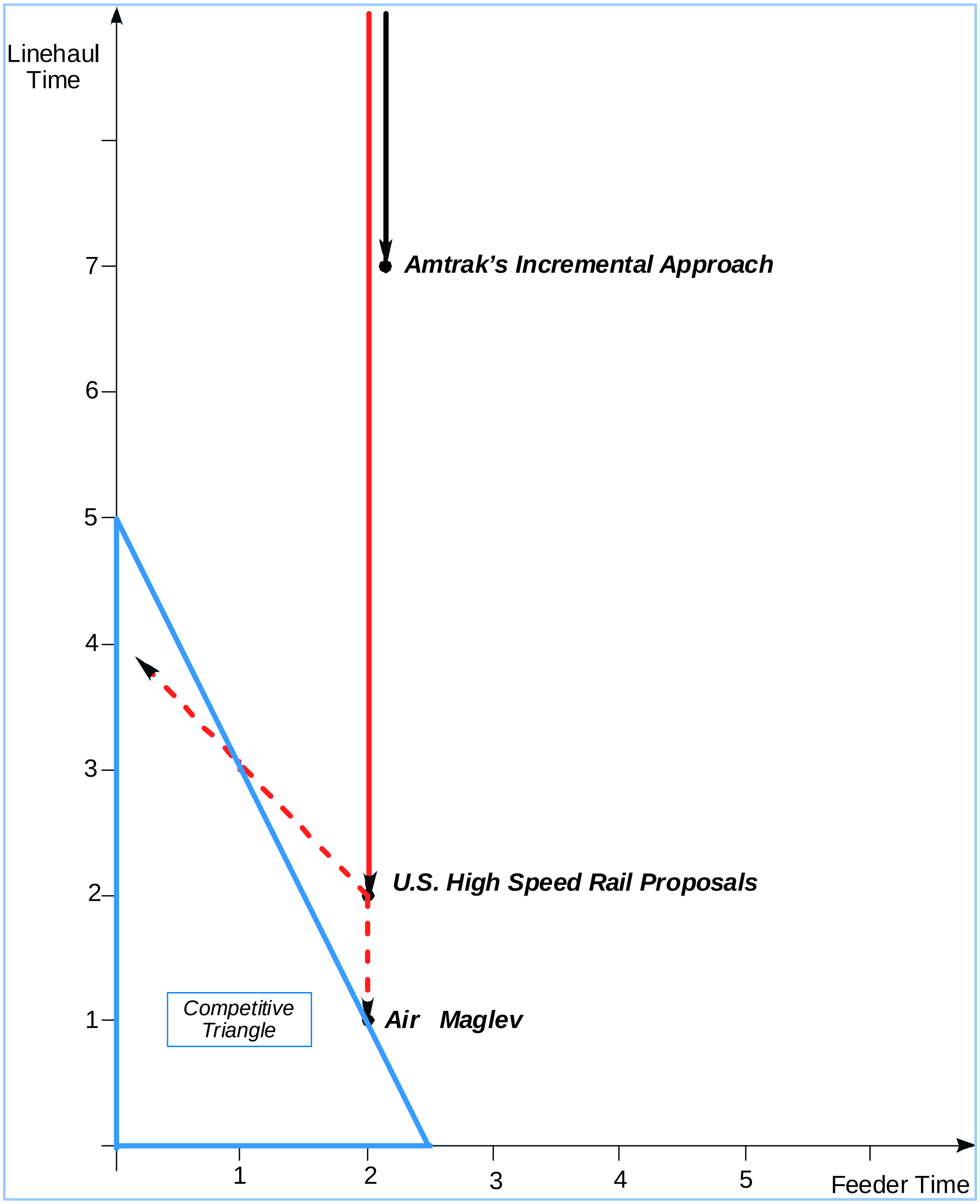
Essentially, the only improvement over conventional rail that high speed rail needed to make was to reduce the linehaul time from 7 h to typically 3 h. That alone made HSR competitive with air and therefore, by default, with other modes as well. Reducing the linehaul time from 7 h to 3 h was sufficient to "get into the competitive triangle." By contrast, land use patterns in the United States with generally low density, multi-centered urban areas would suggest that rail does not enjoy a similar access advantage over air.

Figure 8 illustrates that even using what has been informally called "Very High Speed Rail" systems (maximum speed 350 km/h) that might reduce intercity travel time to as little as 2 h will probably not be sufficient for rail to be competitive with air. The only two alternatives are either building a maglev system which in the future may well approach the linehaul time of air, or improving access by serving more suburban rail stations at the expense of longer linehaul times.

Figure 8 also points to the futility of Amtrak's "incremental approach." Amtrak is spending billions of dollars to upgrade American freight railroad track to European and Japanese passenger rail standards, enabling it to provide a similar service as conventional rail in those countries. That is very similar to saying: "Our competitors are offering the widget for $ 2.- Let's spend billions of dollars in new capital equipment so we can offer widgets for $ 4.- instead of previously $ 7.-" This strategy is dependent on continued heavy subsidies.

Railway Age reported that more than $ 1.2 billion had been invested by the State of California to upgrade right-of-way and rolling stock, and "hundreds of millions additional dollars are in the pipeline." The fare box return, which does not include interest and depreciation on Amtrak owned equipment, is 48% for the San Diego – Santa Barbara corridor, 42% for the Oakland – Bakersfield one, and 32% for San Jose – Sacramento (Wolinsky, 2000).

Figure 8 - High Speed Rail in the United States



For more information on the pitfalls of Amtrak's incremental strategy, the reader might want to refer to a book, well known in the industry, and written by a former president of the High Speed Ground Transportation Association and one time Amtrak executive (Vranich, 1997).

For rail to be self sufficient it needs to offer a combination of access and linehaul time that is inside of the "competitive triangle." Figure 7 demonstrates that by reducing access time while at the same time increasing linehaul time, HSR could improve its competitive position in Europe or Asia even further. Reducing access time by increasing linehaul time is optional in Europe and Asia (Figure 7), but it is required in the United States (Figure 8).

Trains can reduce access time while increasing linehaul time by using their competitive advantage vis-à-vis air, described in Section 1.1 beginning on page 3. The trade-off between access time and linehaul time depends on the Marginal Rate of Substitution, which is the same as the slope of the indifference curve. It can be estimated in mode choice models.

$$MRS = -\frac{\partial V}{\partial X} \Big/ \frac{\partial V}{\partial L} \tag{7}$$

In our hypothetical example the Marginal Rate of Substitution is $-\alpha_1 / \alpha_2 = -2$.

If access time were reduced by 30 min, as long as the resulting increase in linehaul time was less than 60 min, rail would improve its competitive position. Please note that this simple application of our model employs a rather conservatively low marginal rate of substitution of –2 given the results of the literature review.

A 30 min feeder time cut is not unrealistic. Recall from Section 3.1 that Shinagawa station shortened access time by 20 – 30 min for travelers in the southwestern suburbs of Tokyo (Knutton, 2004). This access time reduction can be very inexpensive, if all that is required is adding one additional stop at a station that has already been built. On the other hand, reducing linehaul time by 60 min is almost always very expensive, since a running time reduction of that magnitude can usually only be achieved by building a new line. Deutsche Bahn AG (German Rail) uses € 50,000,000 internally as a rough estimate for the cost of one minute in linehaul time reduction when a new line is required. That means the cost for a 60 min linehaul time reduction in Germany is about € 3 billion, which is a little more than $ 3 billion. The reader surely noticed that two extreme cases were compared: no capital outlays for a 30 min reduction in access time achieved by stopping at an already existing suburban station, versus a $ 3 billion cost in Germany for the equivalent 60 min decrease in linehaul time. Germany very likely has the most expensive high speed rail system in the world (Clever, 1996). We should also remember that an extra stop only benefits those travelers that get on or off at that station. The other passengers on the train experience an increase in linehaul time, often in excess of 5 min, without compensation. In spite of these caveats one would expect a railroad company to exhaust all feeder time savings options before embarking on a new line construction program, in other words offer its services high in the upper portion of the competitive triangle. In reality however, feeder time reductions are almost never discussed in the context of a high speed rail project.

This leads us to believe that the result of this simple model (a 30 min reduction in feeder time is equivalent to a 60 min decrease in linehaul time) is *not* intuitive.

Note that this model might give the reader the impression that we always have to trade a longer linehaul time for a shorter access time. That is correct for the individual train that makes the additional stops, but not for the HSR system as a whole. For the system the tradeoff is usually between accessibility and frequency. The finite demand in a particular travel market only supports a limited number of hourly frequencies, and if one of these frequencies is used to serve intermediate stations, it is not available for non-stop service anymore.

Because of high demand Japan does not have to trade off frequency against accessibility, as explained in Section 3.3.7 Frequency below.

3.2.3 Access/Egress Time versus Other Accessibility Metrics

The previous section deals exclusively with the sum of access and egress time. That is sufficient for the simple model presented, but for our in-depth analysis we have to take *all* qualities and traits of travel options related to accessibility into consideration. Recall our definition of accessibility: Accessibility is to be understood as the weighted sum of *all* attributes of every mode of transportation commonly used to travel from the trip origin to a linehaul terminal or from a linehaul terminal to the final destination. The weights are proportional to the perceived importance that each traveler assigns subjectively to each mode and mode attribute.

We need to balance our desire to collect all the data available in an ideal world with the resources at our disposition. The previous chapter of this thesis helps us strike this balance.

From the literature we know that it is not objective attributes which determine mode choice but the *perception* of these objective attributes (please see Brög (1982) on page 36). We have also learned that access time is not a very good proxy for accessibility, certainly not for accessibility by public transportation (please see Sobieniak et al. (1979) on page 20). When modeling the choice of access mode to intercity transportation terminals Sobeniak found that access at trip origin, service frequency, and baggage handling were much greater determinants of mode choice. Lunsford and Gosling (1994) similarly noted that travel time does not provide a comprehensive representation of ground access quality at an airport. They cited other researchers, especially Harvey (1986a; 1986b; 1987; 1988), who emphasized the importance of stratification between business and non-business travelers, as well as the disutility of handling extra baggage.

Harvey's findings also suggest that non-business females evaluate the concept of "accessibility" to an intercity terminal completely differently than males. They emphasize the need to include **sex** as an important socio-economic variable when modeling the effect of accessibility on linehaul market share (please see page 22).

We can draw an important conclusion from the above synopsis:

It would not be a very good use of time to collect access and egress times for all modes between every zone and every HSR station and every airport in our analysis area. First of all, access time does not appear to be a very good metric for accessibility. Secondly, we would assume a level of information on the part of the traveler which is clearly unrealistic. It is well known that consumers seldom use all the available information before making buying decisions causing their behavior to fall short of economic rationality. Mathematical philosopher Alfred North Whitehead stated:

> "It is a profoundly erroneous truism … that we should cultivate the habit of thinking of what we are doing. The precise opposite is the case. Civilization advances by extending the number of important operations which we can perform without thinking about them." (Barwise & Meehan, 2004)

The data sets available to us will be discussed in detail in Chapter 5. Suffice it to say that we have the great circle distance between the administrative center of all of the access and egress zones and all our linehaul terminals. This is a reasonable proxy for access time and access cost since most of the zone sizes are fairly small.

We do want to use all the socioeconomic information and trip specific details that are available to us, because they could greatly impact the perception of accessibility in the mind of the traveler.

Our data distinguishes between three trip purposes: company business, personal business, and vacation. We also know whether the traveler returned on the same day or needed to stay at least one night away from home. With this level of detail the data implicitly takes the amount of baggage into account. Furthermore, the trip maker's age and sex, as well as, on a zonal level, household income and vehicle ownership are known to us. By interacting these personal and trip characteristics with access and egress distance respectively, we are able to compose a fairly sophisticated, stratified representation of the accessibility of each linehaul terminal from a particular zone.

There are two additional qualities that usually influence the perception of accessibility: frequency of service and schedule reliability of the access modes. As was discussed in Section 3.1, Japan has the highest mode share of transit rail of all industrialized nations (Figure 2 on page 58). Rail dominates intra-urban travel. Private railroad companies are competing for passengers traveling within metropolitan areas (Jackson, 2004c). It is highly unlikely that we could notice different levels of accessibility between two linehaul terminals due to differences in frequency of service and/or schedule reliability of the access mode. Because of high levels of multicollinearity the inclusion of these two attributes would have a tendency to make our model estimation unstable.

To summarize the previous discussion: great circle distance between the linehaul terminals and the administrative center of each origin and destination zone is the only accessibility related supply-side variable available to us. The same is true for egress. However, when combining both access and egress distance with the rich set of socioeconomic variables available to us, we are able to quantify the notion of accessibility in the eyes of each individual trip maker very precisely, since we can take all major choice determinants into account which the literature identified for us. The lack of feeder times and fares is therefore not as large a limitation as might be expected at first glance.

3.2.4 Distinction between Access and Egress

The analysis thus far has focused on the sum of access and egress. However, access and egress need to be considered separately since travelers may value them differently. Hutchinson found that for daily commuters access time to the transit system from their homes is far less important than egress time from the transit system to their offices. Go Transit commuter rail in Toronto provides a good illustration for Hutchinson's findings. In spite of being directly connected to one of the most efficient subway systems in North America, Go's ridership potential is basically limited to the number of work locations within an approximately 700 m radius around the main railroad station. This is a good demonstration of the home versus non-home effect of access/egress, which we are examining in this dissertation.

The previously cited survey of intercity rail passengers in Germany exemplifies the beginning/end of trip effect. It found that if a transfer was unavoidable, 55% of the respondents would prefer it at the beginning of the journey, but only 22% at the end (please see Infratest Sozialforschung (1992) on page 28). Both examples give us ample reason to suspect a different valuation of access and egress on the part of the traveler.

As mentioned earlier, if egress is far more important than access, it would give high speed ground transportation systems a competitive edge over air, since rail systems can get travelers close to big clusters of office buildings and hotels. However, high speed rail systems cannot do the same for residential areas. This was an important reason for our earlier definition of access and egress.

Our data includes the residential zone for each traveler in addition to the trip origin and destination. This is very unusual for an intercity travel survey. The residential zone code makes it possible to clearly distinguish between access and egress, as defined above, therefore allowing us to derive more meaningful results than simply regarding access to be the trip *from* the one-way trip point of origin to the origin terminal and egress *from* the destination terminal to the one-way trip destination.

3.3 Other Mode Choice Determinants

3.3.1 Basic Considerations

Before deciding which additional supply side factors and socio-economic descriptors to include in our model, it is important to remind ourselves of the following fact: It is better to include and keep (even if the t-statistic is less than 2) a variable that does not belong, than to omit a variable that belongs. A variable that does not belong may create additional noise, larger standard errors, and less efficient estimates, but no bias. An omitted variable that belongs creates biased estimates of other parameters. The effect of the omitted variable goes partially into coefficients of variables correlated with it, causing its coefficient estimates to be biased, and partially into the error term, causing dependence in the error term and IIA to fail. This means the model will be wrong.

Having said that, we also have to realize that there is no absolute standard to choose between bias and efficiency. Multicollinearity in theory does not cause any problems, since it does not violate any regression assumption. Goldberger (1991) even compared multicollinearity to what he called "micronumerosity" (lack of data). But high multicollinearity causes less precise estimates due to large standard errors. This leads to large confidence bands making it more likely for the researcher to have to accept the "zero null hypothesis" (Gujarati, 1995, p. 327). In practice, researchers often omit variables that are highly correlated to other variables in the model to obtain more precise estimates, paying the price of biased estimates. Due to our large data set we do not have to be seriously concerned about inefficient estimates. When balancing between bias and efficiency we generally favored the greater number of regressors, thereby minimizing bias.

The list of twelve non-feeder related supply-side attributes together with socio-economic variables in this subchapter is intended to be a comprehensive list of mode choice determinants which should at least be considered when specifying intercity travel models. For variables that are omitted, a short discussion as to the possible effect of that omission on the final model result should be included in the final analysis. Maybe due to space constraints in journal articles, this is unfortunately almost never done at the present time.

A brief summary of the twelve non directly feeder related mode choice determinants in light of our model estimation results is presented in Chapter 7 Conclusions.

3.3.2 Definitions

We first need to understand the precise meaning of high speed rail (HSR):

- Passenger service using steel wheel on rail technology with a maximum speed in excess of 200 km/h in scheduled revenue service is considered **High Speed Rail** (HSR).

- Passenger service using steel wheel on rail technology with a maximum speed of 200 km/h or below in scheduled revenue service is referred to as **Conventional Rail**.

The first high speed train in the world, the Bullet Train, started operating in Japan between Tokyo and Osaka on 1 October 1964 with a top speed of 210 km/h. 200 km/h can be visualized as about double the average American freeway speed of 100 km/h or 62 mph. Two equations make the definitions above more succinct:

$$HSR = V_{max} > 200 \text{ km/h} \tag{8}$$

$$CR = V_{max} <= 200 \text{ km/h} \tag{9}$$

where HSR and CR is a 0/1 dummy variable indicating High Speed Rail and Conventional Rail respectively, and V_{max} denotes the maximum speed in revenue service.

Figure 9 shows some of the UIC designations on a typical conventional European intercity coach. The maximum speed of 200 km/h is clearly marked. The UIC (Union International des Chemins de Fer) is the European equivalent of the Association of American Railroads (AAR).

In the United States the Metroliner between New York and Washington is the fastest conventional train (V_{MAX} = 200 km/h or 125 mph). The Acela Express reaches a top speed of over 200 km/h on a section of track between Boston and Providence in scheduled service and is therefore considered to be High Speed Rail.

200 km/h is also an important technical boundary. At the beginning of the 21st century, only electric traction has a power to weight ratio that makes commercial operation above this speed economically feasible. The critical advantage is its ability to use remotely produced energy fed through overhead wires. All other propulsion methods have to produce their own energy on board and carry the necessary raw materials with them at all times.

Figure 9 - UIC Markings on Conventional European Intercity Coach

To complete this section, the difference between High Speed Rail (HSR) and High Speed Ground Transportation (HSGT) should also be clarified:

- Any guided ground transportation passenger service, regardless of the technology used, with a maximum speed in excess of 200 km/h in scheduled revenue service is considered **High Speed Ground Transportation** (HSGT).

At the present time the only other high speed ground transportation technology deserving serious attention is magnetic levitation (maglev). So for all practical purposes the following equality holds:

HSGT = HSR + maglev

Maglev has its own set of competitive advantages vis-à-vis high speed rail. Those are mostly related to high frequency, high speed regional applications like connecting two airports within the same metropolitan area. This dissertation concerns itself exclusively with intercity travel. Maglev's competitive advantages vis-à-vis high speed rail are therefore of no interest to this research and are incorporated only by reference (Clever, 2005).

Maglev's competitive advantages vis-à-vis air are the same as HSR's competitive advantages vis-à-vis air (please see page 3 in Section 1.1). All discussions in this dissertation apply equally to HSR and maglev which means that for the purposes of our research we can use the terms HSR and HSGT interchangeably.

Finally, the reader should be warned that official U.S. definitions of High Speed Rail, accessible through the *Dictionary* tab on the website of the Bureau of Transportation Statistics www.bts.gov, are confusing at best. The American Transit Association (APTA) states:

> "High Speed Rail: A rail transportation system with exclusive right-of-way which serves densely traveled corridors at speeds of 124 miles per hour and greater." (*Transit Fact Book*, 1996)

Under this definition the Metroliner would be considered High Speed Rail. The author knows of no person or organization (except for APTA) who would regard the Metroliner as high speed rail.

The Code of Federal Regulations (49CFR37) explicitly includes maglev in the definition of high speed rail, even though maglev does not use rail, but a concrete guideway.

3.3.3 Classes of Service

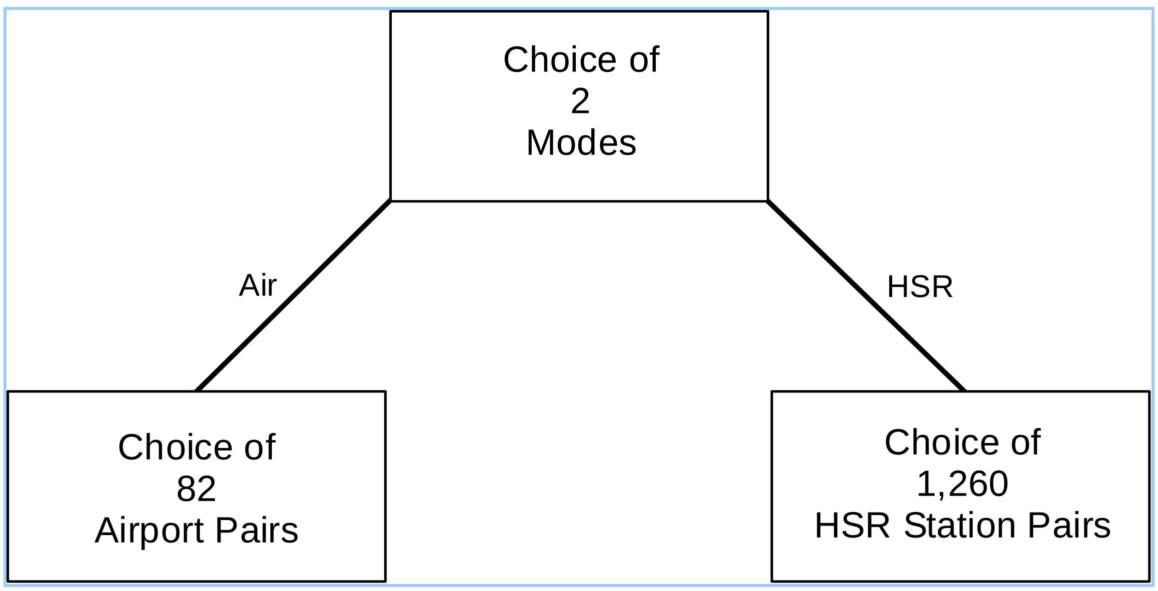
Travelers not only choose between modes, but also between classes of service within that mode. International air travelers can select First Class, Business Class, or coach, in addition to a multitude of fare types, each with different restrictions. Rail travelers' options include several types of train categories. A passenger between New York and Philadelphia can choose between New Jersey Transit's commuter service with very frequent stops, Amtrak's Regional with up to six intermediate stops, the more luxurious and faster Metroliner with no more than 4 stops, and the fastest and most expensive Acela Express which, most of the time, only stops in Newark. The choice of train category is closely related to station pair choice and the feeder service of each station. That means train categories should only be analyzed in conjunction with station pair choice.

Classes of service, or train categories, or fares which would allow to impute the class of service or train category, are almost never included in intercity travel surveys for any of the modes. Our Japanese data set is no exception. It does not even distinguish between conventional and high speed rail. However it contains train station information, and we know which train categories each station is served by. A one-to-one mapping to the three different Japanese HSR train categories is not possible, since many stations are served by more than one category (see the New York to Philadelphia example). What we *can* model is the origin/destination station pair choice. Since it is station pair choice and not train category or class of service choice which is more strongly linked to access and egress considerations, what we *CAN* model is also what we would *WANT* to model.

An air traveler's choice of a particular airport pair is as strongly linked to the feeder service choices as for rail passengers, though the choice set is usually considerably smaller.

These factors establish the basic model structure (Figure 10).

Figure 10 - Basic Model Structure



Choice set selection is covered in Chapter 5.

3.3.4 Linehaul Time

It is self-evident that linehaul time is the most important mode choice determinant when modeling a binary choice between air and high speed rail. All the other attributes remained relatively constant with the introduction of high speed rail in Europe and Japan.

So it was very likely the linehaul time which caused the shift in modal share. This is also the most easily obtainable supply side attribute because of widely available published time tables.

More than 40 years after the introduction of HSR in Japan and almost 25 years after its introduction in Europe certain "universal" laws (at least where Europe and Japan are concerned) emerged. The most important one is that for high speed rail to achieve a dominant share of the combined air/rail market, HSR's linehaul time must be kept below three (3) hours. For corridors up to about 600 km this is not generally a problem since it only requires an average speed of a little above 200 km/h, easily achievable at the usual HSR maximum speed of 300 km/h on newly built lines. The 700 km corridor Madrid – Barcelona, however, had to be designed from its inception for a 350 km/h maximum speed to stay within the critical three hour time frame. This line is under construction now and will be the first HSR line operating above 300 km/h. Note that the distance Madrid – Barcelona is about equivalent to Los Angeles – San Francisco.

JR East's Shinkansen lines east (and north) of Tokyo were built into regions with low population densities, and "the railway cannot afford to share its market with the airlines; it needs to achieve total domination" (Jackson, 2004b). Here it is especially critical to stay within the three hour window and in order to achieve a 3 hour journey time from Tokyo to Aomori (670 km) JR East is working on a target of 360 km/h. The last extension to Hachinohe, opened on 1 December 2002, has been successful. ANA dropped its Tokyo – Misawa service, but the real challenge will be the next extension to Aomori in 2010 (Jackson, 2004b).

The importance of the three hour travel time window, considered such a common wisdom that it is recited in many journal articles about HSR, may indicate that the relationship between HSR linehaul time and market share is not linear. To explore this relationship is beyond the scope of this dissertation. It is on the list of recommended future research.

The three hour window may be psychological, but it may also derive from the fact that in a typical corridor, where HSR is competitive with air (400 km – 700 km), a trip by air takes about three hours door to door, including all security checks. If another mode can offer a downtown to downtown connection within the same time frame, it becomes the preferred mode for a variety of other reasons discussed below, like higher flexibility and/or comfort, etc.

Linehaul time is probably air's strongest competitive advantage. In markets where air competes vigorously with HSR, it tends to defend its edge by offering point to point flights with relatively high schedule reliability as discussed further below.

3.3.5 Fare Levels

Conventional trains with a maximum speed of 200 km/h and an average speed clearly above 100 km/h are competitive with other modes and are able to gain a significant market share as seen in the Northeast Corridor. The New York – Washington air market is much smaller than that of the California Corridor. A part of the reason is that Amtrak takes about 40% of the combined air/rail market in the former, but only about 1% in the latter case. The more budget conscious market segment is served by rail in the

Northeast and by low cost airlines such as Southwest in California. Fares are clearly an important mode choice determinant in the Northeast Corridor, as they are for corridors served by conventional rail in Europe and Asia.

For rail to be truly able to dominate a corridor about 400 km to 700 km in length, as is the case between Tokyo and Osaka or Paris and Lyon, high speed train sets have to operate on newly built, dedicated high speed lines. Between Tokyo and Nagoya the fastest trains achieve an average speed of almost 225 km/h, only 45 km/h or 17% below the maximum speed of 270 km/h (Table 4 on page 63). After the introduction of the new N700 train set with tilt technology the numerous 2500 m-radius curves can be passed through at the full line speed of 270 km/h. In addition the acceleration has been increased to match the standard rate of acceleration of 2.6 km/h/s for commuter trains ("First N700 Shinkansen Train Starts Testing," 2005). At that point trains can run at full speed for almost the complete journey increasing average speed even further. In the extreme case of the Tokyo – Nagoya market (about the same distance as New York – Washington), which airlines have abandoned completely, rail fares are irrelevant. Between Tokyo and Osaka high speed trains still control about 80% of the combined air/rail market. Fares are not very good predictors of mode choice in this case either, as illustrated by the fact that the fare for the fastest train between Tokyo and Osaka, the Nozomi, is higher than the comparable air fare. Fare levels overlap in this market.

As mentioned in the previous section, fares are almost never included in intercity travel surveys. They are also very difficult to estimate since in general travelers have a multitude of fare options to choose from and there is no direct way to impute from available data which one of the fares was the most likely chosen by the trip maker. Additionally, travel surveys are often collected over a period of time which makes it impossible to even determine the fare options available to the particular traveler. Average fares over a certain time period can sometimes be used as a proxy when they are available for each market. However when fares vary over a wide range, as it is usually the case, averages are not very meaningful. What comes as a consolation to the researcher is that in those markets, in which high speed rail enjoys a major competitive advantage, and which therefore account for the vast majority of actual and surveyed rail travelers in the analysis area, fares tend to be not a very important mode choice determinant as stated above.

Average air fares are available to us and are included in the models. An attempt was made to calculate fares between all HSR stations. However it was so highly correlated with linehaul time, that the two effects were indistinguishable. The linehaul time variable for HSR includes rail fares.

3.3.6 Flexibility

The issue of flexibility is very much related to Douglas and Miller's (1974) definition of *stochastic delay*. They defined it as:

> "expected length of delay a potential passenger faces because of the chance that his most preferred scheduled departure will be booked up and he will have to select another and possibly even a third or fourth, and so on."

It would appear that flexibility is an important mode choice determinant because high speed rail has a natural advantage over air making it an important point of distinction between the two modes:

- Trains often do not require advance reservations, and when they do tend to operate at lower load factors than air services, making last minute reservations more feasible; that may be because the marginal cost of putting an extra seat on a train is lower than putting an extra seat on an airplane. The desired train service is less likely to be sold out than the corresponding air service.
- Trains can accommodate standees. Seats may sell out, but space on a train seldom does, so their stochastic delay would be zero.

It is true that the marginal cost of putting an extra seat on a train is lower than putting an extra seat on an airplane for *conventional* trains. But for high speed trains, this competitive advantage is disappearing very fast. While the first generation ICE and TGV train sets still had a power unit at each end, and simply coaches in the middle, newer high speed trains all use *distributed power*. That means different parts of the locomotive are distributed over the entire train set, as with electric multiple units (emus). While first generation ICE's and TGV's were not intended to be separated in normal operation, next generation high speed trains cannot be separated under any circumstances. That means the only flexibility a train operator is left with is to either run one unit by itself, or couple two of them together. Coupling two sets together does not affect overall line capacity, since only one train path is being used. But that is the only economic advantage over airlines on the issue of adapting capacity to demand.

In theory, high speed trains could accept standees, guaranteeing that anybody who absolutely wanted to ride on a particular train would be allowed to do so. Practically, this is only possible in Germany and Japan. With the exception of Germany, and as of very recently, Japan, every high speed train operator requires a reservation just like airlines do.

In Japan customers have always been able to buy and reserve standee tickets. Additionally, since October 1, 2003 three cars in each Nozomi super fast express train were made available for passengers without reservations. Since the travel survey on which our data set is based took place in 1995, the latest special enhancement to flexibility could not be taken into account in our analysis. It is interesting to note that none of the other operators even take advantage of this particular competitive advantage they might enjoy over airlines.

Germany is somewhat of a special case among high speed train systems. The German ICE system, unlike any other, trades off speed against frequency. It offers pulse scheduling (hourly or half hourly) and short average distances between stops (about 90 km) in order to serve every station with high frequency. This, and the fact that HSR lines do not bypass medium size cities as in other countries, causes Germany to have the slowest high speed rail system in the world (Clever, 1996; see Figure 9, page 162). Other European systems feature demand responsive scheduling, low frequencies, but long non-stop segments, resulting in high average speeds. French TGV's especially are like point-to-point airlines offering fast, but often infrequent service. German ICE's are not unlike hub airlines, offering frequent, but considerably slower service. Due to the extreme high demand, Japanese trains are able to offer both high frequencies and high speeds.

The flexibility German Rail affords its high speed travelers in terms of not requiring a reservation and running on a fixed interval schedule results in low *average* load factors.

Table 5 - Average Load Factors of European High Speed Trains

High Speed Train (Country)	Average Load Factor (1992)
AVE (Spain)	66%
TGV (France)	65%
ICE (Germany)	51%

Source: (Jänsch, 1994)

The table above shows that the load factors for the French and Spanish high speed rail systems are very much comparable to those of airlines. That means the desired train service will probably be sold out just as often as the corresponding air service resulting in equal or at least very similar stochastic delays. In summary, high speed trains allowing travelers more flexibility than airlines is a non-issue everywhere except in Germany and Japan, and in Japan for the period of interest (1995) only in the form of reserved standee tickets. The mode choice determinant flexibility was not included in the models for this dissertation.

3.3.7 Frequency

Since the two supply-side attributes frequency and flexibility are so interrelated (combined under the term *schedule delay* by Douglas and Miller), part of the discussion about frequency was already covered in the previous section. We saw that Germany offers the traveler high frequency and high flexibility (in terms of low load factors), but pays the cost of sporting the slowest high speed rail system in the world. It trades off speed to achieve high frequency and, because of the large number of intermediate stops, accessibility. France trades off frequency and accessibility to realize high speed, offering a few but fast non-stop connections between Paris and provincial centers. Japan is in the enviable position to offer all of the above, speed, frequency, and accessibility. Line capacity is continuously being improved, and after the introduction of digital ATC (automatic train control) in 2004/2005, JR Central was able to increase its peak frequency to 13 trains/h – eight Nozomi, two Hikari, and three Kodama (Jackson, 2004a). That means one super fast Nozomi leaves Tokyo or Osaka every 7½ minutes, causing the frequency delay at least in peak hours to be zero.

13 trains/h means one train every 4.6 min which in and of itself would not indicate an extraordinarily efficient use of line capacity. In France, all high speed trains from a particular section of the country (e.g. the southeast) are bundled on the same high speed line approaching Paris. In this bottleneck the railway can handle 20 trains/h running at 300 km/h. That is one train every 3 min. However the French railroad doesn't have intermediate stops to contend with. Studying Table 4 on page 63 a little more closely and only considering the peak hour frequency of eight Nozomi and three Kodama we can see just by inspection that the Kodama trains on their 2:49 h journey from Tokyo to Nagoya are being passed continuously by faster Nozomi, which cover the same distance in only 1:38 h. Notice that this line only has two tracks, with sidings facing outside platforms at each intermediate stop.

It would not be possible to offer such a high frequency with three different train categories without train operators following their schedules extremely closely. This is the topic of the next section.

Airlines are not able to offer such dense frequency due to limited airport capacity. Our level of service database, described in detail in Section 5.2, indicates 29 daily flights between Tokyo Haneda and Osaka Itami airports. Assuming a 16 hour day between 6 AM and 10 PM, this works out to about two flights per hour.

As we shall see later, it was not possible to obtain an exact frequency count for all HSR station and service class (see Section 3.3.3 above) combinations. That means the effect of HSR's superior frequency will be absorbed into the coefficient for the inclusive value of rail in our nested model.

3.3.8 Schedule Reliability

Reliability is an important consideration for the traveler, because the better the on-time performance, the closer the trip maker is able to schedule the arrival time to the actual time he or she needs to be at the destination. If air and rail compete vigorously in a particular market and rail has an excellent reliability, air service will have to approximately match that performance lest its shorter linehaul time advantage will become completely meaningless. And this is what generally tends to happen in those markets.

Before the Japanese National Railroad (JNR) was privatized in 1987, the average deviation from schedule on the Tokaido (Tokyo – Osaka) Shinkansen line was 3.1 min. This improved in the first decade after privatization to 0.7 min. For the following five years this delay fell to 0.4 min. In 2003/2004 the average deviation from schedule, including delays caused by natural disasters such as earthquakes and typhoons, fell to a staggering 0.1 min (6 s). The railway admits that this was mostly due to an absence of natural disasters during that period. After this unusual year the average delay returned to the previous level (Jackson, 2004a).

Except for one or two days a year, when a natural disaster causes the whole system to be delayed by a few hours, driving up the average, trains are operating within a few seconds of schedule. Therefore the average delay of 0.4 min is misleading, and the exceptional mean of 0.1 min in 2003/2004 much closer to the median in all years and more representative of the delays experienced by passengers.

Even if the mean delay of 24 seconds were meaningful, it would still be barely noticeable to the passenger. It comes as no surprise that domestic airlines also have a very good on-time performance. We know for example that in 2000 the domestic carrier All Nippon Airways (ANA) reported 97.7% of all of its flights between Tokyo Haneda and Osaka (Itami and Kansai) were either on time or no more than 15 minutes late. The equivalent figure for Tokyo – Fukuoka was 97.1%. Flight delays are often weather or maintenance related and therefore more uniformly distributed throughout the year.

The corresponding statistic from the Airline Service Quality Performance (ASP) data for all flight arrivals in February 2001 at Los Angeles International (LAX) and San Francisco International (SFO) airports is 60% and 60.3% respectively (*Airport Capacity Benchmark Report 2001*, 2001). To be able to meaningfully compare these different reliability data, we need to assume that the delays are correctly modeled by a Poisson process. Since we only have aggregate level data, it is not possible to verify this assumption.

A random time T has an exponential distribution with rate or parameter λ, if T has the probability density function:

$$f(t) = \lambda e^{-\lambda t} \quad (t \geq 0) \tag{10}$$

with t denoting individual values of random variable T.

Its cumulative distribution function is:

$$P(T < t) = 1 - e^{-\lambda t} \quad (t \geq 0) \tag{11}$$

This enables us to calculate the parameter λ for the schedule reliability statistic of ANA's Tokyo – Osaka flights in 2000:

$$P(t < 15) = 1 - e^{-15\lambda} = 0.977 \tag{12}$$

$$\lambda = 0.25 \tag{13}$$

Figure 11 and Table 6 directly compare the reliability statistics mentioned in this subsection. It is not a major inconvenience for an air traveler between Tokyo and Osaka to add a little more than 15 min to the schedule to insure his or her arrival by a certain time. Notice that delays of up to 15 min are considered "on-time" for many on-time performance measures in the aviation industry (e.g.: U.S. Airline Service Quality Performance or ASQP).

However the one hour and 15 min extra time required by passengers using Los Angeles International Airport would cut the linehaul time advantage of air over a potential high speed rail system by more than one half, significantly reducing its competitiveness. The same can be said about San Francisco International Airport. We could reasonably expect airlines to assign a higher priority to on-time performance in the face of direct HSR competition. At this point, U.S. high speed rail proposals generally do not consider the likely benefit of much improved on-time performance on the part of domestic carriers, not even in the corridor where a high speed rail line would be built.

Because of the consistently high level of schedule reliability for both air and rail in Japan, this supply-side attribute did not have to be taken into account in our mode and terminal choice models.

Figure 11 - Cumulative Distribution Functions of Schedule Delays

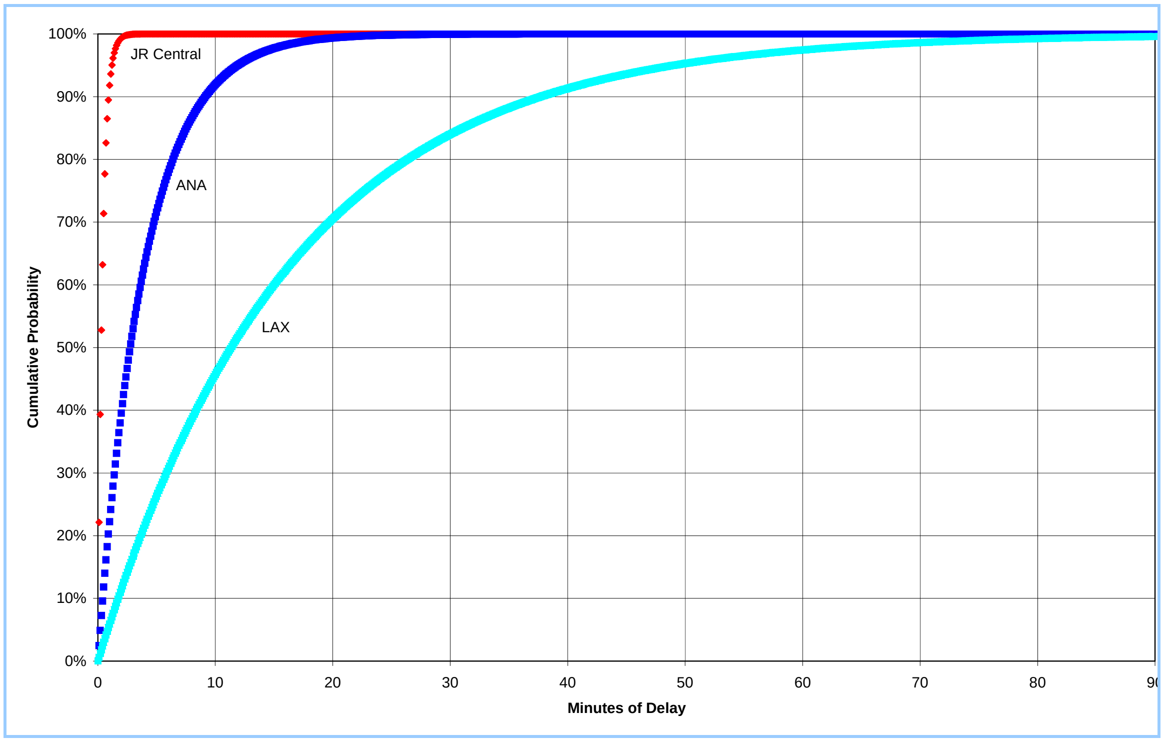


Table 6 - Table of Equivalent Delay Statistics Based on Exponential Distribution

	JR Central	ANA	LAX
λ	2.5	0.25	0.06
$\mu = \frac{1}{\lambda}$	0.4 min	4 min	16 min
Less than 15 min of Delay	100.0%	97.7%	60.0%
99th Percentile Delay	2 min	18 min	75 min

3.3.9 Safety/Fear of Flying

With the sole exception of Germany, where a high speed train set derailed on a conventional line between Hamburg and Hannover, no high speed rail passenger has ever been killed, making it the safest mode available to date. Because of frequent earthquakes Japan has developed a very sophisticated early warning system, which brings high speed trains to an immediate stop before the earthquake shockwave has reached the respective train positions. This system, however, was of no help during an earthquake in 2004 near Niigata which occurred right underneath a train traveling at high speed. The train derailed but remained upright and no passengers were injured. Similarly, a train traveling at 300 km/h derailed between Paris and Lille because a tunnel dating back to World War I had weakened the foundation of the rail bed. Again, the train remained upright and no passenger was hurt. These two events illustrate the resilience of this mode to potentially very dangerous situations.

In the absence of any accident in Japan in which a HSR passenger was injured since HSR's inception in 1964, it may be useful to briefly consider a train crash on Japan's conventional network.

The extreme attention to schedule adherence discussed in the previous section is not only an integral part of high speed rail but also conventional rail operations in Japan. The accident on 25 April 2005 of a suburban commuter train near Osaka focused attention on the high price that may have had to be paid for exemplary on-time performance. Note that JR West also operates the Sanyo Shinkansen line between Osaka and Fukuoka:

> The Ministry of Transport has ordered JR West to draw up a strategy to improve safety and driver training. This follows allegations that JR West reprimands drivers too harshly when they are believed to have caused delays. According to union officials, punishments include pay cuts, demotions, or a long-winded re-education process. JR West says it may change its re-education policy (*Investigation Into Japanese Crash Begins*, 2005).

That train operators were even being fired for what we would consider minor delays was well known before the accident. The facts, that the derailment occurred in a 300 m-radius curve restricted to 70 km/h and that passengers reported unusually high speeds for a train which was running late, called this policy into question. The train was running 90 s late because the driver had overshot the previous station by 40 m, forcing him to back up to the platform. At the time of the accident the train was traveling at a speed of 108 km/h (Briginshaw, 2005). Please note that accident investigations are still ongoing at the time of this writing. After this disaster a change in operating procedures, which will also affect the HSR network, is very likely.

When considering safety issues one must also pay attention to the possible captivity effect caused by fear of flying. Few people would not consider flying under any circumstances. Most people for whom fear of flying is an issue have some kind of trade-off between fear and travel time. To determine what this trade-off is and to what extent travelers are captive to a surface mode in a 300 – 700 km corridor with a high speed rail option is beyond the scope of this research. McCarthy found a likelihood of captivity for intercity car travel of 48%, while the likelihood of captivity for rail and bus was less than 1% (please see McCarthy (1997) beginning on page 48). We do not expect captivity resulting from fear of flying to be significant and did not consider it for our choice models.

3.3.10 Socio-economic Variables

Section 3.2.3 already discussed socio-economic variables in the context of feeder related mode choice determinants. One issue remains to be discussed that is not directly related to access or egress.

Morrison and Winston emphasized that any specification of an intercity demand model should attempt to capture the effect of the number of household members on a trip (as well as trip distance and household income), since these variables had the potential to contribute significantly to the predictive power of the model. The need to include number of household members on a trip was reiterated by David Hensher and Mark Wardman. Hensher pointed out that the appeal of car travel is a single cost, independent of party size (please see Morrison & Winston (1985) p. 39, and Hensher (1997), p. 50).

The number of household members on a trip is extremely important when the choice set includes travel by private automobile. When modeling the choice between air and High Speed Rail this variable is of lesser importance. As further explained in Chapter 5, the automobile data of the 1995 Japanese Intercity Travel Survey are not suitable for discrete choice modeling. Moreover, the size of the travel party is not part of the air and rail data set either. We do have the average household size in the home zone of each traveler. Since this would only be a very imperfect substitute for the number of people traveling together on that particular trip in the sample, the mode choice determinant number of household members on a trip was not included in our analysis.

3.3.11 Comfort

Train seats are generally more spacious and allow for a variety of activities, whereas air economy seats tend to be too restrictive to allow for much more than reading a book or a magazine. This is very much related to the concept of "lost time" discussed previously on page 29.

Even though the 9/11 investigations would like to have us believe that cell phone calls from airplanes are possible (at least over Pennsylvania), they are not (von Bülow, 2003, pp. 122-123). Technology that would allow air passengers to use their cell phones in flight is only now being developed. On the other hand, HSR passengers can remain in constant contact with their office or clients and even long conversations are economical. This adds to the fact that time spent on a train can be used more productively than time spent on a plane.

It is fairly easy to test whether travelers consider one hour spent on a train as less onerous than one hour spent on a plane, as long as the variable linehaul time is specified to be alternative specific in a binary mode choice model. In other words, the effect of comfort is already being picked up in our model, and we do not have to give it additional consideration.

3.3.12 Intensity of Competition

Due to capacity constraints at airports airlines may not be competing vigorously with HSR, which would mean the high market shares for rail are misleading. This certainly appears to be the case in the Paris – London market. British Airways is a major investor in the new high speed line from the channel tunnel to London. It decided to contribute 10% to the first phase, and 34.8% to the second phase (Machefert-Tassin, 1999, p. 45). Like other European airlines, BA wants to use its scarce landing slots for more profitable long range flights. Partnerships between airlines and railroads are very common in Europe. On the other hand, Jombard (2000) reports that one of the challenges Eurostar had to overcome in its 5 year history was a more difficult than anticipated competitive situation due to low cost airlines.

Airport capacity in Europe (and Japan) is extremely restricted. The result of that is that airlines most often focus on how high speed rail services could *complement* their own schedules. Speaking in terms of broad generalizations, while American carriers prefer to offer frequent service with relatively small equipment (e.g. B-737), European carriers usually offer less frequent service with larger equipment (e.g. A-300). The regulatory context in addition to airport capacity restrictions encouraged them to do so. "Competing" carriers in a market were allowed to *pool* all revenues collected in that particular market and then distribute them amongst each other in relationship to the capacity offered. But even after deregulation capacity restrictions kept the European air transportation system from becoming very similar to the U.S. air transportation system.

There are many examples of cooperative agreements between rail operators and airlines in Europe. The first one or two coaches of 10 daily Thalys trains between Brussels and Charles de Gaulle airport in Paris are reserved for Air France passengers. Those Thalys trains also have Air France flight numbers and replace the previous 10 round trip flights between Brussels and Paris. A similar service is available between Frankfurt airport and Stuttgart Central Station on the hourly InterCityExpress trains, which made short haul flights in that market redundant. Code share agreements are only possible because of the integrated rail stations at the airports (Sharp, 2005).

Contrary to Europe there is no direct high speed rail connection to any of the Japanese airports, and there are no cooperative agreements between rail operators and airlines in Japan. As mentioned above, airlines were able to increase their share of the combined Tokyo – Osaka air rail market from 14% to 20% over 5 years and railway officials complain about unfair competition. Also, JR Central, which operates the HSR service in the Tokaido (Tokyo – Osaka) corridor, and JR West, responsible for the Sanyo (Osaka – Fukuoka) corridor, are both privatized, independent from the Japanese government, and highly profitable (Jackson, 2004a, 2004c).

In summary, there are no indications that air and rail are not competing vigorously against each other in Japan. Market distortions like subsidies or non-competition agreements do not appear to affect terminal or mode choice decisions. This is another advantage of using Japanese versus European data.

3.3.13 Number of Transfers

This topic was covered in detail in the literature review and also briefly in the context of feeder related mode choice determinants. In addition to the American and European studies already cited beginning on page 26, which quantified the disutility of a transfer in the middle of a long-distance trip (M. Hansen, 1990; Infratest Sozialforschung, 1992; Vilmart & Paix, 1994; Vrtic & Axhausen, 2003), we want to draw the reader's attention to similar experiences in Japan.

Due to policy differences between JR West and JR Central until October 2003, many of the Nozomi and Hikari services terminated in Osaka. That meant many passengers who wanted to travel between a station on the Tokaido and a station on the Sanyo line had to change trains in Osaka. JR West reports that it was able to increase its market share on the Sanyo corridor by 10% in the 4 years since the introduction of improved Hikari service in March 2000. Due to the policy differences with JR Central, however, that success could not be repeated for services to the capital. These problems were overcome with the opening of the Shinagawa station and the time table change in October 2003. Not only were more through services offered between Nozomi stops on the Tokaido and Nozomi stops on the Sanyo lines. More importantly, several stations west of Osaka, which were formerly served only by local Kodama and semi-fast Hikari now received through Nozomi services to Tokyo. As an example, Shin-Yamaguchi – Tokyo business increased by 30% in the first 6 months because of "passengers' natural disinclination to change trains en route when traveling with heavy baggage," and the one hour journey time savings compared to the previous semi-fast Hikari (Jackson, 2004c).

JR West was willing to break the pattern of fast trains only stopping at certain stations (see Table 4 on page 63) mainly because of the ridership penalty of a transfer in the middle of a long distance trip.

As described in more detail in Chapter 5, the number of transfers required for each HSR station combination and service class was calculated and became an integral part of our model development. However, transfers for air travel were not considered at all, because in our survey data air trips requiring a transfer were never chosen. This again is described in more detail in Chapter 5.

If we looked at all the possible origins and destinations (O and D’s) in the United States, we would find most O and D combinations had no direct air service. But if we allowed one transfer, the number of O and D's with air service would increase dramatically due to the hub and spoke system of most airlines. We can surmise that there is no hub and spoke system for domestic air travel to speak of in Japan in our area of interest on the main island of Honshu and on the smaller island of Kyushu. That is probably due to the fact that flight connections through hubs are not only unattractive for short distances, but also for medium distances like Tokyo – Fukuoka (comparable to New York – Atlanta) when the competing HSR mode only takes 5 hours.

A domestic hub and spoke system in Japan is not only undesirable, it also largely infeasible. Since most of the daily flights at a hub airport operate within a few relatively short time windows, the space requirements are so extensive (Table 7), that building a new airport of that size in Europe or Asia is only possible in exceptional circumstances, as e.g. in Hong Kong (Clever, 2005). Please note that major international hubs like Osaka Kansai have only one runway. Tokyo Narita did not open its second runway until 2002, and that second runway is barely over 2000 m long.

Table 7 - Space Requirement of Hub Airport Compared to Area of German Cities

	Population	Area [km²]
Siegen	108,397	114.67
Osnabrück	164,195	119.08
Denver International Airport		**135**
Krefeld	239,559	137.73
Bonn	306,016	141.22

Source: (*Gemeindeverzeichnis*, 2002) Population as of 31 Dec 2001

3.3.14 Other

No mode choice model, no matter how sophisticated, could capture all attributes that contribute to a terminal or mode choice decision. In order to capture an inherent bias that is not picked up by any other variable, mode choice models need mode specific constants. To model an inherent preference or dislike for a particular airport or station, terminal pair choice models need fixed effects. Both type of models may benefit from other geographic fixed effects, which in our example may include prefecture or city.

There are certain airports or HSR stations which for a large variety or reasons are perceived by travelers as being more attractive than other airports or stations. That may be because of special membership lounges, shorter than average wait times when clearing customs, shorter distances from the gate or platform to the curb, a more modern or cleaner surrounding, better ancillary services like bookstores or restaurants, or simply because they are easier to navigate for the traveler because he or she is more familiar with them. What all these attributes have in common is that their effects on the traveler's decision are not easily picked up by other variables in the model. Since we cannot simply expect all those effects to be orthogonal to our other variables, we need to include fixed effects to pick up these effects and prevent them from influencing the estimation of the other variables, thereby biasing them.

3.4 Summary

3.4.1 Mode Choice Determinants Used in this Research

Table 8 lists all the mode choice determinants discussed in this chapter. It also indicates which ones are explicitly used in our model estimation.

Table 8 - Variables Used Explicitly in Model Estimation

	Air	HSR
Access Distance	yes	yes
Egress Distance	yes	yes
Access Time		
Egress Time		
Access Fare		
Egress Fare		
Classes of Service		
Linehaul Time	yes	yes
Fare Levels	yes	
Flexibility		
Frequency	yes	
Schedule Reliability		
Safety/Fear of Flying		
Socio-economic Variables	yes	yes
Comfort		
Intensity of Competition		
Number of Transfers		yes
Fixed Effects	yes	yes

We now have all the individual components (variables) for the estimation. The Literature Review of the previous chapter helped greatly in this selection. Before we continue with the model specification in the following chapter, a brief summary of the factors and their interrelationships which helped create Japan's passenger transportation system is presented in the next section.

3.4.2 Japan's Secret: Few Trade-offs

We are now in a position to summarize the "unique convergence of factors," referred to at the beginning of this chapter, which explain Japan's exceptional passenger transportation system.

At the end of Section 3.2.2 (page 79) we mentioned that while an individual train needs to trade longer linehaul time for better accessibility due to the extra stops required, the high speed rail system as a whole will usually trade frequency for better accessibility. Finite demand only supports a fixed number of hourly trains. If one of those trips serves intermediate stations, it cannot be used for non-stops. A good example of this kind of HSR operation is France with a limited number of fast non-stop connections between Paris and provincial centers. At the opposite end of the spectrum, Germany trades speed for accessibility and blankets the country with frequent multi-stop intercity trains which are almost operated as an above ground subway system.

The uniqueness of Japan's transportation system stems from the fact that two thirds of its population, or almost 100 million people, live in a narrow, densely populated corridor along the south shore of Honshu island between Tokyo and Fukuoka. This corridor is ideally suited for rail operations. Due to the extreme economies of density, the Japanese HSR system in the Tokaido and Sanyo corridors does not need to trade frequency for accessibility. With extremely high schedule reliability it can also operate several train categories on the same track, not having to trade speed for accessibility.

We can now see how the different supply-side attributes discussed so far are interrelated:

1. Japanese railway companies are in the enviable position to be able to offer high **speed**, high **frequency**, and high **accessibility** due to very strong demand for their services in a narrow densely populated corridor ideally suited for rail operations.

2. Since their customers experience no schedule delay and can enjoy very fast downtown to downtown connections, **fare** levels for HSR overlap with those for air in highly competitive markets like Tokyo – Osaka.

3. Offering super-fast Nozomi service every 7½ min in peak hours while at the same time operating two other train categories with a much larger number of intermediate stops requires severe attention to **schedule reliability**.

4. This however could lead to a **safety** problem in the future, as the recent commuter train accident has illustrated.

Japanese railway companies, of course, do not consciously trade off safety in order to achieve high speed, frequency, and accessibility simultaneously, although the recent accident seems to indicate that they may have pushed the safety envelope a little too far. However, what is being compromised is **capacity**. 13 trains/h are noticeably fewer than 20 trains/h achieved in France. The Tokaido line has been consistently operating at capacity and the options for a capacity increase (better signaling system, higher speed, etc.) have almost been exhausted. Airports operate close to capacity and expansion possibilities are limited. Domestic airlines already employ the largest equipment available, 747's with high density seating. This is the reason for the Railway Technical Research Institute www.rtri.or.jp conducting research on super-conducting maglev technology at the Yamanashi test track, which will eventually become an integral part of the Chuo maglev line between Tokyo and Osaka.

In order to compete with potent ground transportation, airlines focus on point to point flights. A large domestic hub and spoke system would neither be desirable nor feasible. Another strategy for airlines to maintain their linehaul time advantage is to minimize schedule delays.

4 Model Structure

4.1 A Multi-Disciplinary Approach

As briefly mentioned in Subsection 3.3.3 Classes of Service on page 89 we will calibrate terminal pair choice models. These then form the basis for upper nest mode choice model estimation.

Various disciplines, especially Mathematics, Statistics and Econometrics, have grappled with the challenge to suitably model data with discrete dependent variables. While in the final analysis the approaches used are fairly similar, it is surprising to note that the three disciplines developed their solutions independently and during different time periods, each one focusing its attention on a different subset of problems. So it comes as no surprise that a statistician well versed in the application of generalized linear models will have never heard of nested logit models, while an econometrician with the corresponding depth of knowledge would be uncomfortable with the concept of a link function. Each discipline attacks the basic challenge (Figure 12) from a different angle, so each discipline provides valuable insights leading to the resolution of the questions posed in this dissertation.

Ever since Gottfried Wilhelm Leibniz (1646-1716) renaissance men or women have been either rare or non-existent (Anton, 1992, p. XXVII). While some readers might see their patience tested by certain details in this chapter, for the same reason will the large majority find material exposing a fresh perspective or new “camera angle” leading to new ways to confront challenging questions in the future. The main benefit likely to be gained will be new ways of interpreting estimation results.

A deceptively simple approach to modeling data with a binary response is the linear probability model (Figure 12):

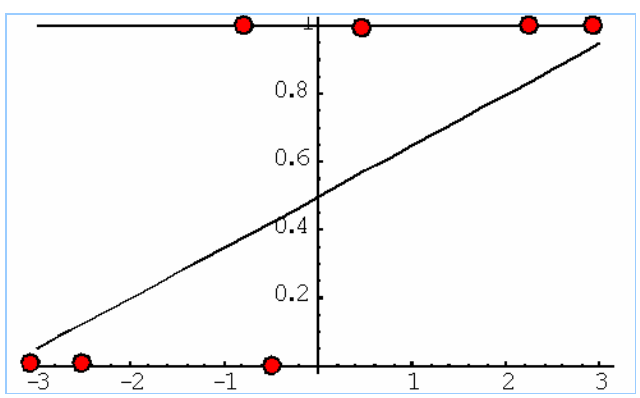
$$\pi(x) = \alpha + \beta x \tag{14}$$

where $\pi(x)$ denotes the probability of success, e.g. of making a sale, dependent on the explanatory variable x. The parameter β corresponds to the increase or decrease in probability per unit change of x. α is the intercept.

Figure 12 illustrates how inadequately a simple linear probability model represents data with binary responses. The linear probability model has two major flaws:

1. The range of its function is not bound by 0 and 1 as probabilities are, leading to nonsensical predictions outside this range.
2. Its underlying assumption that a unit increase/decrease in x leads to a constant *absolute* increase/decrease in probability seems unrealistic. We would expect the probability to increase/decrease less and less as it approaches 0 or 1 (Agresti, 1990, p. 74-75).

Figure 12 - Linear Probability Model for Data with Dichotomous Response Variable



Mathematics offers the most general solution to related problems and its approaches are presented first. Statistics then focuses on a special case of the logistic equation developed by mathematicians and integrates it into the unifying framework of generalized linear models designed to solve a wide variety of statistical problems. Econometrics adds a behavioral component and applies it to a limited subset of generalized linear models best suited to its data and research questions.

4.2 Mathematical Origins of the Logistic Curve

4.2.1 Graphic Derivation

Applying the reciprocal to a set of real numbers between 1 and infinity yields a new set of real numbers bound by 0 and 1. That is the basic idea behind the graphic derivation of the logistic curve depicted in Figure 13.

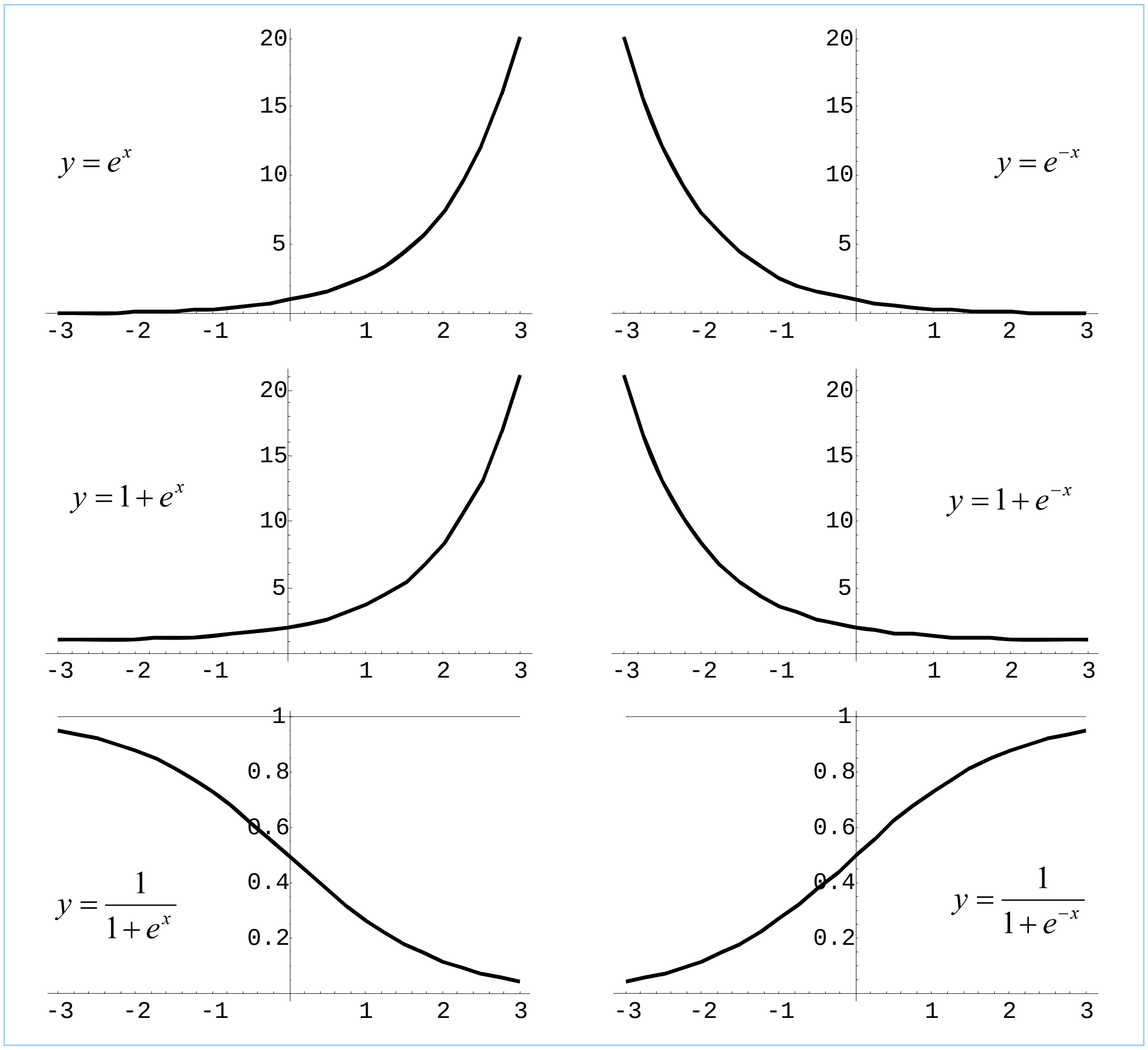


Figure 13 - Graphic Derivation of the Logistic Curve

Result:

$$y = \frac{1}{1+e^{-x}} \tag{15}$$

The graph of the function $y = \frac{1}{1+e^{-x}}$ is commonly known as the S-Curve. It enjoys numerous applications especially in disciplines like marketing. This curve succinctly replicates the three stages of the product cycle. During the market introduction phase great efforts are expended yielding relatively meager results in terms of market share increases. However once the new product has "caught on" additional promotions cause a disproportionately large increase in market penetration. In the final market saturation phase new advertising campaigns effect successively decreasing returns on investment.

In aviation the S-Curve realistically models how airlines with a low share of the total number of flights in a given market will achieve a less than proportional market share while the opposite is true for airlines on the other side of the spectrum. This explains the airlines' propensity to err on the side of too many flights and to "overschedule," resulting in frequent delays and flight cancellations.

Marketers follow a surprisingly similar strategy. With the average consumer seeing about 1 000 000 marketing messages a year the only way to break through the clutter, in other words move to the right side of the inflection point, is to "overspend" (Godin, 1999, pp. 29 and 36).

4.2.2 The Logistic Equation

The Logistic Differential Equation came into being as a generalization of the law of natural growth (or decay) modeling population growth.

The differential equation:

$$\frac{dN}{dt} = rN \tag{16}$$

where $\frac{dN}{dt}$ denotes the instantaneous rate of change of population size N with respect to time t, and r the rate of growth, is based on the simple premise that the rate of change of N is proportional to the current value of N. r is therefore also known as the constant of proportionality. (16) describes exponential growth. Its general solution is

$$N(t) = Ce^{rt} \tag{17}$$

where C is an arbitrary constant.

Since ideal conditions cannot continue indefinitely, it is imperative to relax the assumption of indefinite growth. In the modified equation

$$\frac{dN}{dt} = f(N)N \tag{18}$$

the constant r is replaced by a function of N. We want the growth rate to depend on the population size in such a way that $f(N) \cong r$ for small N, that $f(N)$ diminishes as N grows larger, and, in case of the initial population to be larger than its sustainable size, the growth rate to be negative. This is achieved with

$$\frac{dN}{dt} = r\left(1 - \frac{N}{K}\right)N \tag{19}$$

where K symbolizes the saturation level or environmental carrying capacity. r is renamed intrinsic growth rate, i.e. growth rate without limiting factors. As the population size approaches its limit K, the population in essence ceases to grow, or, in case of a negative growth rate, bottoms out.

(19) is one of several equivalent forms of the Logistic Equation, which models what is known to be *logistic* as opposed to *exponential* growth. Its general solution is

$$N(t) = \frac{r}{K^{-1} + Ce^{-rt}} \tag{20}$$

We can see by inspection that the functional form of the logistic curve (15) is a special case of (20), where the intrinsic growth rate r, the saturation level K and the arbitrary constant C are all equal to 1.

(19) is sometimes referred to as Logistic Differential Equation to distinguish it from the analogous Logistic Difference Equation

$$N_{t-1} = r\left(1 - \frac{N_t}{K}\right)N_t \tag{21}$$

where N_t and N_{t-1} stand for the population size at time t and $t\text{-}1$ respectively.

Many phenomena are modeled more suitably with discrete rather than continuous models. An example is the population growth of a species that propagates only at special times during the year.

The British economist Thomas Malthus is credited as being the first researcher to have observed that many species increase at a rate proportional to their population size. His first paper on population was published in 1798.

The Belgian mathematician and biologist P.F. Verhulst introduced (19) in 1838 to model human population growth. He coined the term logistic growth, hence the name

Logistic Equation. The theory was empirically verified in the first half of the 20^{th} century with fruit fly and flour beetle populations (Boyce & DiPrima, 1992, pp. 52-58) (see also Anton, 1992, p. 593).

4.2.3 The Underlying Differential Equation

The perhaps most important reason to estimate consumer choice models is to be able to determine the extent by which the probability changes in response to a variation in one of the observed factors. We are interested in the partial derivative of the choice probability of alternative *i* for individual *n*, P_{in}, with respect to attribute *k* of alternative *i* for individual *n*, x_{ink}. We wish to evaluate $\frac{\partial P_{in}}{\partial x_{ink}}$.

Looking at logit models from the perspective of differential equations gives us valuable insights into the most important features of the model and greatly simplifies the calculation of derivatives.

Adapting the Logistic Differential Equation (19) to describe a choice probability, i.e. $K = 1$, dependent on a single explanatory variable *x*, and assuming the intrinsic growth rate *r* to be 1, we obtain:

$$\frac{dP}{dx} = P(1-P) \tag{22}$$

This is the underlying differential equation of all logit models used in statistics and econometrics. Probability changes fastest when $P = 1 - P = \frac{1}{2}$. As either *P* or *1-P* close in on 0, the probability largely stays the same in the face of sizeable variations of the independent variable x. Intuitively, the effect of a change in the predictor variable is largest at the highest level of uncertainty, when both P and $1-P$ are about ½. When the choice is almost certain, adjustments to the explanatory variable produce very little result (Train, 1986, p. 38).

(22) summarizes the salient characteristic of the logistic curve which gives it its distinctive sigmoid shape.

The general solution of (22) is

$$P = \frac{1}{1+e^{-x}} \tag{23}$$

This is the simplest case pictured by the 2-dimensional graph in Figure 13 on p. 120.

To generalize (22) we only need to apply the chain rule.

$$\frac{\partial P_{in}}{\partial x_{ink}} = \frac{\partial P_{in}}{\partial V_{in}}\frac{\partial V_{in}}{\partial x_{ink}} = P_{in}(1-P_{in})\frac{\partial V_{in}}{\partial x_{ink}} \tag{24}$$

V_{in} is the systematic utility of alternative i for decisionmaker n. Utilities U are used by economists to measure preferences among bundles of goods. The preference between air and rail for decision maker n might be modeled as simply as

$$U_{An} = \beta_1 x_{An1} + \beta_2 x_{An2} + \varepsilon_{An} \tag{25}$$

$$U_{Rn} = \beta_1 x_{Rn1} + \beta_2 x_{Rn2} + \varepsilon_{Rn} \tag{26}$$

where subscript A denotes air and subscript R rail and x_{An1} could be the travel time by air and x_{An2} the travel cost by air. Many more factors will influence the consumer's mode choice decision which we are unable to quantify for this model. Generally speaking the error term ε absorbs random errors and unobserved characteristics that affect the choice. We can partition utility U into a systematic or deterministic component V and a random element ε.

$$U_{An} = V_{An} + \varepsilon_{An} \tag{27}$$

$$U_{Rn} = V_{Rn} + \varepsilon_{Rn} \tag{28}$$

where

$$V_{An} = \beta_1 x_{An1} + \beta_2 x_{An2} \tag{29}$$

$$V_{Rn} = \beta_1 x_{Rn1} + \beta_2 x_{Rn2} \tag{30}$$

Decision maker n will prefer air over rail if

$$V_{An} + \varepsilon_{An} > V_{Rn} + \varepsilon_{Rn} \tag{31}$$

The decision maker maximizes his or her utility based on the U_{in}. The decision maker's choice is deterministic. The researcher can only maximize utility based on V_{in} and assign probabilities to various choices.

The inequality (31) is the basis of all discrete choice models. This chapter only considers logit models, and in those models the error term ε is assumed to be distributed Gumble. The error term distribution will be discussed in more detail in Subsection 4.4.1 below.

Now it becomes self evident why the partial derivative of the choice probability of alternative i by decision maker n with respect to attribute k of alternative i for decision maker n is of particular interest. Recalling the short derivation of Equation (24) via the chain rule, the final result to remember becomes very intuitive:

$$\boxed{\frac{\partial P_{in}}{\partial x_{ink}} = \frac{\partial V_{in}}{\partial x_{ink}} P_{in}\left(1 - P_{in}\right)} \tag{32}$$

Economists often prefer elasticities over derivates because they are unitless. To state the elasticity of the choice probability P_{in}, with respect to attribute k of alternative i for individual n, x_{ink}, we only need to add one term

$$\mathrm{E}_{x_{ink}}^{P_{in}} = \frac{\partial V_{in}}{\partial x_{ink}} P_{in}\left(1 - P_{in}\right)\frac{x_{ink}}{P_{in}} \tag{33}$$

which simplifies to

$$\mathrm{E}_{x_{ink}}^{P_{in}} = \frac{\partial V_{in}}{\partial x_{ink}} x_{ink}\left(1 - P_{in}\right) \tag{34}$$

4.3 Statistical Approaches to Categorical Data Analysis

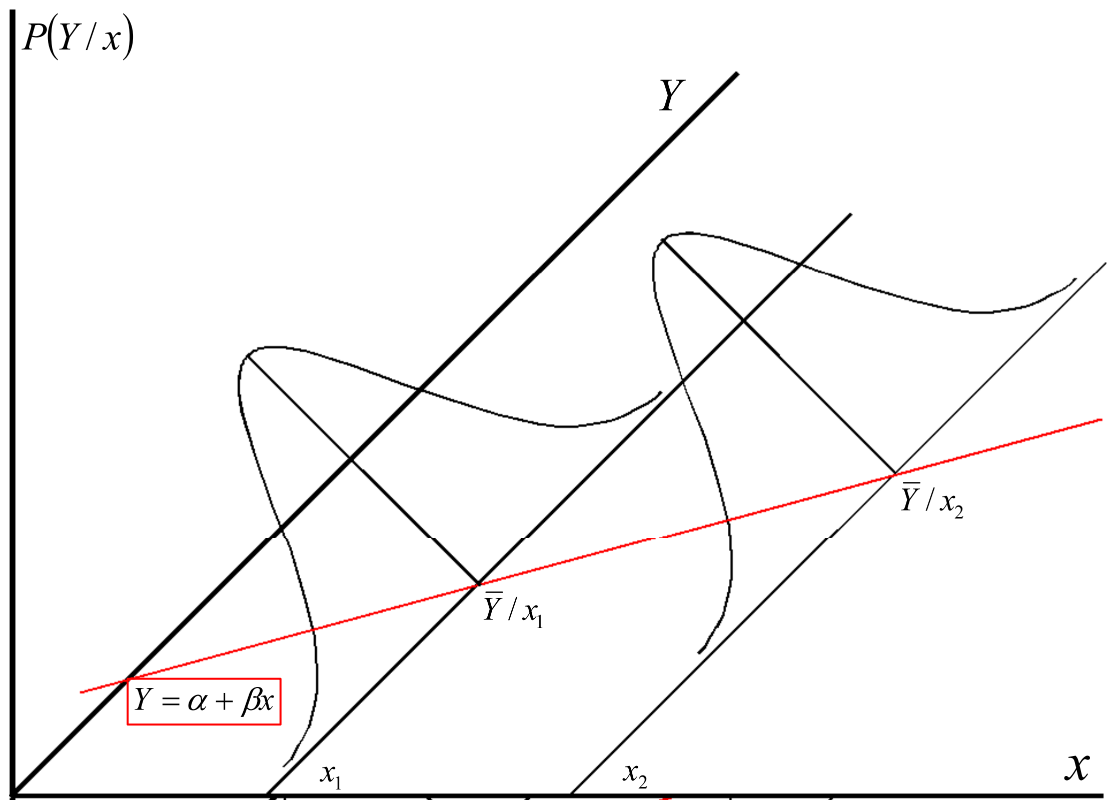
4.3.1 Generalized Linear Models

Figure 14 depicts a classical normal regression model (Wonnacott & Wonnacott, 1979, p. 24). For each value of the covariate x, the observed values of Y are assumed to be normally distributed around the mean on the true regression line. While their means differ, we presume their variance to be the same. To emphasize the fact that we are estimating the mean conditional on x, we write the equation for the true regression line as:

$$\mu(x) = \alpha + \beta x \tag{35}$$

In this special case we are modeling the mean $\mu = E(Y)$ directly.

Figure 14 - Classical Normal Regression Model



In the following discussion μ will always refer to the mean of a random variable.

When modeling probabilities we cannot model the mean directly with a linear combination like $\alpha + \beta x$ because that would be inconsistent with the laws of probability. We can only model a transformation of the mean $g(\mu)$ which maps the unit interval onto the whole real line $(-\infty, \infty)$ (McCullagh & Nelder, 1998, pp. 107-108). One such transformation is the logistic function.

$$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right) \tag{36}$$

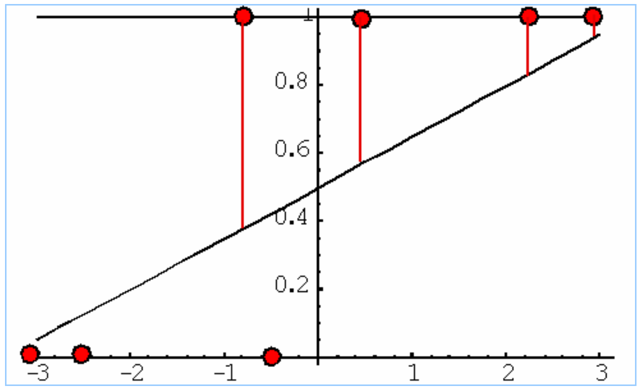
Probabilities are bound by 0 and 1, odds or odds ratios like $\frac{0.75}{0.25} = \frac{3}{1} = 3:1$ can assume any nonnegative value, and the log of the odds ratio, or the logit, can take on any real value. We cannot model probabilities or proportions directly with a linear model, but can model the log of the odds ratio:

$$\log\left(\frac{\mu}{1-\mu}\right) = \alpha + \beta x \tag{37}$$

In ordinary regression (Figure 14) the random component is assumed to have a normal distribution with fixed variance, and we are estimating the mean directly. Generalized linear models (GLM) relax the distributional assumption and allow the modeling of functions of the mean. Both are important generalizations for categorical data.

Figure 15 illustrates that the assumption of a constant variance cannot be maintained when modeling probabilities. The variance is dependent on x.

Figure 15 - Nonconstant Variance with Categorical Responses



Suppressing observation subscript n and alternative subscript i a GLM is of the form:

$$g(\mu) = \alpha + \beta_1 x_1 + \ldots + \beta_k x_k \tag{38}$$

All generalized linear models have three components (Agresti, 1996, pp. 72-73):

1. a random component
2. a systematic component, and
3. a link function.

Let $Y_1, \ldots, Y_N$ denote the individual observations of the response variable Y in a sample of size N. The response variable Y is called the *random component*. Commonly used distributions of the response variable Y are Gaussian for continuous observations (Figure 14), Poisson for non-negative counts, and binomial for binary response data. Note that binary responses are also called Bernoulli variables to emphasize their Bernoulli(π) distribution where π is the probability of success. For n independent observations with binary response the number of successes are distributed binomial(n,π).

The randomness is seen in the sample. Let us assume a subset of *n* travelers in a survey are identical in all observable parameters. Let us further suppose that these trip makers have probability $\pi = 0.8$ to choose air and probability $1 - \pi = 0.2$ to take HSR. The response variable Y_n indicating whether an individual decision maker within this group chose air or HSR is distributed Bernoulli(0.8). The number of times air was chosen within this subset is distributed binomial(n, 0.8). The value of π can change as the observable parameters change. We want to model $\pi(x)$.

The *systematic component* specifies the linear combination of the explanatory variables.

The *link function* maps the constrained response variable, e.g. a probability, to the real line, making it possible to estimate linear models. The logit link function was introduced in Equation (36). Other commonly used link functions are the identity link $g(\mu) = \mu$ for ordinary regression models (35), and the log link. Loglinear models are employed when μ cannot be negative as e.g. with count data. In this case the random component is assumed to be distributed Poisson:

$$\log(\mu) = \alpha + \beta_1 x_1 + \ldots + \beta_k x_k \tag{39}$$

To summarize, the random component is the response variable *Y*. What we wish to model is its mean μ. In the case of ordinary linear regression we can model the mean μ directly with a linear combination of explanatory variables. With categorical response variables, however, we can only model a function of the mean. This function of the mean is called the link function because it links the random and the systematic components of a generalized linear model (GLM).

Each probability distribution has one special function for the mean which is called its natural parameter. The natural parameter for the Gaussian is the mean itself, for the Poisson the log of the mean, and for the binomial the log of the odds ratio. A link function using the natural parameter of the distribution is called the canonical link (Table 9). Many link functions are possible, but canonical links are by far the most popular (Agresti, 1996, pp. 72-74).

Table 9 - Commonly Used Generalized Linear Models

Random Component	Canonical Link Function	Appropriate Data
Gaussian	$g(\mu) = \mu$	continuous
Poisson	$g(\mu) = \log(\mu)$	counts
Binomial	$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$	probabilities

Nelder and Wedderburn (1972) first formulated generalized linear models. An important realization was that many of the “nice” properties of the normal distribution apply to the whole family of exponential distributions. This allowed the application of methods so far limited to ordinary linear regression to a much wider set of problems including categorical responses and/or discrete explanatory variables. Besides important theoretical insights the practical benefit gained was that the same computational tools could be used in many different situations (Dobson, 2001, p. 43). The model fitting algorithm is the same for any GLM. Consequently, GLM software can fit a wide variety of useful models.

As ordinary linear regression generalizes the mean, generalized linear models generalize ordinary linear regression.

4.3.2 Logistic Regression

Following a common convention in statistics we will use π or $\pi(x)$ instead of μ when the population parameter denotes a proportion or a probability.

One of the two fatal flaws of the Linear Probability Model shown in Figure 12 on p. 118 is the constant *absolute* increase in probability. In order for Δx to achieve a constant *relative* increase in probability, we have $\log(\pi)$ linearly depend on x (Wonnacott & Wonnacott, 1979, pp. 132-133).

At the lower end of the scale, where $\pi \approx 0$, changes in $\log(\pi)$ dominate changes in $\log(1-\pi)$. E.g., when π doubles from 0.01 to 0.02 $\log(\pi)$ increases by 70% while $\log(1-\pi)$ only changes by 1%. So we specify:

$$\log(\pi) \cong \alpha + \beta x \quad \text{when } \pi \text{ is small} \tag{40}$$

Similarly, for the upper end of the scale or large π we use:

$$\log(1-\pi) \cong \alpha + \beta x \quad \text{when } \pi \text{ is large} \tag{41}$$

Combining both ends with a minus sign so both terms increase when π increases:

$$\log(\pi) - \log(1-\pi) = \alpha + \beta x \tag{42}$$

$$\boxed{\log\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta x} \tag{43}$$

(43) is called the logistic regression function.

The log of the odds ratio is known as logit:

$$\boxed{\operatorname{logit}(\pi) \equiv \log\left(\frac{\pi}{1-\pi}\right)} \tag{44}$$

Analogously to linear regression, when $\operatorname{logit}(\pi)$ is dependent on several independent variables it is referred to as multiple logistic regression.

Solving the logistic regression function for π we first obtain an expression for the odds ratio. We will refer back to this equation because it leads to an important interpretation of logistic regression results:

$$\frac{\pi}{1-\pi} = e^{\alpha+\beta x} \tag{45}$$

$$\pi = e^{\alpha+\beta x} - \pi e^{\alpha+\beta x} \tag{46}$$

$$\pi\left(1 + e^{\alpha+\beta x}\right) = e^{\alpha+\beta x} \tag{47}$$

$$\pi = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}} \tag{48}$$

$$\boxed{\pi = \frac{1}{1 + e^{-(\alpha+\beta x)}}} \tag{49}$$

(49) is easily recognized to be of the same form as (15), the function describing the S-Curve in Figure 13. (49) is the inverse function of (43) and vice versa.

As early as 1938 Ronald A. Fisher and Frank Yates suggested $\log\left(\frac{\pi}{1-\pi}\right)$ as one of several possible transformations for analyzing binary data. In 1944, Joseph Berkson, a medical doctor and statistician, first used the term “logit” for this transformation. He also showed that the logistic regression model fitted similarly to the already well-known probit model (Agresti, 1996, p. 261).

4.3.3 Cumulative Distribution Functions

Let X be a random variable, x one potential value for X, and $P(X \leq x)$ denote the probability that the value of the random variable X is less than x. The cumulative distribution function (cdf) for X, $F_X(x)$ is defined as:

$$F_X(x) = P(X \leq x) \qquad -\infty < x < \infty \tag{50}$$

These functions often have the appearance of the S-curve which would suggest a new class of models for binary responses where

$$\pi = F_X(x) \tag{51}$$

or

$$F_X^{-1}(\pi) = x \tag{52}$$

The logistic distribution fits into this category. It has cdf

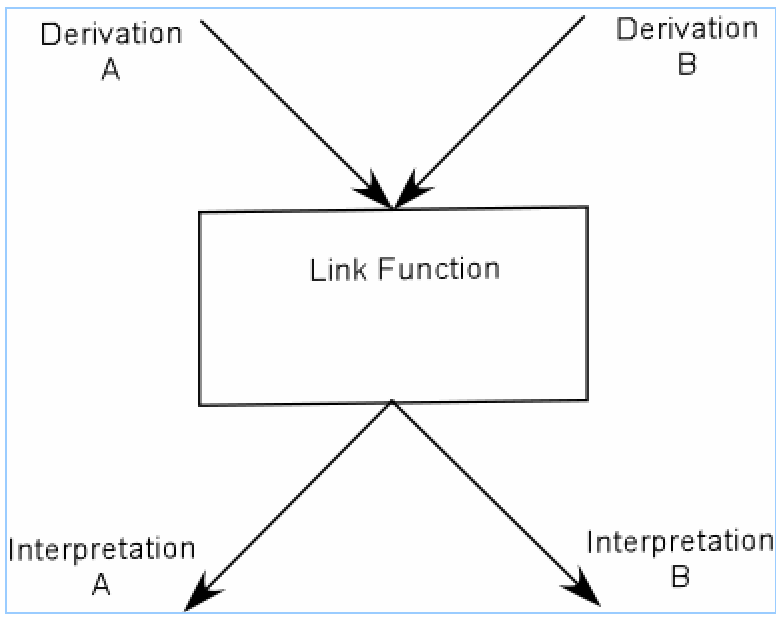
$$F_X(x) = \frac{1}{1 + e^{\mu(\eta - x)}} \tag{53}$$

where η is a location parameter and μ is a positive scale parameter.

For convenience we set $\eta = 0$ and $\mu = 1$ in binary and multinomial logit models. In this special case equation (53) again reduces to that of the logistic curve (15). The assumptions $\eta = 0$ and $\mu = 1$ will be discussed in Section 4.4.

The congruence between (53) and (49) suggests that the actual numeric calculation of the regression parameters is the same. The two only differ in the derivation of the formula and in its interpretation (Figure 16). As a practical consequence there are at least two distinct approaches to the interpretation of logistic regression results.

Figure 16 - Only Derivation and Interpretation Changes



Logit models in econometrics and by extension transportation engineering only use the *cdf* perspective for the interpretation of estimates.

The other important model in the category of cumulative distributions as probability transform is the probit model. We replace the π notation with $\pi(x)$ to avoid confusion with the mathematical constant π.

Let s be the variable of integration in the definition of the Gaussian Distribution with mean μ and standard deviation σ, and let Φ indicates the cumulative probability function of the standard normal $N(0,1)$ distribution..

$$\pi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{s-\mu}{\sigma}\right)^2} ds \tag{54}$$

$$\pi(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \tag{55}$$

The model

$$\boxed{\Phi^{-1}[\pi(x)] = \alpha + \beta x} \tag{56}$$

is the probit model (Agresti, 1990, p. 103). This leads to the definition of probit:

$$\boxed{\text{probit}[\pi(x)] \equiv \Phi^{-1}[\pi(x)]} \tag{57}$$

The standard normal *cdf* takes z-scores as input and outputs cumulative probabilities. Its inverse function reverses the process. The probit transform therefore maps probabilities to standardized z-scores. This leads to a natural interpretation of probit models.

We know that a cumulative probability of 50% is zero standard deviations away from the mean. With this information we can calculate the mean of the true population normal distribution which models the data in our probit model, and which, in other words, is our regression curve, in terms of the true regression parameters α and β.

$$\text{probit}(0.5) = \alpha + \beta x = 0 \tag{58}$$

$$\beta x = -\alpha \tag{59}$$

$$x = -\frac{\alpha}{\beta} \tag{60}$$

In the statistical literature the distribution of the random component in the probit model is not understood to be normal but binomial. The inverse of the normal *cdf* is just a probability transform or a link function. In the same way, the random component of the logit model is seen to be distributed binomial and not logistic. The inverse of the logistic *cdf*

$$g(\pi) = \log\left(\frac{\pi}{1-\pi}\right) \tag{61}$$

again is just the link function. Alan Agresti gives a good summary of the probit model:

> "The probit model is a GLM with binomial random component and probit link." …
>
> "The probit transform maps $\pi(x)$ so that the regression curve for $\pi(x)$ (or $1 - \pi(x)$ when $\beta < 0$) has the *appearance* of a the normal *cdf* with mean $\mu = -\frac{\alpha}{\beta}$ and standard deviation $\sigma = \frac{1}{|\beta|}$."
>
> (Agresti, 1996, p. 79, emphasis added).

Whether a curve has the *appearance* of a normal *cdf* with mean $\mu = -\frac{\alpha}{\beta}$ and standard deviation $\sigma = \frac{1}{|\beta|}$, as Agresti articulates it, or whether a curve _is_ a normal *cdf* with mean $\mu = -\frac{\alpha}{\beta}$ and standard deviation $\sigma = \frac{1}{|\beta|}$, the equation is the same.

The difference is mainly semantic.

4.3.4 Relevant Applications of the Insights Gained

Statistics approaches the problem depicted in Figure 12 on page 118 from the perspective of Generalized Linear Models (GLM). Generalized linear models provide a unifying framework for many commonly used statistical techniques. It gives us a "bird's eye perspective" of a wide variety of models. As an example, we don't only look at the final result of a derivation, e.g. the formula for the choice probability P_{in} but also at its inverse function. This much broader perspective greatly facilitates detecting errors, interpreting results, and solving problems.

1. Consider again the linear probability model shown in Figure 12 on page 118. Without much reflection it may seem like an ordinary least squares regression. Looking again we realize that the response variable (probability of success for each subject) is distributed Bernoulli(π). Its variance $\pi(1 - \pi)$ is not constant. It depends on x because π depends on x. Figure 15 vividly illustrated the nonconstant variance. This violates assumption 2 of the Gauss Markov Theorem (Goldberger, 1991, p. 163ff.). OLS estimators are not BLUE (Best Linear Unbiased Estimator). Maximum likelihood estimators may have a smaller standard error. The linear probability model is a GLM with binomial random component and identity link function (Agresti, 1996, p. 75). From the perspective of generalized linear models it is easy to see that the linear probability model is not an OLS regression model.

2. As each independent variable has its own coefficient estimate, so does it have its own odds ratio estimate. There is a direct relationship between

coefficient and odds ratio. Both describe the same fact, but from a different perspective. The terms odds and odds ratio can be used interchangeably.

The logarithm of the odds (the logit) has a simple linear relationship to the explanatory variable:

$$\log\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta x \tag{62}$$

The logit increases by β units for every 1-unit increase in x (Agresti, 1990, p. 108). Since it is difficult to think in terms of logarithms this interpretation is not very helpful.

An expression for the "odds of success" was first given in Equation (45).

$$\frac{\pi}{1-\pi} = e^{\alpha+\beta x} \tag{63}$$

Rearranging terms we obtain:

$$\frac{\pi}{1-\pi} = e^{\alpha}\left(e^{\beta}\right)^{x} \tag{64}$$

The exponential relationship implies that each unit increase in x has a multiplicative effect on the odds ratio (Agresti & Finlay, 1986, p. 485). For example, in an airport pair choice model let the odds ratio for airport access distance be 0.951, and let airport access distance be measured in km. We could say that each additional km of access distance reduces the odds of choosing that particular airport pair by 5%. In case of a 2 km increase we would see the odds reduced by about 9½% ($0.951^2 = 0.9044$).

To find the expression for the change in odds as x is increased by one unit, we set $\alpha = 0$ and $x = 1$ and obtain:

$$\frac{\pi}{1-\pi} = e^{\beta} \tag{65}$$

(65) suggests an easy calculation of the odds ratio estimate for each regressor using its coefficient estimate. All we need to do is exponentiate the coefficient estimate.

Odds ratios are so intuitive that they are frequently used by people with no formal training in probability.

3. Most derivations presented in this chapter are noticeably simpler than the equivalent derivations laid out in the majority of textbooks. That is because we consider all the different perspectives and then choose the most convenient starting point. This example analyzes the effect of reversing the definition of failure and success starting with the logit definition:

$$\boxed{\operatorname{logit}(\pi) \equiv \log\left(\frac{\pi}{1-\pi}\right)} \tag{66}$$

$$\operatorname{logit}(1-\pi) = \log\left(\frac{1-\pi}{\pi}\right) \tag{67}$$

$$\operatorname{logit}(1-\pi) = -\log\left(\frac{\pi}{1-\pi}\right) \tag{68}$$

$$\boxed{\operatorname{logit}(1-\pi) = -\operatorname{logit}(\pi) = -\alpha - \beta x} \tag{69}$$

The effect of switching the order of the response level is to reverse the signs of the parameter estimates.

Using as a starting point only the standard multinomial logit formulation of econometrics / transportation engineering, it is certainly possible to derive elasticities, obtain the formula for the calculation of odds ratios, or predict the effect of response level switching. But the route is cumbersome, more complicated than it needs to be, and, as a direct result, much less intuitive.

4. The fourth advantage of the statistical perspective may be the most important. For a particular problem at hand econometric software might not be available. In that case the only options would be to write specialized software for just that one application, not to estimate, or to use statistical software. However the latter is not understandable without knowledge of the concepts of generalized linear models. The research at hand provides a good illustration of this predicament.

The following subsection discusses the different classes of multinomial logit models specifically emphasizing the distinction between generalized and conditional logit models (please see Figure 17 on page 149 and Table 10 on page 150). Most economic and transportation related textbooks do not differentiate between the two. This is unfortunate because the routines written for and data transformations required by each class of model are substantially different in many software packages. The deeper understanding of multinomial logit models also helps with model results interpretation.

4.3.5 Multinomial Logit Models

The terms logistic regression and multiple logistic regression generally only refer to binary response models. For multicategory models with multinomial distribution of the response variable we need to distinguish between multinomial *response* models (Agresti, 1990, p. 306), which could use any link function, and multinomial *logit* models (Kuhfeld, 2003, p. 345) based on the logit link. Many different types of methods are available to fit categorical response models with a natural ordering (e.g.: good, better, best). Loglinear models which use the log link $g(\mu) = \log(\mu)$ are frequently used to solve such problems. Logit models only comprise part of the tool set available.

Logit models for nominal responses (Figure 17 on page 149) present a special challenge. While generalized logit models (precisely defined later in this subchapter), just like classical regression models can contain both discrete and continuous variables, they only allow the choice to be dependent on the characteristics of the chooser. In qualitative choice analysis, however, it is more critical to model choice as a function of the alternative's characteristics which vary both among alternatives and choosers. An important question to be answered would be: 'How does mode split change if travel time of alternative 1 is reduced?' We are not so much interested in how mode split changes as trip makers grow older or their income increases.

McFadden (1974) proposed the conditional logit model for explanatory variables which describe each alternative's attributes. Not only each alternative's characteristics but also the choice set can vary among individuals. For each pair of choices only the difference in choice attributes enters the model. Let π_{in} denote the probability of individual n choosing alternative i. For subject n and choice i, let $\mathbf{X}_{in}$ represent the values of the K

explanatory variables. The logit form of the conditional logit model for each pair of alternatives i and j can then be expressed as (Agresti, 1990, p. 316)

$$\log\left(\frac{\pi_{in}}{\pi_{jn}}\right) = \boldsymbol{\beta}'\left(\mathbf{X}_{in} - \mathbf{X}_{jn}\right) \tag{70}$$

where $\boldsymbol{\beta}$ is a single vector of coefficients.

In order to better understand McFadden's conditional logit model we need to distinguish it from the generalized logit model. We first partition $\mathbf{X}_{in}$ into $\mathbf{S}_n$ and $\mathbf{Z}_{in}$.

Let $\mathbf{S}_n$ represent the characteristics of individual n, and let $\mathbf{Z}_{in}$ symbolize the attributes of alternative i for individual n. Also, let C define a set that includes all potential choices for a given population and J indicate the total number of elements in this master choice set. We let $C_n \subseteq C$ designate each individual's feasible choice set and $J_n \leq J$ be his or her number of feasible choices.

For example, for each of the J_n travel modes available to individual n, $\mathbf{S}_n$ may be (1, 47)' where 1 is the dummy variable for the intercept (alternative specific constant) and 47 the trip maker's annual income in thousands of dollars. $\mathbf{Z}_{in}$ could contain travel time and travel cost for each individual and every alternative available to him or her.

The generalized logit model only uses individual characteristics as explanatory variables (Kuhfeld, 2003, p. 347).

$$\pi_{in} = \frac{e^{\boldsymbol{\gamma}_i'\mathbf{S}_n}}{\sum_{j\in C_n} e^{\boldsymbol{\gamma}_j'\mathbf{S}_n}} = \frac{1}{\sum_{j\in C_n} e^{(\boldsymbol{\gamma}_j-\boldsymbol{\gamma}_i)'\mathbf{S}_n}} \tag{71}$$

$\boldsymbol{\gamma}_j$ are vectors of unknown regression parameters corresponding to each of the available choices. Since $\sum_{j=1}^{J_n} \pi_{jn} = 1$ the parameter values are not unique unless we set one of the vectors to null. By fixing $\boldsymbol{\gamma}_{J_n} = \mathbf{0}$ the other vectors represent the effect of $\mathbf{S}_n$ on the probability of choosing the j^{th} alternative relative to the last one.

By contrast, in McFadden's conditional logit model the impact of a unit of **Z** is assumed to be constant across alternatives. As a result, we only obtain one single vector $\hat{\boldsymbol{\theta}}$ with parameter estimates for, continuing with our previous example, travel time and travel cost and no alternative specific constants (Kuhfeld, 2003, p. 347).

$$\pi_{in} = \frac{e^{\boldsymbol{\theta}'\mathbf{Z}_{in}}}{\sum_{j \in C_n} e^{\boldsymbol{\theta}'\mathbf{Z}_{jn}}} = \frac{1}{\sum_{j \in C_n} e^{\boldsymbol{\theta}'\left(\mathbf{Z}_{jn} - \mathbf{Z}_{in}\right)}} \tag{72}$$

Section 3.2.3 stated that we will compose a fairly sophisticated, stratified representation of the accessibility of each linehaul terminal from a particular zone. Socioeconomic variables are clearly needed for this task because of their power to explain differences in choices among travelers with identical trip characteristics. Just as we were not able to model probability directly with a linear combination of explanatory variables and had to model a logit transformation of probabilities instead, we cannot enter socioeconomic variables directly into a conditional logit model. We need to first transform individual characteristics into alternative specific variables. This is easily done by creating interaction terms like "egress distance × young age" where young age would be an indicator variable denoting membership in a particular group of travelers age 29 or younger. The parameter estimate will increase or decrease the slope of egress distance for that group.

As previously mentioned Equation (70) expresses the conditional logit model in *logit form*. Equation (72) states the same model first in *standard form* and then in *difference form*. It is hard to make a distinction between the generalized and the conditional logit model in standard form. Comparing the different forms of Equations (71) and (72), we see that parameter estimates for generalized logit models are the result of a subtraction. They correspond, for example, to the effect of income on the probability of choosing auto versus transit. On the other hand, parameter estimates in a conditional logit model are global and indicate the incremental effect in terms of a unit of that particular attribute (e.g. minute, dollar, mile). Once a socioeconomic variable has been transformed to be alternative specific (egress distance × young age) its coefficient changes the slope of the interaction variable with respect to utility and therefore takes on its inverse unit (in this case km^{-1}).

Reversing the partition of $\mathbf{X}_{in}$ into $\mathbf{S}_n$ and $\mathbf{Z}_{in}$ and understanding that

$$\mathbf{X}_{in} = \mathbf{Z}_{in} + h(\mathbf{Z}_{in,} \mathbf{S}_n) \tag{73}$$

where $h(\cdot)$ is some function transforming socioeconomic to alternative specific variables, we will now state the terminal choice model in difference form because of its more instructive arrangement of variables.

$$\pi_{in} = \frac{1}{\sum_{j \in C_n} e^{\boldsymbol{\beta}'(\mathbf{X}_{jn} - \mathbf{X}_{in})}} \tag{74}$$

To make the generalization from the binary logit model more apparent, we write (74) as:

$$\boxed{\pi_{in} = \frac{1}{1+\sum_{j\in C_n \text{ and } i\neq j} e^{\boldsymbol{\beta}'(\mathbf{X}_{jn}-\mathbf{X}_{in})}}} \tag{75}$$

Finally, we state (75) in a form more commonly used by economists and therefore transportation engineers and distinguish between airport pair choice model (76) and HSR station pair choice model (77):

$$P_{in} = \frac{1}{1+\sum_{j=1,i\neq j}^{82} e^{\boldsymbol{\beta}'(\mathbf{X}_{jn}-\mathbf{X}_{in})}} \tag{76}$$

$$P_{in} = \frac{1}{1+\sum_{j=1,i\neq j}^{1260} e^{\boldsymbol{\beta}'(\mathbf{X}_{jn}-\mathbf{X}_{in})}} \tag{77}$$

Notice that some indices have been suppressed to improve readability. The subscript n in the airport choice model (76) indexes all non-excluded air travelers in the survey. Their total number N is 30 126 (Table 19 on page 181). By the same token for the HSR choice model (77) $N = 18\,395$.

Figure 17 recaps the multinomial logit models discussed in this section and graphically illustrates their relationship to each other.

Table 10 summarizes and organizes the most important equations covered in Section 4.3.

For more detailed information on generalized and conditional logit models as well as other variable transformations the reader is referred to Kuhfeld (2003) pp. 345-353.

The building blocks for the design matrix $\boldsymbol{X}$, in other words the alternative specific and socioeconomic variables employed to model airport and station pair choice, were discussed in detail in the previous chapter. Table 8 on page 112 lists factors and covariates available for each of the two modes.

Figure 17 - Classes of Multinomial Logit Models

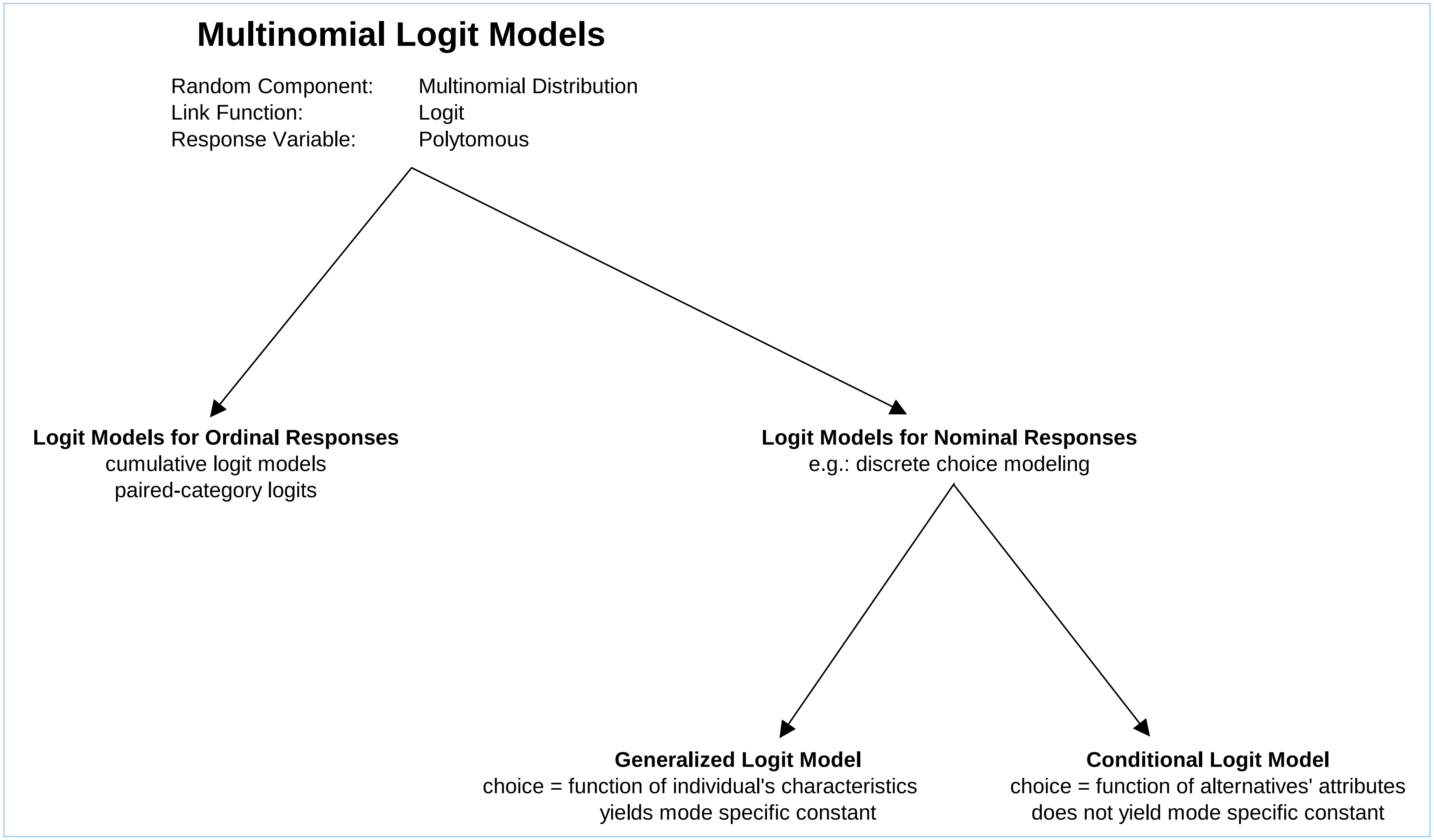


Table 10 - Forms of Logit Equations

Distribution of Random Component	Binomial	Multinomial	
Model Name	**Binary Logit**	**Generalized Logit**	**Conditional Logit**
Logit Form	$\log\left(\frac{\pi}{1-\pi}\right)=\alpha+\beta x$	$\log\left(\frac{\pi_{in}}{\pi_{jn}}\right)=\left(\boldsymbol{\gamma}_j-\boldsymbol{\gamma}_i\right)'\mathbf{S}_n$	$\log\left(\frac{\pi_{in}}{\pi_{jn}}\right)=\boldsymbol{\theta}'\left(\mathbf{Z}_{in}-\mathbf{Z}_{jn}\right)$
Standard Form	$\pi=\frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$	$\pi_{in}=\frac{e^{\boldsymbol{\gamma}_i'\mathbf{S}_n}}{\sum_{j\in C_n}e^{\boldsymbol{\gamma}_j'\mathbf{S}_n}}$	$\pi_{in}=\frac{e^{\boldsymbol{\theta}'\mathbf{Z}_{in}}}{\sum_{j\in C_n}e^{\boldsymbol{\theta}'\mathbf{Z}_{jn}}}$
Difference Form	$\pi=\frac{1}{1+e^{-(\alpha+\beta x)}}$	$\pi_{in}=\frac{1}{\sum_{j\in C_n}e^{\left(\boldsymbol{\gamma}_j-\boldsymbol{\gamma}_i\right)'\mathbf{S}_n}}$	$\pi_{in}=\frac{1}{\sum_{j\in C_n}e^{\boldsymbol{\theta}'\left(\mathbf{Z}_{jn}-\mathbf{Z}_{in}\right)}}$

4.4 Econometric Derivation from Utility Theory

4.4.1 A Behavioral Approach

Economists use similar models as statisticians but with a different emphasis. All models need a human behavioral component and need to be unified with and derived from utility theory. Econometricians and by extension transportation engineers draw only on a small subset of the models employed in statistics, but are very careful about their behavioral justification and in this context add many model variants.

We know from Equation (31) on page 126 that decision maker *n* will prefer air over rail if

$$V_{An} + \varepsilon_{An} > V_{Rn} + \varepsilon_{Rn} \tag{78}$$

Generalizing this to the multinomial case and solving for P_{in}, the probability of individual *n* choosing alternative *i*, we obtain

$$P_{in} = P\left(\boldsymbol{\beta}_i \mathbf{X}_{in} + \varepsilon_{in} \geq \max_{\forall j \neq i}\left(\boldsymbol{\beta}_j \mathbf{X}_{jn} + \varepsilon_{jn}\right)\right) \tag{79}$$

Note that the parameter vector $\boldsymbol{\beta}$ has been redefined to be a column vector, allowing us to dispense with the prime notation to indicate a transpose. This change of notation is effective only for this and the following subsection.

We need to assume a distribution for the random components ε. Ideally this distribution would have the desirable property that the maximum of a set of randomly drawn values from this distribution has the same distribution as the values from which the set was drawn. The extreme value type I distribution, also referred to as Gumbel distribution, has these properties.

The probability density function (pdf) $f(\varepsilon)$ and the cdf $F(\varepsilon)$ of the Gumbel distribution have the form

$$f(\varepsilon) = \mu e^{-\mu(\varepsilon-\eta)} e^{-e^{-\mu(\varepsilon-\eta)}} \tag{80}$$

$$F(\varepsilon) = e^{-e^{-\mu(\varepsilon-\eta)}} \qquad \mu > 0 \tag{81}$$

Its variance is

$$\sigma^2 = \frac{\pi^2}{6\mu^2} \tag{82}$$

Assuming that all ε are independently and identically distributed (IID) Gumbel variates with location parameter $\eta = 0$ and common scale parameter μ it can be shown that the systematic component of the maximum in Equation (79) can be expressed as:

$$\boldsymbol{\beta}^* \mathbf{X}_n^* = \frac{1}{\mu} \ln \sum_{\forall j \neq i} e^{\mu \boldsymbol{\beta}_j \mathbf{X}_{jn}} \tag{83}$$

That means

$$P_{in} = P\left(\boldsymbol{\beta}_i \mathbf{X}_{in} + \varepsilon_{in} \geq \boldsymbol{\beta}^* \mathbf{X}_n^* + \varepsilon_n^*\right) \tag{84}$$

$$P_{in} = P\left(\boldsymbol{\beta}^* \mathbf{X}_n^* + \varepsilon_n^* - \boldsymbol{\beta}_i \mathbf{X}_{in} + \varepsilon_{in} \leq 0\right) \tag{85}$$

Since the difference between 2 IID Gumbel variates with common scale parameter μ is logistically distributed, we may write:

$$P_{in} = \frac{1}{1 + e^{\mu\left(\boldsymbol{\beta}^* \mathbf{X}_n^* - \boldsymbol{\beta}_i \mathbf{X}_{in}\right)}} \tag{86}$$

Rearranging terms, substituting (83), and setting μ = 1 we obtain the standard form of the multinomial logit (MNL) model:

$$P_{in} = \frac{e^{\boldsymbol{\beta}_i \mathbf{X}_{in}}}{\sum_{\forall j} e^{\boldsymbol{\beta}_j \mathbf{X}_{jn}}} \tag{87}$$

(Washington, Karlaftis & Mannering, 2003, pp. 260-263)

For an alternative derivation of the multinomial logit model from the Gumbel distribution please see Ruud (2000) pp. 780-781.

The assumption $\eta = 0$ is not very restrictive as long as each systematic utility has a constant term. However the postulation $\mu = 1$ requires careful consideration. (82) implies that μ plays the same role as variance in normal distributions, but it is inversely proportional to standard deviation. In binary and multinomial logit models we just set μ = 1 since the scale of the individual utilities is meaningless, only the order of the differences matter. μ just multiplies the difference of the utilities (86). Yet the scale of the binary and multinomial logit models manifests itself through the magnitude of the coefficients. If μ is small it means the coefficients are small. We find a "shallow" model with coefficients barely larger than their standard errors. Due to the inverse relationship with variance the errors are big in relation to the coefficients. Because of different scales we cannot compare coefficient estimates of different models directly, only ratios of coefficients.

The disturbances being distributed IID represents an important restriction.

- It forces the scale parameter μ of all disturbances to be equal, meaning all disturbances have the same variance.
- The error terms have to be independent which may be difficult to defend.

In summary, through the postulation $\mu = 1$ "we reflect the assumption of homoscedastic disturbances; if this assumption is inappropriate for the population in question, it is necessary to take suitable measures …" (M. E. Ben-Akiva & Lerman, 1985, pp. 104 and 107).

Those suitable measures will be discussed in Section 5.5 below starting on page 189.

4.4.2 Independence of Irrelevant and Relevant Alternatives

An important feature of MNL models results from the assumption of IID error terms. It is euphemistically called Independence of Irrelevant Alternatives (IIA). This property should really be called Independence of Irrelevant and Relevant Alternatives (IIRA). As we can easily see from

$$\frac{P_{in}}{P_{kn}} = \frac{\dfrac{e^{\boldsymbol{\beta}_i \mathbf{X}_{in}}}{\sum_{\forall j} e^{\boldsymbol{\beta}_j \mathbf{X}_{jn}}}}{\dfrac{e^{\boldsymbol{\beta}_k \mathbf{X}_{kn}}}{\sum_{\forall j} e^{\boldsymbol{\beta}_j \mathbf{X}_{jn}}}} = \frac{e^{\boldsymbol{\beta}_i \mathbf{X}_{in}}}{e^{\boldsymbol{\beta}_k \mathbf{X}_{kn}}} \tag{88}$$

the ratio of probabilities only depends on i and k. It is independent from all other alternatives in choice set C_n. This is both a good and a bad property of MNL.

It is good when the other alternatives are truly "irrelevant," because it dramatically simplifies the estimation. Because we believe IIA is appropriate, only two of possible six intercity modes were considered in this dissertation. The focus is on the competition between air and high speed rail. Conventional rail, auto, intercity bus, and ferry are not included in the model.

We were also able to take advantage of the IIA feature by randomizing choice sets (Section 5.6 starting on page 199). However, that was only possible after first correcting for heteroskedasticity in the error term. The composition of a fairly sophisticated, stratified representation of accessibility heavily relies on access distance to the terminal (Subsection 3.2.1). Since access distance is calculated between the terminal and the administrative center of the particular zone, and different zones vary greatly in size, error terms are not identically distributed. As previously mentioned, suitable measures will be discussed in Section 5.5 below.

IIA does not hold when the other alternatives are "relevant." Subway and commuter trains are similar enough that they will share unobserved attributes which are common to all rail transit modes. That means the assumption of independent error terms cannot be maintained. The classic green bus/blue bus example illustrates an extreme violation because the error terms are exactly the same (Train, 1986, pp. 18-21).

Terminal pair choice models for air and rail cannot be combined for exactly the same reason. HSR station pairs on one hand and airport pairs on the other would each have similarities in their respective error terms, violating the assumption of independently distributed Gumbel variates if we combine them. Nested logit models overcome this limitation and are discussed in the next subsection.

4.4.3 Nested Models

Nested models allow us to separate all choices with common unobserved attributes into different nests, calculate a composite utility for each nest and estimate the coefficients of these composite utilities in the upper nest. Informally, the scale parameter μ, which could be ignored in MNL models, is explicitly estimated in nested logit models.

The composite utility is known as the inclusive value or the logsum. It is the log of the sum of the estimated utilities for all of the choices available within a nest to decision maker *n*. So the inclusive value for air, e.g., is summed over all 82 airport pair choices. Let I_{An} and I_{Rn} denote the inclusive value of traveler *n* for air and rail respectively, then

$$I_{An} = \log \sum_{j=1}^{82} e^{\hat{\boldsymbol{\beta}}' \mathbf{x}_{jn}} \tag{89}$$

$$I_{Rn} = \log \sum_{j=1}^{1260} e^{\hat{\boldsymbol{\beta}}' \mathbf{x}_{jn}} \tag{90}$$

The "inclusive value" (McFadden, 1978b) represents the expected received utility from a nest. It measures the expected value of the highest utility among all the alternatives in the nest for each choice maker. It describes "the utility of the best alternative in a subset of choices as a summary of the 'value' of that subset to an individual." (M. E. Ben-Akiva & Lerman, 1985, p. 282)

That the inclusive value is a generalization of the systematic utility becomes more intuitive when we consider the special situation of a decision maker only having one alternative, say *LF*, available to him or her within a particular nest *F*. In this case:

$$I_F = \log e^{V_{LF}} = V_{LF} \tag{91}$$

In contrast to the airport pair (76) and HSR station pair (77) choice models on page 147, subscript n in equations (89) and (90) indexes both air and HSR travelers. However air travelers for whom the HSR alternative is not available (e.g. those traveling solely within the island of Kyushu) are excluded. It would not be possible to calculate an inclusive value of rail for these trip makers. The same holds for HSR trips without an air alternative like Tokyo – Nagoya. For equations (89) and (90) N = 38 255 (Table 19 on page 181).

The inclusive values make up the design matrix $\boldsymbol{X}$ whose coefficient vector $\boldsymbol{\gamma}$ we estimate in the upper level choice model.

$$V_{An} = \gamma_A I_{An} \tag{92}$$

$$V_{Rn} = \alpha + \gamma_R I_{Rn} \tag{93}$$

$$P_{Rn} = \frac{1}{1 + e^{-(V_{Rn} - V_{An})}} \tag{94}$$

It is instructive to distinguish between three different cases (Table 11).

Table 11 - Types of Nested Logit Models

Inclusive Value Coefficient	Interpretation	Distribution of Error Term
$\gamma_A = \gamma_R = 1$	The Nested Logit collapses to the MNL model.	IID Gumble
$\gamma_A = \gamma_R$	The error terms in both nests have the same variance. They are identically distributed, but not independent. That is why Nested Logit is used.	ID Gumble
$\gamma_A \neq \gamma_R$	The error terms are not identically distributed. They have different variances.	Gumble

It should be understood that *within* a lower level nest, error terms have to be always IID Gumble to meet the condition for MNL to be used.

The airport pair choice models are substantially different from the HSR station pair models estimated for this dissertation. We can therefore expect $\gamma_A \neq \gamma_R$. Table 8 on page 112 shows that fare levels and frequency are available to estimate airport pair, but not HSR station pair choice. It is also important to remain open to the possibility that some variables may be significant for the terminal pair choice of one but not the other mode. Fixed effects are also different as long as they are not based on geographic criteria (Tokyo metropolitan area), but are terminal specific (Tokyo Haneda Airport, Tokyo Central Station). These distinctions help terminal pair choice models to be more mode specific and therefore more precise.

What may have helped model precision the most was the ability to use different data subsets for lower and upper level models. 18 395 records are available for HSR only models, of which only half describe trips where air was a viable alternative (Table 19 on page 181). All 18 395 observations were made use of to estimate the terminal pair choice model. That means all observations contributed to the coefficients employed in the construction of the inclusive value of rail for each traveler.

4.5 Model Summary and List of Assumptions

Table 12 is laid out as a nested logit model diagram like Figure 10 on page 90. It should be read from the bottom up.

Table 12 - Model Summary

All Travelers in Markets where Air and HSR Compete

$$P_{Rn} = \frac{1}{1+e^{-(V_{Rn}-V_{An})}} \qquad (94)$$

$$V_{An} = \gamma_A I_{An} \qquad (92)$$

$$V_{Rn} = \alpha + \gamma_R I_{Rn} \qquad (93)$$

All Travelers in Markets where Air and HSR Compete

$$I_{An} = \log \sum_{j=1}^{82} e^{\hat{\boldsymbol{\beta}}'\mathbf{X}_{jn}} \qquad (89)$$

All Travelers in Markets where Air and HSR Compete

$$I_{Rn} = \log \sum_{j=1}^{1260} e^{\hat{\boldsymbol{\beta}}'\mathbf{X}_{jn}} \qquad (90)$$

All Air Travelers

$$P_{in} = \frac{1}{1+\sum_{j=1,i\neq j}^{82} e^{\boldsymbol{\beta}'(\mathbf{X}_{jn}-\mathbf{X}_{in})}} \qquad (76)$$

All HSR Travelers

$$P_{in} = \frac{1}{1+\sum_{j=1,i\neq j}^{1260} e^{\boldsymbol{\beta}'(\mathbf{X}_{jn}-\mathbf{X}_{in})}} \qquad (77)$$

The major assumptions for the conditional logit models (76) and (77) used in this dissertation are:

1. The error terms are distributed Gumble.
2. Its location parameter $\eta = 0$.
3. Its scale parameter $\mu = 1$.
4. The impact of a unit of $\boldsymbol{X}$ is constant across alternatives.
5. All choices are subject to the Axioms of Rational Choice (Nicholson, 1992, Chapter 3).
6. For the study of the competition between air and HSR between Tokyo and Fukuoka in Japan the other intercity modes are irrelevant, that means IIA holds.
7. For the HSR station choice models estimated in this research the IIA property reflects reality. It is therefore appropriate to estimate model parameters on a subset of alternatives for each sampled traveler.

For the nested logit model (94) only assumption 3. is relaxed.

Other chapters in this dissertation are related to the list of assumptions. Specifically, section 3.2.3 starting on page 79 discusses the boundaries imposed on the construction of the accessibility index by data availability.

5 Data Sources and Sampling Issues

5.1 The 1995 Japanese Intercity Travel Survey

5.1.1 Level of Detail

The 1995 Japanese Intercity Travel Survey is the second survey of its kind conducted under the auspices of the National Land Agency. The first survey was conducted in 1990, and the third one was carried out in 2000. It includes data on all intercity modes used in Japan, air, rail, bus, automobile, and ferry (Table 13). Each observation corresponds to an actual intercity trip taken in the Fall of 1995. For an approximate English translation of the survey instrument used for rail passengers please see Appendix 1 on page 280.

As already noted in Subchapter 3.2.1 Definitions, each observation in our database describes one trip with one origin and one destination for a single traveler. That means in this dissertation the terms observation, trip, traveler, trip maker, and passenger can be used interchangeably.

Table 13 - The 1995 Japanese Intercity Travel Survey

Mode	Number of Observations
Air	161,596
Rail	74,740
Bus	23,104
Car	365,707
Ferry	9,060
Total	**634,207**

The numbering system for the 47 prefectures in Japan starts in the North with 01 for the island of Hokkaido (Sapporo is its major city) and ends in the South with the number 47 given to Okinawa (not shown on Map 1, page 64). This analysis is restricted to intercity air and high speed rail travel within 24 out of those 47 prefectures. Table 14 lists these 24 prefectures and Map 2 shows their location relative to the rest of the country. The prefectures on Japan's main island of Honshu were selected because of the research focus on the Tokaido and Sanyo corridors. The four prefectures on the island of Shikoku (numbered 36 – 39) were excluded since access to this island is limited to a few bridges and ferry lines.

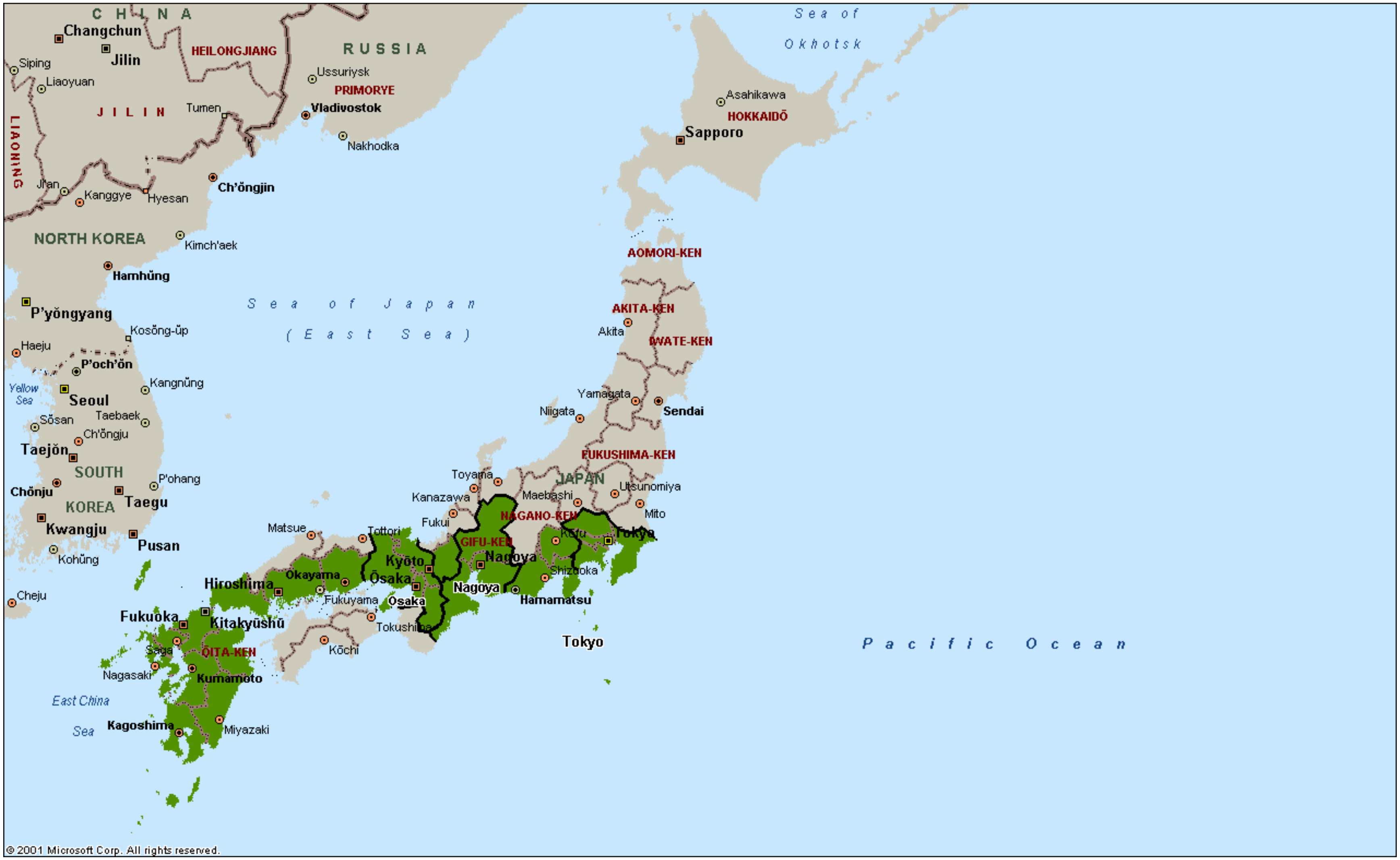
Prefectures numbered 40 to 46 are located on the island of Kyushu. Map 1 on page 64 illustrates that the only Sanyo Shinkansen high speed rail terminal on this island is in Fukuoka. As mentioned earlier, a portion of the Kyushu Shinkansen is in operation today, however even at the point of this writing it is not connected yet to the Sanyo Shinkansen in Fukuoka. In 1995, Kyushu represented a special case inasmuch as there are many more commercial airports on this island than Shinkansen stations. On the main island of Honshu exactly the opposite is true. This allows for interesting comparisons regarding access and egress issues.

"Intercity Travel" was defined as travel crossing prefecture boundaries. The three metropolitan areas of Tokyo, Nagoya, and Osaka were treated as one prefecture each for the purpose of this survey (Table 14). Prefecture boundaries are depicted as dotted lines on Map 2, however the metropolitan areas of Tokyo, Nagoya, and Osaka are demarcated by bold lines. The great advantage of this survey is its detailed description of each trip including up to two intermediate airports and up to four intermediate rail stations. The trip purpose and trip length (one day or overnight) is also included for most records, as is information on the type of access and egress mode. Please see Appendix 2 on page 283 for an informal English version of the database record layout which specifies all the information that is available from this survey.

Table 14 - The 24 out of 47 Prefectures Used in this Analysis

Prefecture			
Number	**Name**	**Metro Area**	**Island**
11	Saitama-ken	Tokyo	Honshu
12	Chiba-ken	Tokyo	Honshu
13	**Tokyo-to**	Tokyo	Honshu
14	Kanagawa-ken	Tokyo	Honshu
19	Yamanashi-ken		Honshu
21	Gifu-ken	Nagoya	Honshu
22	Shizuoka-ken		Honshu
23	Aichi-ken	Nagoya	Honshu
24	Mie-ken	Nagoya	Honshu
25	Shiga-ken		Honshu
26	Kyoto-fu	Osaka	Honshu
27	**Osaka-fu**	Osaka	Honshu
28	Hyogo-ken	Osaka	Honshu
29	Nara-ken	Osaka	Honshu
33	**Okayama-ken**		Honshu
34	**Hiroshima-ken**		Honshu
35	Yamaguchi-ken		Honshu
40	**Fukuoka-ken**		Kyushu
41	Saga-ken		Kyushu
42	**Nagasaki-ken**		Kyushu
43	Kumamoto-ken		Kyushu
44	Oita-ken		Kyushu
45	Miyazaki-ken		Kyushu
46	**Kagoshima-ken**		Kyushu

Trip origins and destinations are coded down to 5 digits, the first 2 digits being the prefecture, the first 3 digits denoting either a large city, like Tokyo, or a large area within a prefecture. There are 207 3-digit zones in Japan. The 5 digits then go down to the level of a ward within Tokyo (Table 15). The data is therefore perfectly suited to study access and egress issues. Even a cursory inspection of the survey instrument presented on page 280 will reveal the intense focus on the chronologically first and last parts of the trip, to and from the first and last railway station or airport.



Map 2 – The 24 out of 47 Prefectures Used in this Analysis

Table 15 - Political Subdivisions in Japan

	2-digit	3-digit	5-digit
Name	prefecture	"207-code"	"5-digit code"
Number of Zones	47	207	3230
Example	13= Tokyo prefecture	131= City of Tokyo	13101-13123= 23 wards of Tokyo

As discussed in Section 3.3.3 Classes of Service starting on page 89, this dataset lacks information on which train category was chosen or which fare was paid. Since we neither have travel times, we cannot distinguish between the semi-fast Hikari and super fast, premium service Nozomi. The dataset does not even differentiate between conventional and high speed rail. We can only surmise that the traveler chose high speed rail by his or her choice of rail stations (the use of conventional rail between two high speed rail stations is extremely unlikely). This means we cannot model the choice of train category, but we can model the station pair choice. Similarly, we modeled airport pair choice for the air mode. With terminal pair choices for each mode modeled in the lower nests we could model mode choice in the upper nest of a nested logit model.

The socio-economic variables in the dataset are sex, age, trip purpose, and trip length (same day or overnight). Other important socioeconomic data had to be inferred from aggregate zonal data. Very fortunately, the survey data does include the residential 5-digit zone code for each traveler, in addition to the 5-digit trip origin and destination zones. This is very unusual for an intercity travel survey. The residential zone code made it possible to link each trip record to specific aggregate zonal socio-economic data.

It also made it possible to clearly distinguish between access and egress as defined earlier.

Each record also contains daily and yearly extension coefficients. Origin and Destination (O&D) trip matrices for a typical Fall day in 1995 and also for the whole year of 1995 can be reconstructed by multiplying each observation by the respective expansion coefficient. Daily extension coefficients contributed to the weights used for model estimation, which is explained in more detail in Section 5.4 below.

Table 16 shows the number of observations available for our 24 prefecture analysis area (unweighted column). The daily extension coefficient gives us the actual number of trips made on a typical Fall day in 1995 that each record represents. By summing up these coefficients by mode we are able to estimate mode shares for intercity trips within the 24 prefectures.

Table 16 - Mode Shares for Intercity Travel within the 24 Prefectures

	Observations		Mode Share
	unweighted	weighted	
Air	57,859	77,339	5%
High Speed Rail	30,765	362,987	24%
Conventional Rail	6,500	65,898	4%
Automobile	123,073	980,324	64%
Bus	7,715	33,776	2%
Ferry	1,986	12,366	1%
Total		1,532,690	100%

5.1.2 Travel Modes Not Considered

Auto

The private automobile is the dominant mode of intercity travel even in Japan and even within the 24 densely populated prefectures chosen for this analysis (Table 16). Unfortunately the majority of the observations only contain information on the toll stations where the car entered and exited the tollway. Without residential zone code and access and egress zone codes not much information can be deduced from them. For this reason, automobile travel is not included in our analysis.

Intercity Bus and Conventional Rail

Observations describing trips where the bus was used as the primary intercity mode are fairly complete, but made reference to hundreds of intercity bus terminals. It was possible to obtain latitude and longitude data for the 18 airports (Table 17) and 36 Shinkansen stations (33 shown in Table 4 on page 63), but it was not feasible to do so for the hundreds of intercity bus terminals, especially since their market share was only 2% (Table 16).

The same problem prevented the inclusion of travel data for the narrow gauge (1 067 mm) conventional railways. These records were excluded because they did not include at least two Shinkansen stations, implying they were made exclusively on the narrow gauge (1 067 mm) network. People still used the conventional train system on routes parallel to the Shinkansen lines in 1995, because (1) it served more stations, (2) it was noticeably less expensive, and (3) it offered overnight and sleeper services in important markets like Tokyo – Osaka. Since these services have been consistently unprofitable they were gradually discontinued until they served almost exclusively as feeders to high speed lines. Even in 1995 7/8th of all intercity trips taken by rail included segments on the Shinkansen line (Table 16).

Ferries

Since the main focus of this dissertation is to study the effect of accessibility to High Speed Rail stations, and ferries do not compete directly with either High Speed Rail or air, the survey data on ferry trips was not considered for further study.

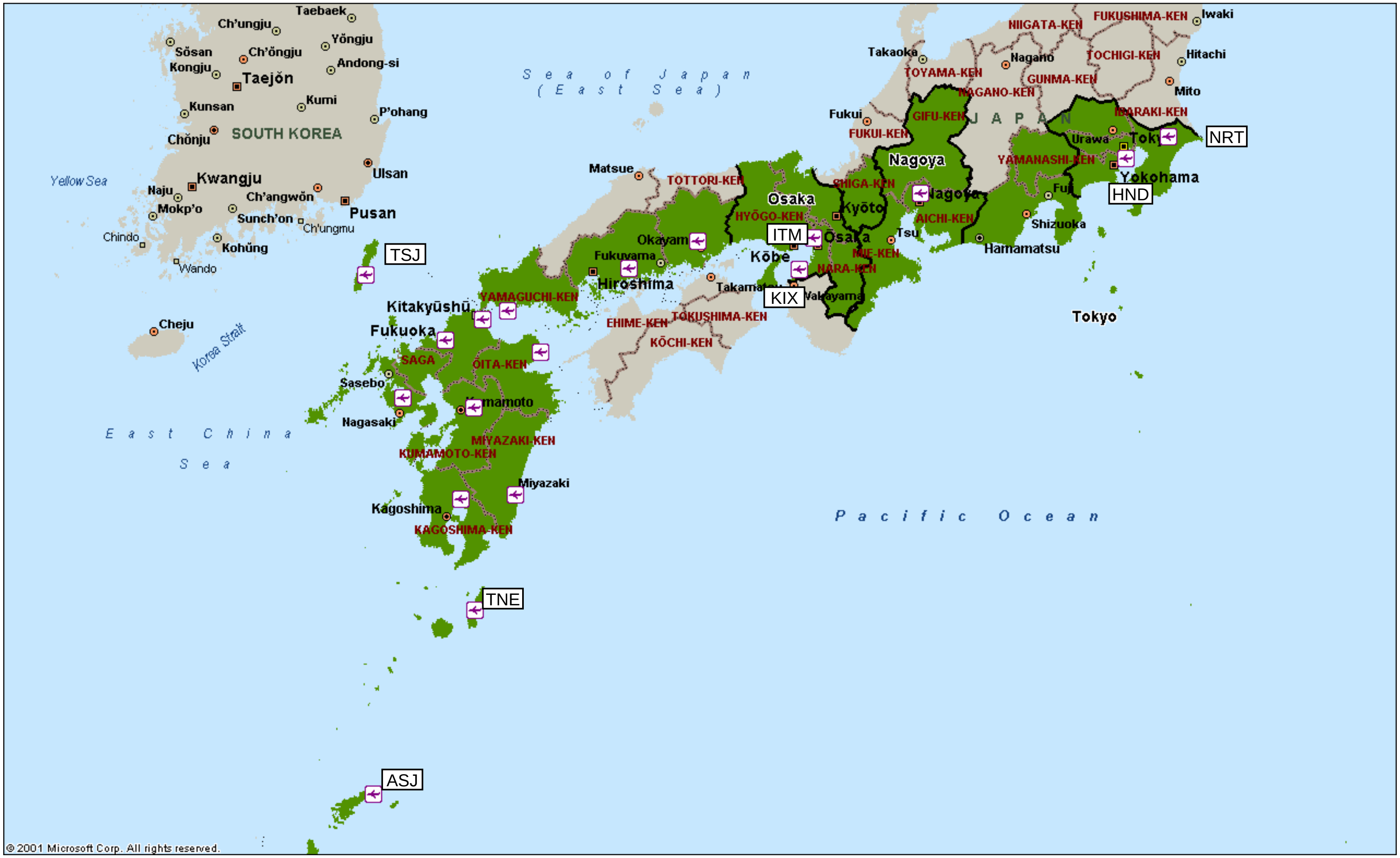
5.2 *Choice Set Generation and Excluded Observations*

5.2.1 Airport Pairs

The 1995 Japanese Intercity Travel Survey references 89 airports. Only 18 of them occur relatively frequently in the trip records for travel within the analysis area (Table 17 and Map 3). Tanegashima, Amami O Shima, and Tsushima are island airports (shima = island). They were therefore not suitable to be included in an airport choice model, which leaves us with the 15 most important airports. Notice that 7 of these 15 airports are on the island of Kyushu. That is particularly noteworthy since Kyushu is only about the size of the Netherlands or the combined size of New Hampshire and Vermont. The list includes both major airports for each of the three metropolitan areas of Tokyo, Osaka, and Fukuoka.

Table 17 - The 18 Most Important Airports in the Analysis Area

Number of Occurrences in Data Subset	Airport		Island
	Name	**Code**	
22798	Tokyo-Haneda	HND	Honshu
9609	Fukuoka-Itazuke	FUK	Kyushu
6874	Osaka-Itami	ITM	Honshu
3248	Kagoshima	KOJ	Kyushu
2734	Miyazaki	KMI	Kyushu
2691	Nagoya-Komaki	NGO	Honshu
2492	Nagasaki-Omura	NGS	Kyushu
2279	Oita	OIT	Kyushu
2263	Kumamoto	KMJ	Kyushu
2238	Hiroshima Intl	HIJ	Honshu
2012	Osaka-Kansai Intl	KIX	Honshu
984	Yamaguchi-Ube	UBJ	Honshu
470	Okayama	OKJ	Honshu
190	Fukuoka Kita-Kyushu	KKJ	Kyushu
149	Tanegashima	TNE	-
142	Amami O Shima	ASJ	-
101	Tsushima	TSJ	-
100	Tokyo-Narita	NRT	Honshu



Map 3 – The 18 Most Important Airports in the Analysis Area

A proprietary Level of Service (LOS) database (discussed in Subsection 5.3.2 below), which contains very detailed information on several alternative routings for each airport pair, was used to determine the optimal routing for each airport pair to be used in the choice set. Routings that included one transfer for airport pairs where direct air service was not available were initially extracted from the consultant database as well. Parameter estimates for number of transfers turned out to be highly unstable and insignificant. With a sample size of 30 433 observations based on travel between the top 18 airports, trips requiring a transfer were only chosen 37 times. After excluding the island airports resulting in the final sample size of 30 126 observations for air travel, transfers were not chosen at all anymore, and airport pairs requiring a transfer were no longer considered as part of the choice set. The number of transfers as a mode choice determinant especially with regards to air transportation in Japan was discussed at the end of Subsection 3.3.13 on page 109.

The 15 major airports considered would theoretically result in 15 x 14 = 210 airport pairs. Using both the proprietary Level of Service (LOS) database and the Official Airline Guide as a source, we were only able to collect level of service data for 82 airport pairs. Part of the reason may be the fact that the International Edition of the Official Airline Guide does not list all domestic flights in Japan. Indeed, four survey records had to be excluded because the itinerary included an airport pair for which there was no air service data.

But the main reason is most likely the lack of short haul flights in corridors where air would have to compete directly with high speed rail. As mentioned earlier, there is no domestic air service between Tokyo and Nagoya, a distance of 366 km, comparable to

New York – Washington. Due to the excellent ground transportation infrastructure and severe airport capacity constraints there are also no international feeder flights like the turbo-prop flights between Burbank, Ontario, or Orange County to Los Angeles International Airport. In areas where the high speed rail system is still under construction, as on the island of Kyushu, short-haul flights are still plentiful. The conventional 1 067 mm railroad is, not even in short-haul markets, able to compete strongly enough to suppress all air service.

5.2.2 HSR Station Pairs

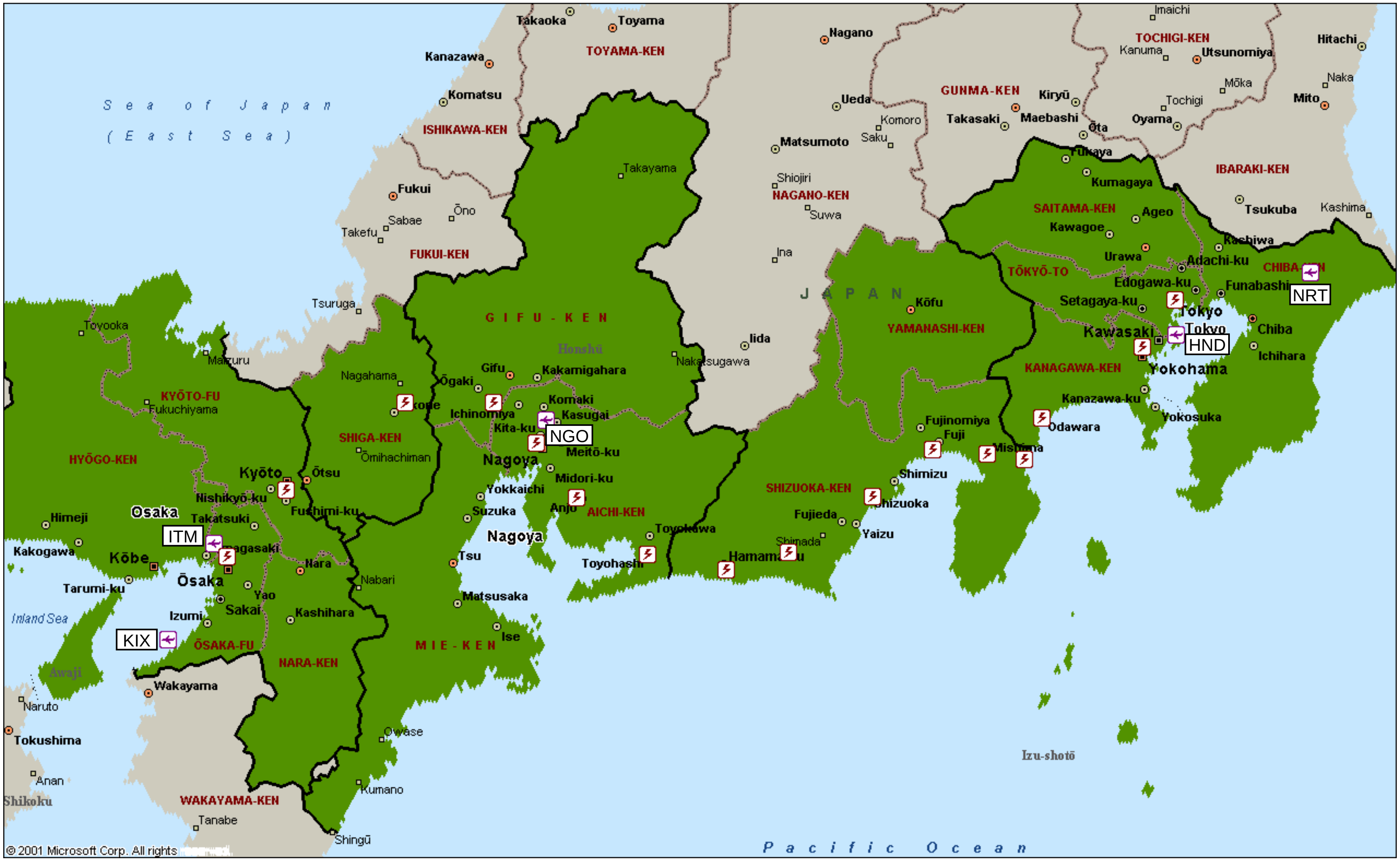
In addition to the 33 Shinkansen stations on the Tokaido (Map 5) and Sanyo lines (Map 6), not counting Asa and Shinagawa which were added after 1995, three Shinkansen stations in the Greater Tokyo Region on the lines to Niigata and Sendai were included in the analysis (Table 18 and Map 4). Without this addition, any traveler originating in or destined to the Greater Tokyo area would have had only three station choices: Tokyo Central, served by all high speed train categories, Yokohama, frequented by Hikari and Kodama, and Odawara, only reachable by Kodama (Map 5 on page 177 and Table 4 on page 63). There are no through-running trains from the East to the Tokaido line, but a transfer at Tokyo Central is fast and convenient. The insertion of these three stations ensures that all contiguous HSR lines and their respective stations in the analysis area of the 24 prefectures are included.

Table 18 - JR East Shinkansen Stations in Tokyo Metropolitan Area

HSR Station Name	Distance to Tokyo Central (km)
Kumagaya	65
Omiya	30
Tokyo-Ueno	4
Tokyo Central	**0**

After this supplement every high speed rail traveler in the analysis area theoretically has a choice of 36 x 35 = 1 260 station pairs. Why all of the station pairs were included in the choice set for every traveler is discussed in the following subsection.

Map 4 – Major Airports and JR East HSR Stations in Tokyo Metropolitan Area



Map 5 – Major Airports and JR Central HSR Stations in Tokaido Corridor

Map 6 – Major Airports and JR West HSR Stations in Sanyo Corridor

5.2.3 No Extreme Observations Discarded

The general principle for excluding observations is to be very careful that the excluding criteria are unrelated to the research question. The research question is roughly: How does the mode share between HSR and air change when we improve accessibility to the HSR station or to the airport? Excluding observations because of coding errors is not related to this question and will therefore not introduce a systematic bias. If we focus our research on only one corridor, then we can confidently only draw conclusions about that one corridor. Excluding travelers who do not live in that one corridor, because we do not have much information about them, will probably not skew the results either. On the other hand, any exclusions related to access and egress distance will most likely bias the parameter estimates for access and egress related variables, because it would systematically exclude travelers who made a particular station or airport pair choice in spite of a large access and/or egress distance.

And that is the reason for the exceptionally large choice set of 1 260 station pairs for rail. While most of these station pair choices may seem outrageously unrealistic for a given traveler, anything other than the complete choice set is not able to represent the data faithfully. Because of computational limitations described in more detail in the section on randomized choice sets, many other smaller choice sets were tried, but found lacking, and therefore discarded.

99% of all air travelers chose an access or egress distance below 139 or 134 km respectively. It would seem sensible to limit airport access and egress distance to 250 km, greatly reducing computer time required for model estimation. However, in one of our earlier estimations, the coefficient for access distance, e.g., changed from -0.052

to -0.060, making access distance appear considerably more important. We had systematically excluded travelers for whom the access distance was less important.

The decision was made not to discard outliers, which is easiest to defend on analytical grounds. From an empirical perspective it is probably the “safest” decision.

5.2.4 Other Exclusion Criteria

As we can see in Table 19, a large number of records had to be excluded due to insufficient 5-digit zone information. That means either origin or destination zone was only coded accurately down to a 3-digit city-level or to a 2-digit prefecture-level. In many instances the 5-digit zone in the 1995 Intercity Travel Survey could not be cross-referenced in the Japanese zone directory (Jinbunsha Editorial Staff, 2000). Table 19 also indicates that many observations lacked complete socio-economic information. Either sex, age, trip purpose, or trip length (same day or overnight) was missing.

Intra-Kyushu air travel was excluded from the mode choice model because the alternate mode (high speed rail) was not available. Intra-Kyushu air travel, however, is included in the 30 126 observations used to estimate the lower level air nest. The 92 HSR observations with both origin and destination zone on the island of Kyushu had to be discarded, because, as previously mentioned, there was no intra-Kyushu HSR service in 1995. These observations presumably involved trip chaining.

Table 19 - Excluded Observations

	Air	HSR	Total
Original dataset	161,596	74,740	236,336
Travel within 24 out of 47 prefectures	-103,732	-37,475	-141,207
Exclude train trips without High Speed Rail segment	0	-6,500	-6,500
Exclude records with missing daily extension coefficient	-5		-5
Number of Records Used to Calculate Mode Share	**57,859**	**30,765**	88,624
Exclude records with home zone outside of 24 prefecture area	-2,813	-627	-3,440
Exclude records with incomplete 5-digit zone information	-23,430	-8,485	-31,915
Delete records with incomplete socio-economic information		-3,088	-3,088
Exclude air records with origin / destination closest to same airport	-209		-209
Exclude HSR records with origin / destination closest to same Kodama station		-27	-27
Exclude records with traffic mode not equal to linehaul mode		-51	-51
Only keep air records where both chosen origin and destination airport is within top 15 airports	-1,277		-1,277
Exclude air records with no air service data for chosen airport pair	-4		-4
Number of Records Used for Air-only models	**30,126**		48,613
Exclude intra-Kyushu air and HSR travel	-1,448	-92	-1,540
Number of Records Used for HSR-only models		**18,395**	47,073
Exclude HSR records with origin / destination closest to the same top 15 airport		-817	-817
Delete HSR records with no air service between the two top 15 airports closest to origin and destination		-8,001	-8,001
Number of Records Used for Upper Nest of Nested Mode Choice Models	**28,678**	**9,577**	**38,255**
Exclude air records w/ no air service data for closest airport pair	-801		-801
Number of Records Used for Binary Mode Choice Models	**27,877**	**9,577**	**37,454**

5.3 *Additional Data Sources*

Distance to the nearest airport and Shinkansen high-speed railway station is measured as great circle distance ("as the crow flies") from the terminal to the administrative office of the 5-digit trip origin and destination zone. Latitude and longitude zone data to the second exact was taken from a Japanese directory (Jinbunsha Editorial Staff, 2000). JR East provided us with latitude and longitude data exact to the thousandth of a second for all Shinkansen stations. Airport data was obtained from a variety of sources, mainly (*Airport Characteristics Data Bank*, 1992) and

http://worldaerodata.com/countries/Japan.php

5.3.1 Socio-economic Information

Our 24 prefecture analysis area contains 1 574 5-digit zones, about half of the 3 320 in all of Japan. As already mentioned, latitude and longitude data (degrees, minutes, and seconds) for the administrative office of the 5-digit origin and destination zone was taken from a Japanese directory (Jinbunsha Editorial Staff, 2000). Population and zone size in square kilometers was also obtained from this directory.

For zone level socio-economic data a Japanese directory of regional economic statistics was consulted (*Chiiki Keizai Soran (Comprehensive Databank of Regional Economics - in Japanese)*, 2000). Zone size in square kilometers, number of households, income, vehicle ownership per 100 population, and zone type (urban or rural) was copied from this directory. The information in neither one of the two Japanese directories was available in electronic format. All the data for the 1 574 5-digit zones had to be entered manually. Since the second directory did not reference the 5-digit zone number, but only the zone name in the Japanese character set, the records from both directories had to be merged using zone sizes. That information was not completely consistent between the two directories. In many cases the differences between zone sizes was not very great and the merging of the records could be done programmatically. But the last 50 zones had to be merged manually by a Japanese speaker.

5.3.2 Level of Service Data Sources for Air

In order to have access to level of service (LOS) information, a special data set, prepared by a consulting company in Japan, has been made available to us on a confidential basis. This large 150 MB dataset contains very detailed information on several alternative routings for each airport pair. Air travel times, air fares, and air frequencies were taken from this data set. Air frequencies and especially air linehaul times were not consistent with the International Edition of the Official Airline Guide and had to be adjusted manually. The May 1994 issue of the international OAG (*OAG Desktop Flight Guide - Worldwide Edition*, 1994) was used because it was the closest available to the Fall of 1995.

5.3.3 Construction of Level of Service Data Matrix for HSR

Railway information in the proprietary LOS-database is limited to level of service data for an O & D matrix of the 207 3-digit zones, which is not precise enough for our purposes. It does not distinguish between narrow gauge (1 067 mm) and high speed

(1 435 mm) rail, nor does it reference railroad stations. Level of service data for trains had to be taken from Japanese time tables (*Overseas Timetable*, 2000).

Using a timetable for a period as close as possible to the Fall of 1995 is not very critical for HSR, since the only service changes were the shortening of running times by a few minutes due to incremental increases in maximum speed. A major time table change did not occur until the opening of the Shinagawa station on October 1st, 2003.

A spreadsheet similar to the one shown in Table 4 on page 63 formed the basis to calculate level of service attributes for a matrix of 36 x 35 = 1 260 station pairs. Great care was taken to distinguish between Hikari and Kodama on one hand and Nozomi and Kodama service levels on the other because of the differences in fares, travel time, and number of transfers required. The semi-fast Hikaris stop more often, resulting in direct service between more station pairs, but are slower than Nozomi. Most station pairs require a short "feeder trip" at the beginning or end of the journey to get to or from the nearest Hikari or Nozomi station. The goal was to construct for each of the 1 260 station pairs and for both of the two service levels – Nozomi/Kodama (more expensive) and Hikari/Kodama (less expensive) – the best combination of travel time and number of transfers. For the tradeoff between transfers and travel time a transfer penalty of 20 minutes was assumed (Vrtic and Axhausen (2003) on page 26).

For each station pair involving both an origin and destination station only served by the slower Kodama trains eight different routings had to be considered: As the first leg of the journey go to the nearest Hikari station in the direction of travel or against the direction of travel (Table 4 on page 63). By the same token the optimal second Hikari/Kodama transfer point could be either to the north or to the south of the final destination station. Then the same alternatives had to be explored for a Nozomi/Kodama service combination. The work was done on two Excel spreadsheets with 8 Visual Basic subroutines each.

Trip times, fare estimates, and number of transfers for the close to 1 260 x 8 = 10 080 alternatives, and the choice of the best option for each of the 1 260 station pairs, and two train categories could be generated programmatically. For number of transfers just a few simple rules had to be kept in mind, e.g. no train goes beyond Tokyo Central Station, and no Kodama goes beyond Shin Osaka, making a transfer unavoidable.

The HSR station choice model did not include frequency effects. In order to obtain frequencies, all the alternatives would have had to be looked up manually in a Japanese time table and the number of daily frequencies counted by hand. Since all train categories operate several times per hour, a frequency variable would not likely have changed the estimates for station pair choice. Thus the considerable extra effort could not be justified.

The significantly longer time required to distinguish between the less expensive Hikari/Kodama and more expensive but faster Nozomi/Kodama services turned out to be spent in vain. In models employing both Hikari and Nozomi travel times and transfers as a determinant for station pair choice, only the coefficients for the Nozomi attributes turned out to be significant, so the Hikari attributes were not used in later models.

The information gathered from these different data sources was merged with the air and rail databases of the 1995 Japanese Intercity Travel Survey and formed the basis for the model estimations.

5.4 Calculation of Weights – Daily Extension Coefficients

5.4.1 A Choice Based Sample

The 1995 Japanese Intercity Travel Survey is a choice based sample. All five data sets, air, rail, car, bus, and ferry were collected using on-board surveys. The results were combined with other sources of aggregate data, such as cordon line counts and passenger ridership of each line. Daily and annual expansion factors were calculated based on these aggregate data. An Extension or Expansion Coefficient "extends" the sample to the original population. With the *daily* extension coefficient the Origin and Destination Table for a typical Fall day in 1995 can be recreated. The *annual* extension coefficient allows the researcher to restore the O & D table for the whole year and takes seasonal variations into account.

Table 20 - Distribution of the Daily Extension Coefficient for Air and HSR

Percentile	Air	HSR
Minimum	0.01	0.03
0.6		1
13.6	1	
5.0	0.92	3.21
25.0	1.06	6.93
50.0	1.25	10.52
Mean	*1.31*	*11.81*
75.0	1.51	20.05
95.0	1.90	21.83
Maximum	6.56	32.36

Table 20 shows the frequency distribution of the daily extension coefficient for air and rail. A daily extension coefficient of less than 1 means over sampling. Almost 14 % of the air records indicate over sampling. 90% of the expansion coefficients fall into a range of 0.9 to 1.9 for air and of 3 to 22 for rail. When compared to the actual number of trips, the air sample is about 10 times as large as the rail sample.

It is important to stress that we have choice based sampling at the level of station pairs and airport pairs, because the data was collected using on-board surveys. That means we have to use weights not only for (upper level) mode-choice models, but also for air-only and HSR-only terminal pair choice models.

5.4.2 Typical Use: Predicting Tokaido Line Ridership

Looking at Table 19 on page 181 we find that our 18 395 HSR observations make up about 60% of the number of original HSR passenger records within the 24 prefectures (30 765). That means we had to drop about 40% of the relevant observations (HSR travelers within the 24 prefectures).

We can use this information to estimate the actual number of daily passengers on the Tokaido (Tokyo – Osaka) line. Out of the 18 395 HSR trips 16 732 either originated or ended at one of the stations between Tokyo and Osaka, in other words made use of the Tokaido line. The sum of the weights of those passengers is 198 090 (Table 21). Those are the number of passengers represented by our sample. If the weights in our database are correct, dividing the 198 090 passengers by 60% should give us a good approximation of the number of daily passengers on the Tokaido line. We estimate 331 000 daily passengers.

Ridership on the Tokaido Shinkansen (Tokyo – Osaka) has been fairly constant at about 31 Million passengers / year or 361 000 passengers / day between 1993 and September 2003 (Jackson, 2004a). The slight difference between the two estimates is not surprising, since the actual number of daily trips is only an average over a 10 year period.

Table 21 - Comparison of Sample with Actual Passenger Volume on Tokaido Line

	Survey Observations		Actual Number of Daily Trips
	unweighted	weighted	
Number of Daily HSR Passengers on Tokaido (Tokyo - Osaka) Line	16,732	198,090	361,000
Sampling Rate	5%	55%	100%

Based on the actual number of daily trips the number of Tokaido line passengers represented by our sample is approximately 55% of the total number of daily passengers (Table 21). However the number of observations relating to Tokaido line passengers constitute only 5% of the actual number of daily trips. The two numbers differ by a factor of about 11, which is the mean daily expansion coefficient (Table 20).

Weights need to be taken into consideration during model estimation. However when using weights, we also need to normalize. Without normalizing the sum of the weights to the actual sample size (in our example 16 732) standard errors would be calculated based on a sample size of 198 090. That means on average standard errors would be underestimated by a factor of $\sqrt{11.81}$ or about 3.4. Normalization has little effect on the standard errors of airport choice models, which are underestimated on average by a factor of 1.1 without normalization.

All statistics in this dissertation from this point on forward are weighed by the daily extension coefficient. That means the distributions are not only representative of our dataset, but also of all air and HSR travel within the 24 prefecture area on a typical Fall day in 1995. The only exception is Table 30 on page 205.

5.5 *Calculation of Weights – Heteroskedasticity*

5.5.1 Zone Size Variability

The other issue related to the concept of using weights for the analysis is the problem of heteroskedasticity in the error term. A cause of concern is the great variability in the zone size, since access and egress distance is measured from only one point in each zone. Table 22 outlines the zone size distributions. 100 km^2 at about the 70th percentile is approximately the size of the city and county of San Francisco:

7	miles	x	7	miles	=	49	square miles
10	km	x	10	km	=	100	km^2

Please notice the interquartile range to be about 100 km^2 as well. Half the zones are about half the size of San Francisco or smaller.

Table 22 - Zone Size Distribution

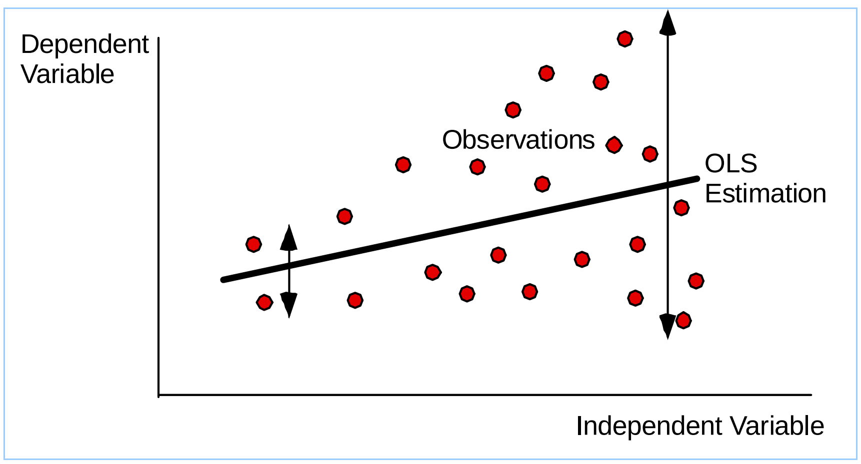
	Air		HSR	
Percentile	**Origin Zone Size** km^2	**Destination Zone Size** km^2	**Origin Zone Size** km^2	**Destination Zone Size** km^2
Minimum	3	2	3	3
5th	10	10	10	10
25th	20	19	22	20
Median	47	39	50	53
Mean	95	95	108	114
70th	**100**	**100**		
71st			**100**	**100**
75th	128	126	112	124
95th	290	290	364	364
Maximum	1146	1146	1146	1146

Table 22 highlights the significant interquartile range of 20 km^2 to 120 km^2.

5.5.2 Solution Outline

We could reasonably expect the error in the feeder distance measurement to be related to the size of the zone. The intuition behind the correction for heteroskedasticity is best illustrated with a simple linear regression example (Figure 18). Notice the higher absolute values of the residuals to the right. This indicates a relationship between the independent variable and the error term. Since ordinary least squares regression minimizes the squares of the residuals, a few additional large positive errors on the right would tilt the regression line up considerably. A few additional large negative errors would have the opposite effect. To overcome this instability we could weigh each observation in proportion to the error. Since observations with large errors now have less influence on the final estimate, different samples would move the regression line up and down on the right to about the same degree as on the left (Kennedy, 1992, pp. 116-117).

Figure 18 - Heteroskedasticity Illustration



The more formal treatment is a continuation of Subsection 4.4.1 ending on p. 154.

> Since the scale of the utilities of the logit model is inversely proportional to the standard error of the random utility components …, the basic assumption of a constant scale for all the observations is the same as the assumption of homoscedastic (or equal variance) random utilities (M. E. Ben-Akiva & Lerman, 1985, p. 204).

As an alternative to the assumption of a constant scale we will consider

$$\mu_n = \alpha_1 + \frac{\alpha_2}{\sqrt{S_n}} \tag{95}$$

where α_1 and α_2 are coefficients to be estimated and $\sqrt{S_n}$ is the square root of the sum of the origin and destination zone size for each individual.

(95) allows us to estimate how much zone size contributes to the error, and how this contribution varies with zone size. That knowledge can be used to construct the appropriate weights for all observations. Notice that in case of $\alpha_1 = 1$ and $\alpha_2 = 0$ $\mu = 1$.

We first estimate a preliminary terminal pair choice model described in the last chapter. The model is based on the assumption $\mu = 1$. Its utility function is

$$V_{in}^{\circ} = \hat{\boldsymbol{\beta}}'\mathbf{X} \tag{96}$$

where V_{in}° denotes the initial systematic utility of person *n* choosing Alternative *i*. The matrix $\mathbf{X}$ describes the observed characteristics of both decision maker *n* and alternative *i*. $\hat{\boldsymbol{\beta}}'$ is the vector of parameters. $\hat{\boldsymbol{\beta}}'$ is estimated using only the sample weights (daily extension coefficients).

However, what we are really interested in is V_{in} :

$$V_{in} = f(\text{zone size})V_{in}^{\circ} \tag{97}$$

We define our new utility function:

$$V_{in} = \left(\alpha_1 + \frac{\alpha_2}{\sqrt{S_n}} \right) V_{in}^{\circ} = \alpha_1 V_{in}^{\circ} + \alpha_2 \frac{V_{in}^{\circ}}{\sqrt{S_n}} \tag{98}$$

With $\alpha_1 > 0$ and $\alpha_2 > 0$ the larger the zone size, the smaller μ_n will be, and, by extension the larger the error σ_n. With μ_n being inversely proportional to σ_n, μ_n is the proper weight to be multiplied by the daily extension coefficient for each trip to form the composite weight.

In an ideal case the estimation result will be $\hat{\alpha}_1 = 1$ and $\hat{\alpha}_2 = 0$. That would mean zone size has no influence on the utility V_{in}, and an adjustment to our model to account for the heteroskedasticity of zone size is not needed. More likely though we will find $\hat{\alpha}_1$ being less than 1 and $\hat{\alpha}_2$ being significantly larger than 0, meaning a small adjustment is called for.

The coefficients of the terminal pair choice model can now be calculated in a third and final estimation using the weight: daily extension coefficient * μ_n. Weights have to be normalized to sum up to the sample size before the first (96) and renormalized before the 3rd estimation.

5.5.3 Estimation Results

This chapter's focus is on methodology. Estimation results are analyzed only from that perspective. Model interpretation is the subject of another chapter (Chapter 6).

Table 23 shows the estimation results for $\hat{\alpha}_1$ and $\hat{\alpha}_2$ for the airport pair choice model. Notice the close proximity to 1, and the low p-value for $\hat{\alpha}_1$. How closely the distribution of μ_n is centered around 1 becomes evident from Table 24. The coefficient estimate for $\hat{\alpha}_2$ is not significant, and a reestimation of the model with the inclusion of μ_n in the weight would not be necessary. Not surprisingly, Table 25 reveals that multiplication with μ_n changes the coefficient estimates very little.

For the HSR station choice model $\hat{\alpha}_2$ is highly significant (Table 26), and μ_n is distributed over a wider range (Table 27), but multiplication with μ_n still does not greatly influence coefficient estimates (Table 28). We can conclude that the air and rail data only exhibit mild heteroskedasticity.

Table 29 displays the global fit statistics for the airport pair and HSR station pair choice models. We can deduce from these global fit statistics what we were already able to discern from comparing coefficient estimates and standard errors: the models are fairly similar and, essentially, accounting for heteroskedasticity does not change the results.

The suffix "-2" in the model designation indicates that a correction for heteroskedasticity was made. All models presented from this point on forward have been corrected for heteroskedasticity and the "-2" suffix is often suppressed. Capital letters at the end of the model number as in R 300-2A or R 300-A denote minor variations in the model specification.

Table 23 - Parameter Estimates of α_1 and α_2 for Airport Pair Choice Model

Variable	Parameter Estimate	Standard Error	*p*-Value
α_1	0.97208	0.01840	<.0001
α_2	0.19149	0.20462	0.3493

Table 24 - Distribution of μ_n for Airport Pair Choice Model

Minimum	Mean	Maximum
0.97701321	0.98974652	1.0236835

Table 25 - Airport Pair Choice Model with and w/o Heteroskedasticity Correction

	A 1-1 Normalized Sample Weights only			A 1-2 Renormalized Sample Weights		
	Parameter Estimate	*Standard Error*	*p-Value*	*Parameter Estimate*	*Standard Error*	*p-Value*
Access Distance	-0.0497	0.0013	<.0001	-0.0498	0.0013	<.0001
log (Access Distance)	-0.6957	0.0484	<.0001	-0.6950	0.0484	<.0001
Acc Dist * Origin Pop Dens	9.43E-07	8.19E-08	<.0001	9.44E-07	8.19E-08	<.0001
Acc Dist * Vacation	0.0157	0.0010	<.0001	0.0157	0.0010	<.0001
Egress Distance	-0.0365	0.0013	<.0001	-0.0366	0.0013	<.0001
log (Egress Distance)	-0.7274	0.0535	<.0001	-0.7241	0.0535	<.0001
Egr Dist * Co Business	-0.0106	0.0011	<.0001	-0.0107	0.0011	<.0001
log (Air Frequency)	0.9898	0.0440	<.0001	0.9897	0.0440	<.0001
Air Time	-0.0374	0.0059	<.0001	-0.0374	0.0059	<.0001
Fare / Avg Zonal Income	-2.33E-04	2.72E-05	<.0001	-2.33E-04	2.73E-05	<.0001
Tokyo Narita	-0.5719	0.1467	<.0001	-0.5745	0.1467	<.0001
Fukuoka Itazuke	-1.4987	0.2929	<.0001	-1.4955	0.2939	<.0001
Osaka Itami	1.2416	0.2760	<.0001	1.2495	0.2767	<.0001
Kagoshima	-1.6147	0.3007	<.0001	-1.6172	0.3017	<.0001
Miyazaki	-1.4152	0.3018	<.0001	-1.4172	0.3028	<.0001
Nagoya	0.4710	0.2140	0.0277	0.4774	0.2145	0.0260
Nagasaki	-1.8139	0.2921	<.0001	-1.8133	0.2931	<.0001
Oita	-1.9748	0.3018	<.0001	-1.9732	0.3027	<.0001
Kumamoto	-2.2217	0.2964	<.0001	-2.2220	0.2973	<.0001
Hiroshima	0.2604	0.3014	0.3876	0.2657	0.3023	0.3795
Osaka Kansai	1.7559	0.2771	<.0001	1.7647	0.2777	<.0001
Yamaguchi Ube	-2.8036	0.3152	<.0001	-2.8029	0.3162	<.0001
Okayama	0.0666	0.2860	0.8159	0.0730	0.2868	0.7992
Fukuoka Kita-Kyushu	-3.5809	0.3277	<.0001	-3.5774	0.3286	<.0001
Tokyo Haneda	0.0000	.	.	0.0000	.	.

Table 26 - Parameter Estimates of α_1 and α_2 for HSR Station Pair Choice Model

Variable	Parameter Estimate	Standard Error	*p*-Value
α_1	0.89424	0.01286	<.0001
α_2	1.22214	0.14154	<.0001

Table 27 - Distribution of μ_n for HSR Station Pair Choice Model

Minimum	Mean	Maximum
0.92564586	1.00604716	1.27230181

Table 28 - HSR Sta Pair Choice Model with and w/o Heteroskedasticity Correction

	R 300-1 Normalized Sample Weights only			R 300-2 Renormalized Sample Weights		
	Parameter Estimate	*Standard Error*	*p-Value*	*Parameter Estimate*	*Standard Error*	*p-Value*
Origin Sta = Tokyo Central	0.7385	0.0503	<.0001	0.7346	0.0505	<.0001
Dest Sta = Tokyo Central	0.6736	0.0500	<.0001	0.6838	0.0502	<.0001
Origin Sta = Shin Osaka	0.6743	0.0425	<.0001	0.6682	0.0425	<.0001
Dest Sta = Shin Osaka	0.6050	0.0415	<.0001	0.6007	0.0416	<.0001
Nozomi Origin Sta	1.5041	0.0362	<.0001	1.5077	0.0364	<.0001
Nozomi Dest Sta	1.4458	0.0344	<.0001	1.4440	0.0347	<.0001
Access Distance	-0.0132	0.0011	<.0001	-0.0133	0.0011	<.0001
log (Access Distance)	-0.8636	0.0158	<.0001	-0.8674	0.0160	<.0001
Acc Dist * Co Business	0.0020	6.97E-04	0.0041	0.0020	6.99E-04	0.0048
Acc Dist * Veh Ownership	-6.77E-05	7.90E-06	<.0001	-6.71E-05	7.91E-06	<.0001
Acc Dist * Linehaul Time	4.12E-05	2.34E-06	<.0001	4.13E-05	2.35E-06	<.0001
Acc Dist * Origin Pop Dens	2.06E-07	4.04E-08	<.0001	2.14E-07	4.01E-08	<.0001
Acc Dist * Vacation	0.0045	7.92E-04	<.0001	0.0045	7.96E-04	<.0001
Acc Dist * (Age < 30)	5.44E-04	4.74E-04	0.2508	5.14E-04	4.76E-04	0.2802
Egress Distance	-0.0190	0.0011	<.0001	-0.0191	0.0011	<.0001
log (Egress Distance)	-0.6905	0.0140	<.0001	-0.6951	0.0141	<.0001
Egr Dist * Co Business	0.0028	7.54E-04	0.0003	0.0028	7.59E-04	0.0003
Egr Dist * Male	0.0015	5.79E-04	0.0109	0.0015	5.83E-04	0.0115
Egr Dist * Veh Ownership	-1.66E-05	7.28E-06	0.0227	-1.63E-05	7.31E-06	0.0261
Egr Dist * Linehaul Time	3.08E-05	2.60E-06	<.0001	3.10E-05	2.62E-06	<.0001
Egr Dist * Dest Pop Dens	1.65E-07	4.05E-08	<.0001	1.70E-07	4.04E-08	<.0001
Egr Dist * Vacation	0.0022	8.76E-04	0.0133	0.0022	8.82E-04	0.0140
Egr Dist * (Age < 30)	0.0013	4.59E-04	0.0046	0.0013	4.62E-04	0.0052
Nozomi Time	-0.0187	5.19E-04	<.0001	-0.0187	5.23E-04	<.0001
Nozomi Transfers	-0.3407	0.0325	<.0001	-0.3426	0.0326	<.0001

Table 29 - Global Fit for Airport and Station Pair Choice Models with and w/o Heteroskedasticity Correction

	A 1-1 Normalized Sample Weights only	A 1-2 Renormalized Sample Weights * μ_n	R 300-1 Normalized Sample Weights only	R 300-2 Renormalized Sample Weights * μ_n
Number of Individuals	30,126	30,126	18,395	18,395
Number of Choices per Individual	82	82	300	300
Number of Records	2,470,332	2,470,332	5,518,500	5,518,500
Degrees of Freedom (df)	24	24	25	25
$\ell\ell(0)$	-133,830.685	-133,837.870	-108,492.790	-107,618.180
$\ell\ell(\hat{\beta})$	-13,313.371	-13,298.322	-25,060.760	-24,760.941
Likelihood Ratio Test Statistic	241,034.628	241,079.097	166,864.060	165,714.479
ρ^2	0.9005	0.9006	0.7690	0.7699
$\bar{\rho}^2$	0.9003	0.9005	0.7688	0.7697

$\ell\ell(0)$ is the value of the log likelihood function when all parameters are zero.

$\ell\ell(\hat{\beta})$ is the value of the log likelihood function at its maximum.

Likelihood Ratio Test Statistic (H_0: all β are 0) = $-2[\ell\ell(0) - \ell\ell(\hat{\beta})]$

$\rho^2 = 1 - \frac{\ell\ell(\hat{\beta})}{\ell\ell(0)}$

$\bar{\rho}^2 = 1 - \frac{\ell\ell(\hat{\beta}) - \text{df}}{\ell\ell(0)}$

This outcome is not surprising when compared to previous research. Munizaga, Heydecker, and Ortúzar (2000) distinguish between *heteroskedasticity between options* (e.g. due to different levels of information between frequently and rarely selected options) and *heteroskedasticity between observations* (e.g. in mixed data experiments where both RP and SP survey observations are used). Our case is *continuous heteroskedasticity between observations*, because all routing options for each individual via alternative terminal pairs begin in the same origin zone and end in the same destination zone. Munizaga, et al. used simulated data to test the performance of different model structures both in terms of their ability to recover the known parameters and in terms of their prediction ability related to policy changes. In the case of heteroskedasticity between options, either the heteroskedastic extreme value (Bhat, 1995) or the multinomial probit model had to be used. However, in the case of discrete heteroskedasticity between observations (2 groups of observations) the multinomial logit model performed as well as more complex models. This illustrates the robustness of the multinomial logit model with respect to some heteroskedasticity violations.

After

1. carefully selecting the choice set in order not to bias access coefficients (Section 5.2),

2. accounting for peculiarities of the data collection by using daily extension coefficients (Section 5.4),

3. normalizing the sample weights to obtain correct standard error estimates (Section 5.4), and

4. compensating for the effects of a mildly heteroskedastic error term caused by variations in zone size (Section 5.5),

we can now be confident that all the conditions for the application of the multinomial logit model are met. By using the appropriate weights in the estimation, we have caused the error term to be independently and identically distributed (IID). In the next section we shall see how we have used this IID property to our advantage by using a randomized choice set to estimate a HSR station choice model from a data set with over 23 million records.

5.6 Randomized Choice Sets

5.6.1 Computational Resources

30 126 air travelers with 82 choices each, including the chosen one, result in a data set with 2 470 332 records. The resulting file size is small enough to allow an airport pair choice model using the complete choice set to be estimated in about one hour. The same, however, is not possible for HSR station pair choice models.

There are 18 395 HSR observations for lower nest estimation. Each represents one actual traveler who made a specific station pair choice. Each traveler has a choice set of 36 x 35 or 1 260 station pairs, one of which is chosen, and 1 259 of which are not chosen. This means our database has 18 395 x 1 260 = 23 177 700 records and, depending on the number of variables, can reach a size of up to 10 gigabytes (GB). Estimating a model using the complete choice set on the available equipment was not feasible. We needed to make use of randomized choice sets.

5.6.2 A Consistent Estimator

Taking advantage of the independence of irrelevant alternatives (IIA) property of multinomial logit, McFadden has shown that estimation with just a subset of the alternatives including the chosen one yields consistent estimates (M. E. Ben-Akiva & Lerman, 1985, p. 261ff). It is imperative to keep in mind that consistency is the weakest of the desirable properties of an estimator. It only says that with an infinite sample size the estimator is unbiased. It says nothing about how *fast* the estimator approaches that limit

(Wonnacott & Wonnacott, 1979, p. 65). In asymptotic distribution theory we substitute an infinite sample for our limited sample to derive useful results. So the question arises, how many observations are needed in practice to get useful approximations.

> "The simplest truthful answer is that no one knows, at least not without additional information …" (Ruud, 2000, p. 267).

The sample size of all HSR station pair models is constant, $N = 18\,395$. Only the number of choices varies. However, the estimator for simple random sampling of alternatives exhibits the same properties with an increasing number of alternatives as we would expect with a growing sample size. Some estimates appear to converge on a true value (Figure 19 - Figure 21).

We have found that the speed with which parameter estimates converge on one single value is greatly dependent on the correct specification of the model. The more accurately a model reflects the underlying data the faster the coefficient estimates converge.

A summary of the estimation results of an earlier incorrectly specified model with a slightly larger choice set is presented in Table 30. The decreasing estimates for access and egress distance are obvious. The (incorrectly specified) simple linear model with 1 482 choices was estimated on a Linux machine with 1.5 GB of RAM in 12 hours. The parameter estimates in Table 30 for the randomly selected choices reflect the mean estimates resulting from four independently drawn choice sets for each number of choices. The p-value for each coefficient in each of the 16 estimations was <0.0001.

> "With classical statistical inference methods, the specification of the model is never in question; the model is virtually always assumed to be correct." (M. E. Ben-Akiva & Lerman, 1985, p. 155)

Those *p*-values tell us that based on the assumptions of

a. correct model specification, and

b. an infinite sample size,

the probability of the population parameters to be equal to zero is virtually zero. If only one of the two assumptions were incorrect, the *p*-values would be meaningless. In this case neither one of the two conditions is met.

Now consider Table 31 on page 206. A remarkable result is the accuracy of model R10, which was estimated using only nine out of 1 259 choices, plus the chosen one. All coefficients not only have the correct sign but are also of the proper magnitude. The only exception is access distance times a dummy variable indicating the traveler's age to be below 30 years. With consistently high standard errors it can be safely classified as non-explanatory variable, so the change of sign is of no concern. This level of accuracy with only ten random choices, and the high level of consistency between very different choice sets is only possible with a well specified model.

Expanding the choice set by and large causes the standard errors to become markedly smaller, which is expected. The spread of the parameter estimates with several random draws of the same size is also considerably reduced (not illustrated in Table 31). But most importantly, the standard errors generally become more and more reliable. To illustrate this last point, consider Nozomi Transfers. The lower and upper bounds for the 95% confidence interval in model R10 are –0.71 and –0.40 respectively. Assuming that model R300 is the most accurate model of the four shown in Table 31, and comparing the 95% confidence interval of R10 to the coefficient estimate of R300 (–0.34), we can surmise that the standard error for this variable in model R10 is not very likely to be correct. This is also apparent from Figure 19.

An exception to standard errors becoming more reliable with an increasing number of choices appears to be the variable Access Distance * Vehicle Ownership. While being consistently significant the parameter estimate is unstable, that means dependent on the particular sample (all samples of choices were drawn independently and are not nested within each other). A true 95% confidence interval would be much larger than the standard errors indicate.

An informal convergence criterion for the selection of an appropriate size for the choice set is illustrated in Table 32. 22 of the 25 coefficient estimates of model R100 lie in the corresponding 95% confidence interval of R300. After tripling the sample size we find that our R100 coefficient estimates rest within the ranges of values that we are now 95% confident the true parameters are situated in. That in itself would indicate that a sample size of 100 may be sufficient. It is the three coefficient estimates which do not meet the informal convergence criterion that force us to look beyond a sample size of 100.

With Nozomi Transfers, as with some of the other variables, we can clearly see the parameter estimate increasing or decreasing as the sample size increases. It cannot be ruled out that these are biased estimates converging on a true value. This in itself is enough reason to prefer R300 over R100, even though the former requires 6 h computer time versus 8 min for the latter. In summary, even though the four models in Table 31 show a high level of consistency while using vastly different choice sets, in order to attain the most precise estimates both for parameters and standard errors, we elected to use a choice set of 299 randomly selected station pair alternatives, plus the selected option.

Global fit statistics for models R10 through R300 are presented in Table 33. They illustrate another interesting pattern one should be aware of, especially when evaluating the global fit of a model without the benefit of having estimates based on different sample sizes shown right next to each other. As the loglikelihood decreases with an increasing number of choices (more and more small probabilities are multiplied with each other), the *difference* between $\ell\ell(\hat{\beta})$ and $\ell\ell(0)$, represented by the likelihood ratio, increases, while the *ratio*, represented by ρ^2, decreases. The likelihood ratio allows to test nested, while ρ^2 is helpful to evaluate non-nested hypotheses. Table 33 demonstrates that when comparing models using different sample sizes, a ρ^2 of 0.7699 does not indicate a worse fit than a model with a ρ^2 of 0.8941. Exactly the opposite is true in this case.

When analyzing estimation results attained by randomizing choice sets one cannot rely on a single model blindly trusting confidence intervals. One can only properly understand the significance of individual mode choice determinants and get an idea of the magnitude of their coefficients by looking at a *set* of models like those presented in Table 31.

That is the approach to model interpretation taken in the next chapter.

Figure 19 - Nozomi Transfers Confidence Intervals for Different Size Choice Sets

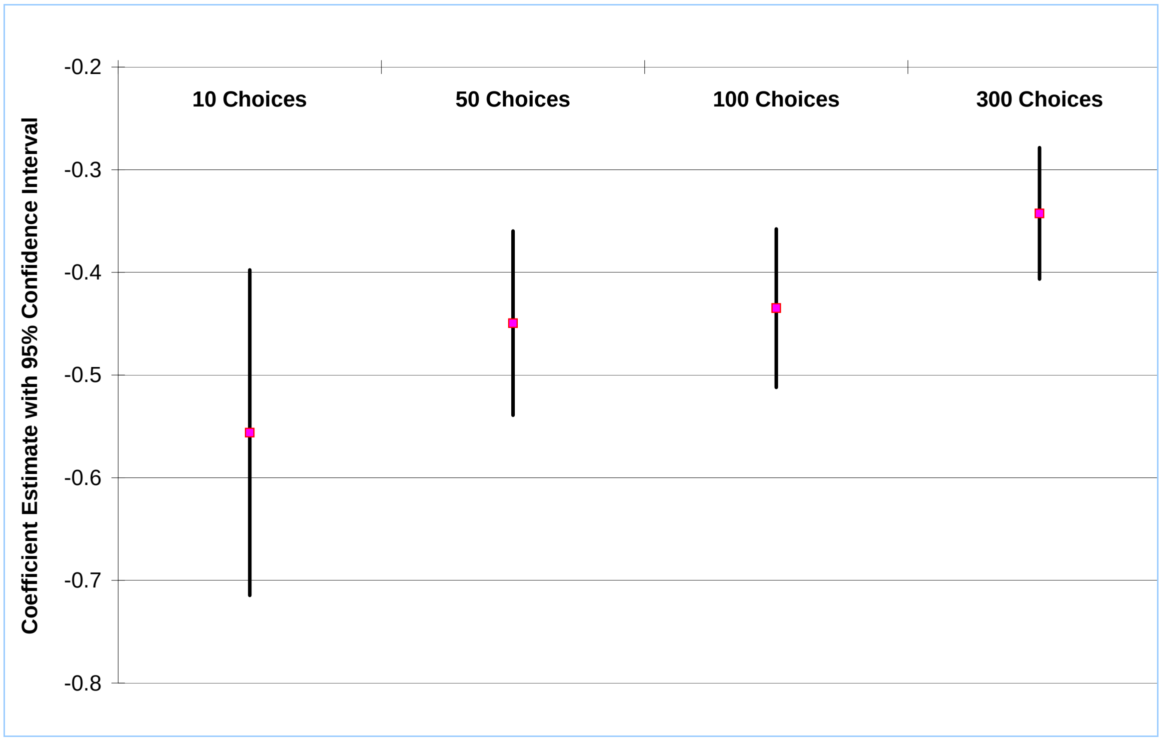


Figure 20 - Nozomi Time Confidence Intervals for Different Size Choice Sets

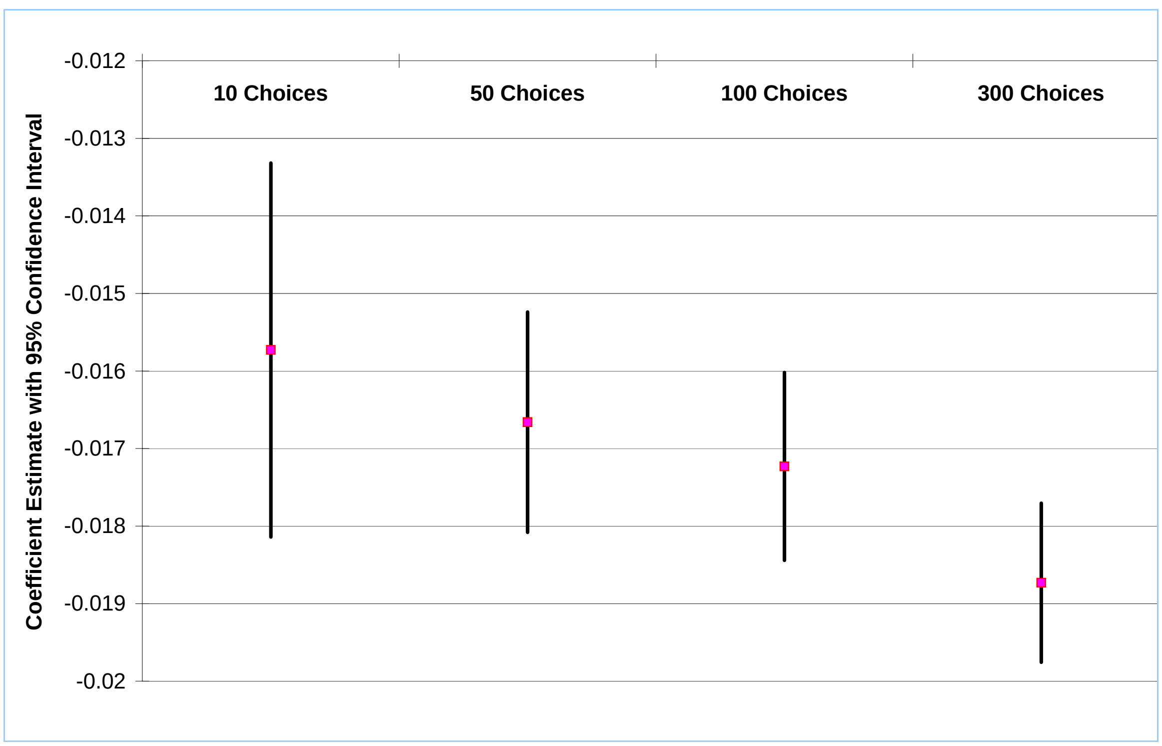


Figure 21 - Shin Osaka Origin Confidence Intervals for Different Size Choice Sets

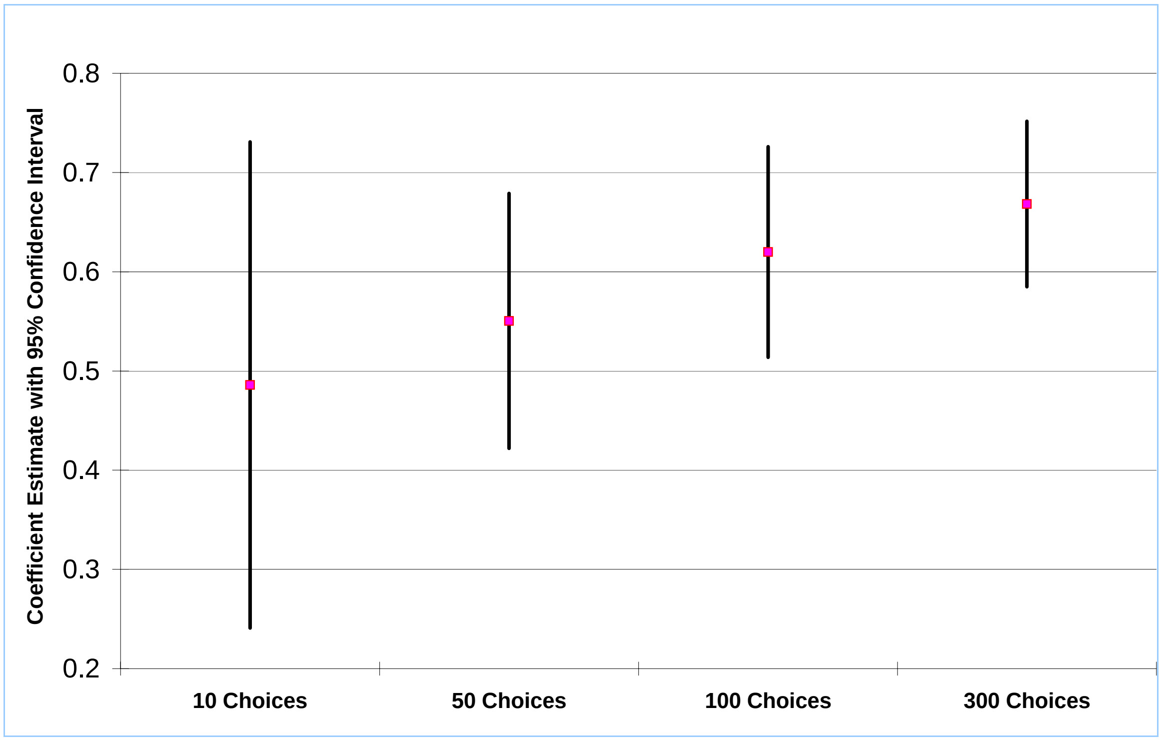


Table 30 - Effect of Number of Random Choices on a Misspecified Model

	Number of Randomly Selected Choices				Complete Choice Set
	10	40	100	300	1,482
Access Distance	-0.141	-0.165	-0.179	-0.194	-0.216
Egress Distance	-0.139	-0.156	-0.172	-0.184	-0.203
Nozomi Origin	1.680	1.631	1.615	1.641	1.625
Nozomi Destination	1.571	1.604	1.551	1.537	1.530
Nozomi Time	-0.004	-0.005	-0.005	-0.005	-0.005
Nozomi Transfers	-0.580	-0.549	-0.569	-0.556	-0.570

Table 31 - Parameter Estimates of HSR Station Choice Models with Choice Sets of Different Sizes

	R 10-2 10 Random Choices		**R 50-2 50 Random Choices**		**R 100-2 100 Random Choices**		**R 300-2 300 Random Choices**	
	Parameter Estimate	*Standard Error*	*Parameter Estimate*	*Standard Error*	*Parameter Estimate*	*Standard Error*	*Parameter Estimate*	*Standard Error*
Origin Sta = Tokyo Central	0.587	0.137	0.711	0.076	0.716	0.063	0.735	0.050
Dest Sta = Tokyo Central	0.588	0.144	0.545	0.077	0.646	0.064	0.684	0.050
Origin Sta = Shin Osaka	0.486	0.125	0.550	0.066	0.620	0.054	0.668	0.043
Dest Sta = Shin Osaka	0.389	0.124	0.587	0.068	0.588	0.054	0.601	0.042
Nozomi Origin Sta	1.552	0.090	1.608	0.052	1.559	0.044	1.508	0.036
Nozomi Dest Sta	1.562	0.092	1.524	0.050	1.494	0.042	1.444	0.035
Access Distance	-0.010	0.002	-0.014	0.001	-0.015	0.001	-0.013	0.001
log (Access Distance)	-1.064	0.045	-0.892	0.023	-0.881	0.019	-0.867	0.016
Acc Dist * Co Business	*1.48E-03*	*1.56E-03*	*1.64E-03*	*9.08E-04*	1.98E-03	7.98E-04	1.97E-03	6.99E-04
Acc Dist * Veh Ownership	-4.52E-05	1.58E-05	-4.18E-05	9.99E-06	-3.58E-05	8.85E-06	-6.71E-05	7.91E-06
Acc Dist * Linehaul Time	2.93E-05	4.56E-06	3.67E-05	3.00E-06	3.77E-05	2.72E-06	4.13E-05	2.35E-06
Acc Dist * Origin Pop Dens	1.73E-07	8.01E-08	2.38E-07	4.97E-08	2.82E-07	4.41E-08	2.14E-07	4.01E-08
Acc Dist * Vacation	3.52E-03	1.74E-03	4.27E-03	1.02E-03	4.22E-03	9.07E-04	4.47E-03	7.96E-04
Acc Dist * (Age < 30)	*-7.54E-04*	*1.07E-03*	*-1.05E-04*	*6.32E-04*	*2.17E-04*	*5.54E-04*	*5.14E-04*	*4.76E-04*
Egress Distance	-0.019	0.003	-0.019	0.001	-0.019	0.001	-0.019	0.001
log (Egress Distance)	-0.760	0.040	-0.697	0.021	-0.673	0.018	-0.695	0.014
Egr Dist * Co Business	*3.48E-03*	*2.04E-03*	*1.83E-03*	*1.01E-03*	1.99E-03	8.75E-04	2.75E-03	7.59E-04
Egr Dist * Male	2.78E-03	1.18E-03	2.22E-03	7.78E-04	1.82E-03	6.61E-04	1.47E-03	5.83E-04
Egr Dist * Veh Ownership	-3.39E-05	1.68E-05	*-6.16E-06*	*9.99E-06*	*-1.05E-05*	*8.56E-06*	-1.63E-05	7.31E-06
Egr Dist * Linehaul Time	2.62E-05	5.14E-06	2.36E-05	3.49E-06	2.83E-05	2.99E-06	3.10E-05	2.62E-06
Egr Dist * Dest Pop Dens	*1.10E-07*	*8.85E-08*	1.29E-07	5.33E-08	1.52E-07	4.69E-08	1.70E-07	4.04E-08
Egr Dist * Vacation	4.60E-03	2.19E-03	*1.58E-03*	*1.14E-03*	*1.35E-03*	*1.00E-03*	2.17E-03	8.82E-04
Egr Dist * (Age < 30)	*9.02E-04*	*1.07E-03*	1.29E-03	6.30E-04	1.42E-03	5.39E-04	1.29E-03	4.62E-04
Nozomi Time	-0.016	0.001	-0.017	0.001	-0.017	0.001	-0.019	0.001
Nozomi Transfers	-0.556	0.081	-0.449	0.046	-0.435	0.039	-0.343	0.033

Table 32 - Informal Convergence Criterion for Selection of Suitable Choice Set Size

R 10-2 **10 Random Choices**	**R 50-2** **50 Random Choices**	**R 100-2** **100 Random Choices**	**R 300-2** **300 Random Choices**
Parameter Estimate within R 300 Confidence Interval	*Parameter Estimate within R 300 Confidence Interval*	*Parameter Estimate within R 300 Confidence Interval*	*Parameter Estimate within R 300 Confidence Interval*
	Origin Sta = Tokyo Central	Origin Sta = Tokyo Central	Origin Sta = Tokyo Central
Dest Sta = Tokyo Central		Dest Sta = Tokyo Central	Dest Sta = Tokyo Central
		Origin Sta = Shin Osaka	Origin Sta = Shin Osaka
	Dest Sta = Shin Osaka	Dest Sta = Shin Osaka	Dest Sta = Shin Osaka
Nozomi Origin Sta		Nozomi Origin Sta	Nozomi Origin Sta
		Nozomi Dest Sta	Nozomi Dest Sta
	Access Distance	Access Distance	Access Distance
	log (Access Distance)	log (Access Distance)	log (Access Distance)
Acc Dist * Co Business	Acc Dist * Co Business	Acc Dist * Co Business	Acc Dist * Co Business
			Acc Dist * Veh Ownership
	Acc Dist * Linehaul Time	Acc Dist * Linehaul Time	Acc Dist * Linehaul Time
Acc Dist * Origin Pop Dens	Acc Dist * Origin Pop Dens	Acc Dist * Origin Pop Dens	Acc Dist * Origin Pop Dens
Acc Dist * Vacation	Acc Dist * Vacation	Acc Dist * Vacation	Acc Dist * Vacation
	Acc Dist * (Age < 30)	Acc Dist * (Age < 30)	Acc Dist * (Age < 30)
Egress Distance	Egress Distance	Egress Distance	Egress Distance
	log (Egress Distance)	log (Egress Distance)	log (Egress Distance)
Egr Dist * Co Business	Egr Dist * Co Business	Egr Dist * Co Business	Egr Dist * Co Business
	Egr Dist * Male	Egr Dist * Male	Egr Dist * Male
	Egr Dist * Veh Ownership	Egr Dist * Veh Ownership	Egr Dist * Veh Ownership
Egr Dist * Linehaul Time		Egr Dist * Linehaul Time	Egr Dist * Linehaul Time
Egr Dist * Dest Pop Dens	Egr Dist * Dest Pop Dens	Egr Dist * Dest Pop Dens	Egr Dist * Dest Pop Dens
	Egr Dist * Vacation	Egr Dist * Vacation	Egr Dist * Vacation
Egr Dist * (Age < 30)	Egr Dist * (Age < 30)	Egr Dist * (Age < 30)	Egr Dist * (Age < 30)
			Nozomi Time
			Nozomi Transfers
10	17	22	25

Table 33 - Global Fit Statistics of HSR Station Choice Models with Choice Sets of Different Sizes

	R 10-2 10 Random Choices	R 50-2 50 Random Choices	R 100-2 100 Random Choices	R 300-2 300 Random Choices
Number of Individuals	18,395	18,395	18,395	18,395
Number of Choices per Individual	10	50	100	300
Number of Records	183,950	919,750	1,839,500	5,518,500
Degrees of Freedom (df)	25	25	25	25
$\ell\ell(0)$	-45,039.288	-74,648.795	-87,400.825	-107,618.180
$\ell\ell(\hat{\beta})$	-4,769.568	-10,450.842	-14,821.073	-24,760.941
Likelihood Ratio Test Statistic	80,539.439	128,395.906	145,159.504	165,714.479
ρ^2	0.8941	0.8600	0.8304	0.7699
$\overline{\rho}^2$	0.8935	0.8597	0.8301	0.7697

$\ell\ell(0)$ is the value of the log likelihood function when all parameters are zero.

$\ell\ell(\hat{\beta})$ is the value of the log likelihood function at its maximum.

Likelihood Ratio Test Statistic (H_0: all β are 0) = $-2[\ell\ell(0) - \ell\ell(\hat{\beta})]$

$$\rho^2 = 1 - \frac{\ell\ell(\hat{\beta})}{\ell\ell(0)}$$

$$\overline{\rho}^2 = 1 - \frac{\ell\ell(\hat{\beta}) - \text{df}}{\ell\ell(0)}$$

5.6.3 Interactions

The previous sections and subsections discussed sample weights (daily extension coefficients), weights based on zone size, normalization, and randomized choice sets. A question may come up on how the randomizing of choice sets may affect the daily extension coefficients and the weights based on zone size. Is there an interaction? The rule is that when taking a subsample, weights are not affected, as long as they are taken into the subsample. A randomly chosen subsample will have, on average, the same proportion of weights as the original sample.

Another issue arises when renormalizing a randomized choice set. Common techniques used to generate a randomized choice set result in a highly varying number of choices for each individual. That means the sum of the weights is no longer meaningful, since each observation does not contribute to the sum of weights in proportion to its own weight. In order not to complicate the renormalization process unnecessarily, we employed an algorithm developed by SAS Institute which guarantees the same number of choices (not the same choices!) for each individual. Please see Appendix 3 on page 287 for more details.

5.7 Randomizing Observations versus Randomizing Choices

The Independence of Irrelevant Alternatives allowed us to randomize choices.

$$\frac{P_{in}}{P_{kn}} = \frac{\dfrac{e^{\boldsymbol{\beta}_i \mathbf{X}_{in}}}{\sum_{\forall j} e^{\boldsymbol{\beta}_j \mathbf{X}_{jn}}}}{\dfrac{e^{\boldsymbol{\beta}_k \mathbf{X}_{kn}}}{\sum_{\forall j} e^{\boldsymbol{\beta}_j \mathbf{X}_{jn}}}} = \frac{e^{\boldsymbol{\beta}_i \mathbf{X}_{in}}}{e^{\boldsymbol{\beta}_k \mathbf{X}_{kn}}} \tag{99}$$

(99) illustrates the peculiar property of multinomial logit models that the ratio of probabilities of choosing two alternatives is independent of all other available choices. This allows us to estimate choice probabilities on only a subset of alternatives. It also allows us to only consider two modes when estimating the nested mode choice model.

When using any other model we have to sample observations. This second method is intuitively unbiased, since drawing a random sample of 1000 individuals from a certain population and then randomly drawing 100 observations from that dataset is equivalent to randomly picking only 100 people in the first place.

The problem with the second method is that when drawing a sample of 100 observations (with 1260 alternatives each), from the total of 18 395 observations that represent all HSR travelers, our data subset with 126 000 records only contains 100 choices. When sampling 10 alternatives our data subset with 183 950 records contains 18 395 choices. Multinomial logit models allow us to obtain more precise estimates with a given expense of computing resources.

6 Model Results

6.1 Airport Pair Choice and HSR Station Pair Choice Models

The previous chapter discussed a number of models to illustrate the methodology used. This chapter refers to some of the same models to draw conclusions related to the topic of this dissertation: Airport and Station Accessibility as A Determinant of Mode Choice.

6.1.1 Base Models

Both the airport pair choice model A 1-2 (Table 34) and the HSR station pair choice model H 300-2 in Table 35 were first introduced in the context of heteroskedasticity correction (5.5.3 Estimation Results on page 192). HSR station pair choice models had to be estimated using randomized choice sets which were discussed in the section directly following it. While it is always prudent to rely on more than one model estimation to reach conclusions, it is almost essential to do this when using a random sample of alternatives. Besides model H 300 (based on a random sample of 299 alternatives plus the chosen one) the coefficient estimates of R 50 and R 100 are also presented in Table 35, allowing us to consider a set of models for the interpretation of results. Model R 10 is not included in Table 35 because a choice set of 10 alternatives did not produce accurate enough estimates.

The airport pair choice model A 1-2 is based on the complete choice set and a larger sample size (almost 30 000 observations). Estimates are therefore less subject to variation which will allow us to draw more precise conclusions than with HSR station pair models. The slightly different presentation format of coefficient estimates for air versus HSR is a visual reminder of this distinction. Table 34 displays odds ratios (4.3.4 Relevant Applications of the Insights Gained on page 139) which help make the interpretation of coefficient estimates very intuitive.

Table 34 - Airport Pair Choice Model with 95% Confidence Interval and Odds Ratio

Model A 1-2

	95% Confidence Interval				
	Lower Bound	*Parameter Estimate*	*Upper Bound*	*p-Value*	*Odds Ratio*
Airport Access Distance	-0.052	-0.050	-0.047	<.0001	0.951
log (Airport Access Distance)	-0.790	-0.695	-0.600	<.0001	0.499
Acc Dist * Origin Pop Dens	7.84E-07	9.44E-07	1.10E-06	<.0001	1.000
Acc Dist * Vacation	0.014	0.016	0.018	<.0001	1.016
Airport Egress Distance	-0.039	-0.037	-0.034	<.0001	0.964
log (Airport Egress Distance)	-0.829	-0.724	-0.619	<.0001	0.485
Egr Dist * Co Business	-0.013	-0.011	-0.009	<.0001	0.989
log (Air Frequency)	0.903	0.990	1.076	<.0001	2.690
Air Time	-0.049	-0.037	-0.026	<.0001	0.963
Air Fare / Home Zone Income	-2.86E-04	-2.33E-04	-1.79E-04	<.0001	1.000
Tokyo Narita	-0.862	-0.575	-0.287	<.0001	0.563
Fukuoka Itazuke	-2.071	-1.496	-0.920	<.0001	0.224
Osaka Itami	0.707	1.250	1.792	<.0001	3.489
Kagoshima	-2.209	-1.617	-1.026	<.0001	0.198
Miyazaki	-2.011	-1.417	-0.824	<.0001	0.242
Nagoya	0.057	0.477	0.898	0.026	1.612
Nagasaki	-2.388	-1.813	-1.239	<.0001	0.163
Oita	-2.567	-1.973	-1.380	<.0001	0.139
Kumamoto	-2.805	-2.222	-1.639	<.0001	0.108
Hiroshima	-0.327	0.266	0.858	0.380	1.304
Osaka Kansai	1.220	1.765	2.309	<.0001	5.840
Yamaguchi Ube	-3.423	-2.803	-2.183	<.0001	0.061
Okayama	-0.489	0.073	0.635	0.799	1.076
Fukuoka Kita-Kyushu	-4.221	-3.577	-2.933	<.0001	0.028
Tokyo Haneda	.	0.000	.	.	.

Table 35 - Three HSR Station Pair Models with 95% Confidence Interval

	R 50-2	R 100-2	R 300-2		
			95% Confidence Interval		
	Parameter Estimate	*Parameter Estimate*	*Lower Bound*	*Parameter Estimate*	*Upper Bound*
Origin Sta = Tokyo Central	0.711	0.716	0.636	0.735	0.834
Dest Sta = Tokyo Central	0.545	0.646	0.585	0.684	0.782
Origin Sta = Shin Osaka	0.550	0.620	0.585	0.668	0.752
Dest Sta = Shin Osaka	0.587	0.588	0.519	0.601	0.682
Nozomi Origin Sta	1.608	1.559	1.436	1.508	1.579
Nozomi Dest Sta	1.524	1.494	1.376	1.444	1.512
Access Distance	-0.014	-0.015	-0.015	-0.013	-0.011
log (Access Distance)	-0.892	-0.881	-0.899	-0.867	-0.836
Acc Dist * Co Business	_1.64E-03_	1.98E-03	6.01E-04	1.97E-03	3.34E-03
Acc Dist * Veh Ownership	-4.18E-05	-3.58E-05	-8.26E-05	-6.71E-05	-5.16E-05
Acc Dist * Linehaul Time	3.67E-05	3.77E-05	3.67E-05	4.13E-05	4.59E-05
Acc Dist * Origin Pop Dens	2.38E-07	2.82E-07	1.36E-07	2.14E-07	2.93E-07
Acc Dist * Vacation	4.27E-03	4.22E-03	2.91E-03	4.47E-03	6.03E-03
Acc Dist * (Age < 30)	_-1.05E-04_	_2.17E-04_	_-4.19E-04_	_5.14E-04_	_1.45E-03_
Egress Distance	-0.019	-0.019	-0.021	-0.019	-0.017
log (Egress Distance)	-0.697	-0.673	-0.723	-0.695	-0.667
Egr Dist * Co Business	_1.83E-03_	1.99E-03	1.26E-03	2.75E-03	4.24E-03
Egr Dist * Male	2.22E-03	1.82E-03	3.28E-04	1.47E-03	2.61E-03
Egr Dist * Veh Ownership	_-6.16E-06_	_-1.05E-05_	-3.06E-05	-1.63E-05	-1.98E-06
Egr Dist * Linehaul Time	2.36E-05	2.83E-05	2.59E-05	3.10E-05	3.61E-05
Egr Dist * Dest Pop Dens	1.29E-07	1.52E-07	9.10E-08	1.70E-07	2.49E-07
Egr Dist * Vacation	_1.58E-03_	_1.35E-03_	4.42E-04	2.17E-03	3.90E-03
Egr Dist * (Age < 30)	1.29E-03	1.42E-03	3.85E-04	1.29E-03	2.19E-03
Nozomi Time	-0.017	-0.017	-0.020	-0.019	-0.018
Nozomi Transfers	-0.449	-0.435	-0.407	-0.343	-0.279

All parameter estimates are significant at the 5% level, except when in *italic*.

Italic variable names **indicate that R 100 coefficient does not fall into R 300 Confidence Interval.**

Section 6.3 introduces variants to these base models. The HSR alternatives (Table 45 on page 246) are completely consistent with those in Table 35 and are included in the set of models used for coefficient interpretation. However the only air model (Table 46 on page 247) without counterintuitive estimates for egress distance is model

A 1A-2, which is very similar to A 1-2 in Table 34. The main reason the other models, including their global statistics (Table 47 on page 248), were presented is to demonstrate that airport pair choice models similar in structure to HSR station pair choice models are inferior to base model A 1-2 in Table 34.

The airport pair choice model lacks number of transfers (3.3.13 Number of Transfers on page 108). However it does include two additional independent variables that are not found in the HSR station choice model, frequency and fare.

Terminal pair choice determinants found to be important for one mode, but not the other, are analyzed in the remainder of this section.

6.1.2 Air Frequency

As discussed in 5.3.3 Construction of Level of Service Data Matrix for HSR on page 183, the HSR station choice model does not include frequency effects. In order to obtain the number of daily frequencies, the information for over 10 000 alternative routings would have had to be looked up manually in a Japanese time table and counted by hand. Since all train categories operate several times per hour, a frequency variable would not likely have changed the estimates for station pair choice and the considerable extra effort could not be justified.

However, air frequency has significant explanatory power because it varies greatly between airport pairs. Airlines in general and especially airlines in Japan tend to concentrate their service offering in a few highly competitive markets for two reasons:

- Unlike HSR competing with air, airlines competing with HSR have to offer almost exclusively non-stop service because the time penalty and incremental cost for an additional stop is comparatively high (please see competitive advantage #4-quick stop on page 3).
- Unlike HSR airlines have to compete with each other. The S-curve (4.2.1 Graphic Derivation of the Logistics Curve on page 120) explains their tendency to "over-supply" competitive markets.

(M. M. Hansen, 1988) found that air frequency is best modeled by its logarithm. This result was used without further research.

6.1.3 Air Fares and Air Time – Two Distinct Effects

Due to asymmetric supply in different markets the correlation between air fares and distance tends to be much weaker than between rail fares and distance. Highly competitive markets generally have lower fares than less competitive ones. As a case in point of the strong correlation (almost a perfect 1) between rail fare and distance consider that the majority of European railways priced their services exclusively on a per km basis. Only with the advent of High Speed Rail did the tariff structure for most long distance connections change to a city pair based system.

Additionally, air time and distance are not as highly correlated as rail linehaul time and distance. In relatively short-haul domestic markets take-off and landing, ascent and descent, and taxiing to and from the gate take up a much greater portion of total air time than cruising. The data used in this analysis shows a range of 60 min for air time versus almost 6 hours for rail linehaul time (Table 36).

Table 36 - Linehaul Time Distribution for Chosen Airport and HSR Station Pair

	Air Time min	Nozomi Time min
Minimum	**35**	**6**
5th	50	37
25th	60	65
Median	70	101
Mean	*70*	*113*
75th	75	152
95th	90	212
Maximum	**95**	**363**

Air fares and air time are also not strongly correlated because each one is driven by different factors. That enables us to distinguish between the two effects in airport pair choice models.

Section 3.3.5 Fare Levels on page 92 already mentioned that we do not have fares for all of the 1 260 Shinkansen station pairs. It also discussed why HSR fare levels are often not as important as one might initially suspect. The fare we estimated using a linear regression on available data was highly collinear with time and distance. A few attempts to use the estimated fare gave counterintuitive results.

The lack of accurate fare data can be a significant shortcoming of an air/rail mode choice model if a HSR operator uses pricing to compensate for longer linehaul times. The fare offered between Tokyo and Fukuoka (HSR market share 12%) might be much lower on a per km basis than between Tokyo and Osaka (HSR market share 86%). However, this was not apparent from the available fare data used for the regression, and we do not have any other indication that this might be the case in Japan. The reason is the significant capacity constraint the Shinkansen system operates under.

Section 3.4.2 Japan's Secret: Few Trade-offs on page 113 explains how Japanese railway companies can offer high speed, high frequency, and high accessibility simultaneously, but only by compromising capacity. Almost all avenues to increase capacity have been exhausted and the next step is to build the Chuo maglev line between Tokyo and Osaka.

6.1.4 HSR Transfers

Transfers only being a terminal pair choice determinant for HSR is the corollary to frequency being a significant explanatory variable only for airport pair choice models. Since airlines cannot offer service requiring a transfer in markets where they compete with HSR because it would destroy their linehaul time advantage, they have to focus on a few city pairs that allow them to fly frequent non-stops. The concentration on major airport pairs leads to a clear distinction between different markets in terms of air frequency and makes this an important terminal pair choice determinant for air.

HSR operators are able to offer service between many different cities, but can only do so by requiring transfers. In order for rail travel to and from smaller and medium size cities to be fast enough, the major portion of the trip has to be spent aboard super fast express trains making relatively few stops. When traveling to smaller destinations this is generally not possible without a transfer in a major city. If potential rail travelers were forced to take local trains, like the Kodama, for the whole length of the trip, many might choose air even if it necessitated driving to and from a distant airport.

The duality between air frequency and HSR transfers is directly related to competitive advantage #4-quick stop on page 3. Quick stops make quick cross-platform transfers possible. Two high speed trains can exchange passengers in five minutes or less. These transfers serve as the great “equalizer” of frequency. If all three train categories, express, semi-fast, and local are operated several times per hour, 16 hours per day, as is the case in Japan, any city combination on the high speed line is automatically served several times per hour. Thus the important terminal choice determinant for HSR is number of transfers, not frequency.

6.2 *Relative Importance of Terminal Pair Choice Determinants*

Due to the interdependencies in our models one has to make assumptions with regard to the attributes of the trip and the trip maker in order to analyze the relative impact of various independent variables on terminal pair choice. The following subsection looks at the distributions of important explanatory variables to develop the concept a "typical" traveler which will be used for the remainder of this chapter.

6.2.1 The Typical Japanese Intercity Traveler in 1995

Studying the distribution of discrete variables in Table 37 we find a 79 : 21 relationship for both trip purpose and sex. Approximately 80% of all travelers are male, and approximately 80% of HSR passengers travel on company business. Not many young or old people use either air or HSR. Almost three quarters of all trip makers fall into the middle of three equally long age ranges. Air and rail statistics are fairly similar; trip duration, however, reflects the difference in average travel distance for each mode. Very noticeable is the dominance of urban zones, not only as origin and destination, but also as place of residence. Note that of the 1 574 five-digit zones in our 24 prefecture analysis area, only 544 or 35% are urban zones.

The margins of Table 38 reproduce the 79% : 21% relationship for both trip purpose and sex which we are already familiar with from Table 37. Table 37 also shows that business travelers make up a higher percentage of HSR passengers than of air passengers. Air leads in vacation trips. This difference is lost in the combined air and rail data of Table 38 because HSR travelers outnumber air travelers by a ratio of almost 6:1.

Table 37 - Distributions of Important Discrete Explanatory Variables

Variable	Category	Air	HSR
Purpose	Company Business	76%	79%
	Personal Business	9%	11%
	Vacation	15%	10%
		100%	100%
Duration	1 Day	17%	45%
	Overnight	83%	55%
		100%	100%
Sex	Female	21%	21%
	Male	79%	79%
		100%	100%
Age	[0,30)	15%	17%
	[30,60)	73%	72%
	[60,100)	12%	11%
		100%	100%
Direction	Leaving Home	54%	53%
	Coming back Home	46%	47%
		100%	100%
Trip Origin Zone	Rural	10%	7%
	Urban	90%	93%
		100%	100%
Trip Destination Zone	Rural	8%	6%
	Urban	92%	94%
		100%	100%
Home Zone	Rural	10%	8%
	Urban	90%	92%
		100%	100%

Table 38 - Relationship between Sex and Trip Purpose

		Vacation or Personal Business		Company Business		Total
Female	*Percent*	13%		7%		**21%**
	Row Percent	65%		35%		100%
	Column Percent		62%		9%	
Male	*Percent*	8%		71%		**79%**
	Row Percent	10%		90%		100%
	Column Percent		38%		91%	
Total		**21%**	100%	**79%**	100%	100%

Of all females (Row Percent) two thirds travel on personal business or are on vacation. Only one third of all females travel on company business. This contrasts to males of which over 90% travel on company business. We would expect males to also be required to travel for personal business, but they are more likely than females to be able to combine it with a company business trip. Most surprisingly, of all people traveling on company business (Column Percent) 91% were male and only 9% female in our analysis area during the Fall of 1995.

Our choice set for air travelers consists of 82 feasible airport pairs. Out of these 82 pairs four account for one third of all the chosen airport pairs (Table 39).

With HSR, the top 4 station pairs make up one quarter of all the possible 1 260 station pair choices (Table 40). In both instances the four pairs are formed by only three terminals. For HSR it is Tokyo Central, Shin Osaka, and Nagoya, and for air Tokyo Haneda, Osaka Itami, and Fukuoka Itazuke.

Table 39 - The Four Most Important Airport Pairs

Origin Airport	Destination Airport	Percent of All Air Travelers within Analysis Area
Fukuoka Itazuke	Tokyo Haneda	13%
Tokyo Haneda	Fukuoka Itazuke	12%
Osaka Itami	Tokyo Haneda	5%
Tokyo Haneda	Osaka Itami	4%
Total		**34%**

Table 40 - The Four Most Important High Speed Rail Station Pairs

Origin HSR Station	Destination HSR Station	Percent of All HSR Travelers within Analysis Area
Shin Osaka	Tokyo Central	7%
Tokyo Central	Shin Osaka	6%
Nagoya	Tokyo Central	5%
Tokyo Central	Nagoya	5%
Total		**24%**

The focus of this dissertation is access and egress related, so it is helpful to be familiar with the distance distribution to and from the chosen airport or rail station (Table 41). The rail distributions have fatter tails, or more very short and more very long access journeys compared to air. The relative high number of short access trips is easy to explain with station locations in the center of urban activity. Very long access trips are most likely made by rail, given this mode's dominance of intercity travel in Japan. High speed rail stations are almost by definition easier to reach by rail than airports, which makes the trip continuation on HSR very sensible. Except for the tails, the distributions of access and egress and of air and rail are very similar. It is worthwhile to remember that the median access/egress distance for air and HSR are only 17 km and 15 km respectively.

Table 41 - Distribution of Access/Egress Distance to/from Chosen Airport or Station

Percentile	Airport Access Distance (km)	Airport Egress Distance (km)	HSR Station Access Distance (km)	HSR Station Egress Distance (km)
Minimum	2	2	0.4	0.4
5th	4	4	1	1
25th	10	9	6	4
Median	**18**	**17**	**15**	**15**
Mean	*27*	*25*	*33*	*36*
75th	35	32	33	36
95th	79	68	138	145
99th	139	129	243	241
Maximum	595	538	889	902

A difference in tail behavior is also discernible from the zone size distribution (Table 42). Larger zones tend to be located in more rural areas, farther away from densely populated urban districts. The 95th and 99th percentile indicate a fatter tail for rail travel, which is also reflected in the larger mean.

Table 42 - Zone Size Distribution

Percentile	Air Origin Zone Size (km^2)	Air Destination Zone Size (km^2)	HSR Origin Zone Size (km^2)	HSR Destination Zone Size (km^2)
Minimum	3	2	3	3
5th	10	10	10	10
25th	20	19	22	20
Median	**47**	**39**	**50**	**53**
Mean	*95*	*95*	*108*	*114*
75th	128	126	112	124
95th	290	290	364	364
99th	361	361	1146	1146
Maximum	1146	1146	1146	1146

We are now able to describe the "typical" Japanese intercity traveler within our analysis area of the Tokaido and Sanyo corridors in the year 1995:

1. He is male.
2. He is traveling on company business.
3. He will stay at least one night away from home.
4. He is in his 30's, 40's, or 50's.
5. His zone of residence, trip origin, and trip destinations are all urban.
6. His median access/egress distance to or from the airport is 17 km with an interquartile range from 10 km to 35 km.
7. His median access/egress distance to or from the HSR station is 15 km with an interquartile range of 5 km to 35 km.
8. His trip origin is likely to be near his residence.
9. If traveling by air he is flying from Fukuoka Itazuke to Tokyo Haneda.
10. If using HSR he is riding from Shin Osaka to Tokyo Central.

6.2.2 Utility Impact Analysis of Terminal Pair Choice Determinants

Besides describing a typical intercity traveler, median values for the other socio-economic and land use variables for either mode were determined. This allowed us to compute utility for three different scenarios:

- Both access and egress distance are at the 25th percentile.
- Both access and egress distance are at the median.
- Both access and egress distance are at the 75th percentile.

The disadvantage of using a typical traveler is that “atypical” intercity travelers, like vacationers or HSR passengers who transfer, are not represented. The results of the utility impact analysis are easily visualized with the help of Figure 22 and Figure 23. Table 43 and Table 44 focus on important findings that would be difficult to derive graphically.

Probably the biggest surprise was the relative insignificance of crossterms. This is worth noting. The literature review made it apparent that many intercity mode choice studies, including recent ones, do not take access or egress related variables into account. One research goal is to offer a suggestion on how to remedy this shortcoming in future work without a disproportionate increase in expenditures. Figure 22 and Figure 23 would suggest that if access or egress were not a major research focus, the efforts spent on collecting land use data like origin or destination zone density could probably not be justified. Figure 23 shows a stack of seven crossterms at the top of the 75th percentile column. The last two, which are the interaction of access and egress distance with linehaul time, account for 59% of the crossterm contribution to positive utility.

On the other side of the spectrum, the extent to which fixed effects impact utility of HSR travel was unexpected. Fixed effects amount to 44% of total absolute utility in the 25th percentile segment of HSR (Table 43). Table 43 highlights the difference between air and HSR utility in terms of fixed effects. The reader may caution that the large contribution of fixed effects shown in our analysis may be somewhat accidental since the chosen station pair, Shin Osaka – Tokyo Central, contains the only two stations explicitly included in our model. But if we chose the fourth most important station pair, Tokyo – Nagoya, we would only lose Origin Station = Shin Osaka, while gaining additional utility from Tokyo Central as the origin station. Most important is the addition to positive utility by a station being one of the select few where every Nozomi express train stops. Fixed effects are covered in greater detail later in this chapter.

Also remarkable in Figure 22 and Figure 23 is not only the large influence of access and egress distance on total utility, but the difference in composition between air and rail. The linear element of access and egress distance is much smaller for HSR than for air. The linear part of HSR access distance is barely visible at the 25th percentile. The logarithmic coefficient for HSR access distance is 65 times larger in magnitude than the linear component. For air the same ratio is only 14 : 1. Table 43 and Table 44 show that, within the interquartile range, the total access and egress distance related portion of utility is very similar for air and HSR. The dissimilar mathematical compositions of access distance disutility make a difference for longer distances. For larger numbers the linear dominates the logarithmic term and a steeper linear slope means a higher marginal disutility. This again supports the notion, first mentioned earlier in this section, that rail is the preferred mode for very large access and egress distances.

Figure 22 – Utility Impact Analysis of Airport Pair Choice Determinants

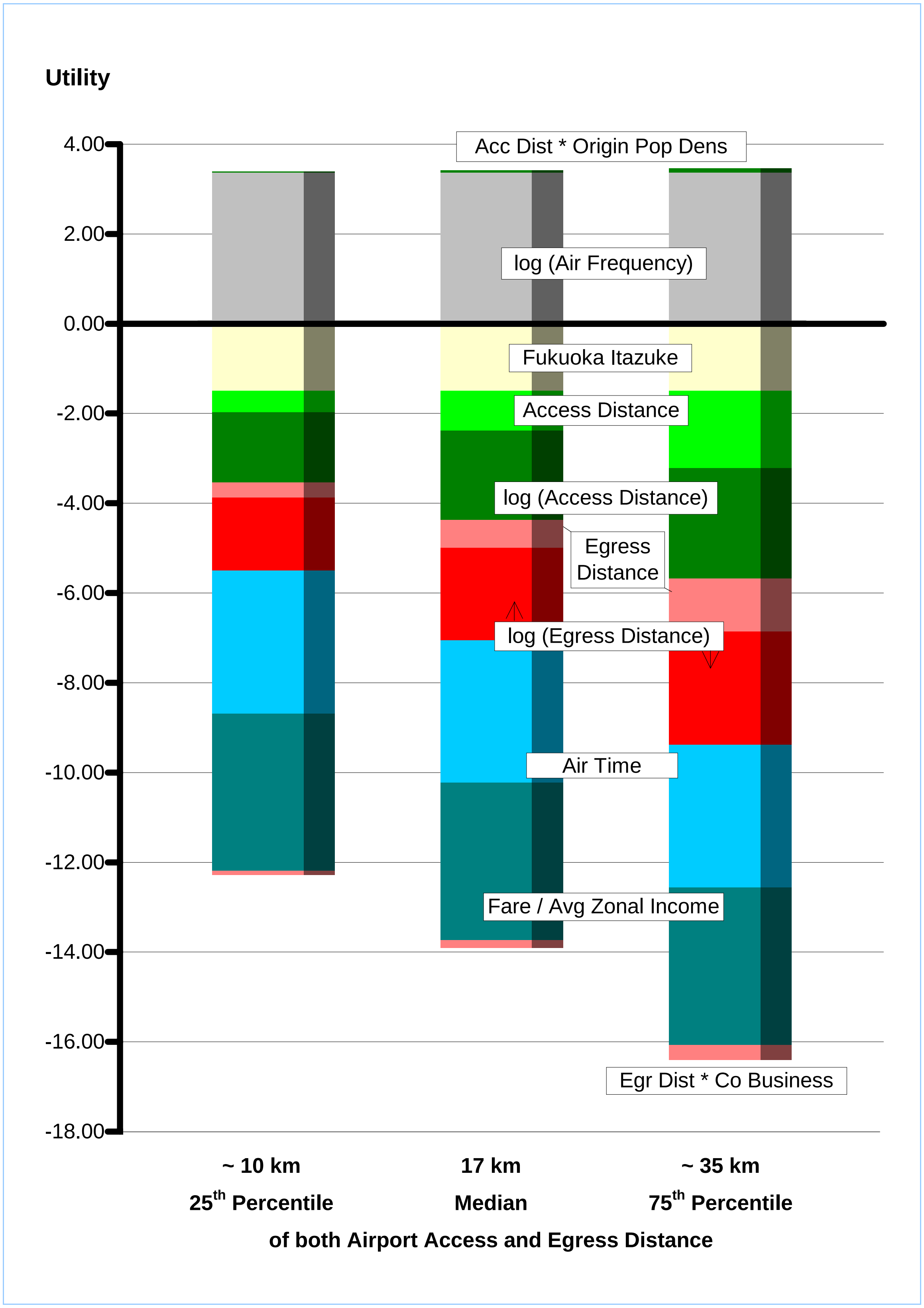


Figure 23 – Utility Impact Analysis of HSR Station Pair Choice Determinants

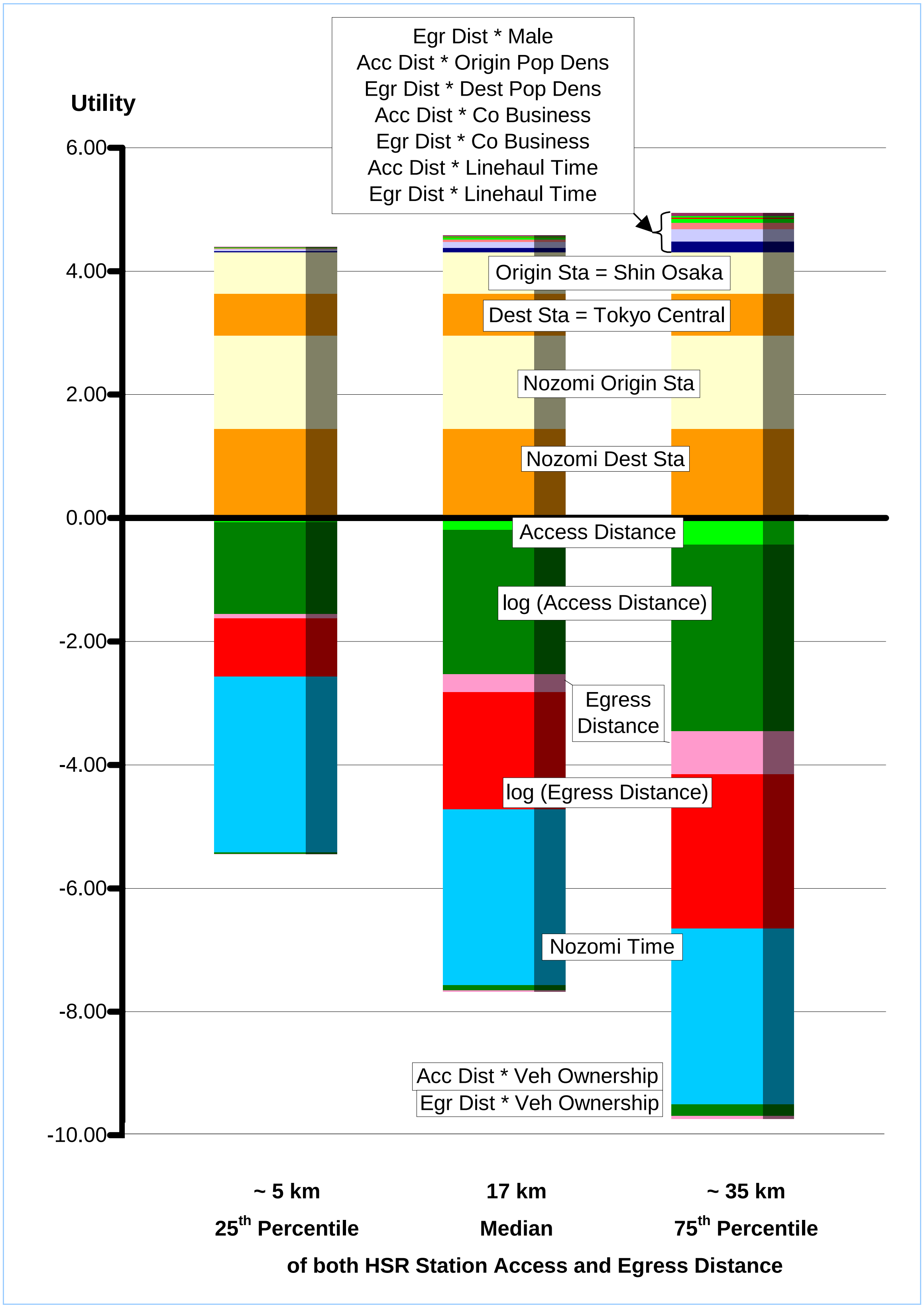


Table 43 - Typical Air Traveler's Utility as a Function of Feeder Distance

	Utility of Airport Access / Egress Dist at		
	25th Percentile	Median	75th Percentile
Total Utility	-8.89	-10.49	-12.94
Total Positive Utility	3.39	3.42	3.47
Total Negative Utility	-12.28	-13.91	-16.40
Total Absolute Utility	15.68	17.33	19.87
Access Distance Related Utility	-2.01	-2.83	-4.08
Egress Distance Related Utility	-2.06	-2.85	-4.04
Total **Access & Egress** Distance Related Utility	-4.07	-5.68	-8.12
Percent of Total Absolute Utility	26%	33%	41%
Utility of **Fixed Effects**	-1.50	-1.50	-1.50
Percent of Total Absolute Utility	10%	9%	8%

Table 44 - Typical HSR Traveler's Utility as a Function of Feeder Distance

	Utility of HSR Station Access / Egress Dist at		
	25th Percentile	Median	75th Percentile
Total Utility	-1.06	-3.09	-4.79
Total Positive Utility	4.39	4.58	4.94
Total Negative Utility	-5.45	-7.67	-9.73
Total Absolute Utility	9.84	12.25	14.68
Access Distance Related Utility	-1.54	-2.48	-3.35
Egress Distance Related Utility	-0.98	-2.07	-2.90
Total **Access & Egress** Distance Related Utility	-2.52	-4.55	-6.25
Percent of Total Absolute Utility	26%	37%	43%
Utility of **Fixed Effects**	4.30	4.30	4.30
Percent of Total Absolute Utility	44%	35%	29%

6.3 Functional Form of Feeder Distance

6.3.1 Log plus Linear Function Composition

Common sense and experience let us to believe that reducing access distance from 2 km to 1 km increases the utility of that choice far more than reducing the distance from 102 km to 101 km. We can presume that a solely linear specification would not represent the data faithfully. Linear functions meet the basic requirement of reproducing a negative first derivative (additional access distance reduces utility more). But we also need a positive second derivative to model decreasing marginal disutility of access distance at least between the minimum and the 99th percentile. For our data the two requirements would have to be met for the range of 0.4 km to about 250 km (Table 41 on page 224).

Estimation results for various specifications of HSR access and egress distance are graphically illustrated in Figure 24 - Figure 27.

Linear, quadratic, cubic, and log plus linear models were estimated initially. The two requirements were met for all functional forms except the linear (by design $f'' = 0$). The quadratic, cubic, and log plus linear functions are falling at a decreasing rate ($f'' > 0$) within the range of common values (0.4 km – 250 km).

Polynomials can give counterintuitive results for some data points. It is important to check their first and second derivatives, at least visually. Both the vertex of the quadratic ($f' = 0$ at 621 km) and the inflection point of the cubic function ($f'' = 0$ at 430 km) were beyond the 99th percentile, but below the maximum (Figure 24). For the log plus linear specification the two requirements are met for any positive value greater than or equal to1, i.e. for any distance of 1 km and above.

Figure 24 - Different Functional Forms To Model Access Distance to HSR Station

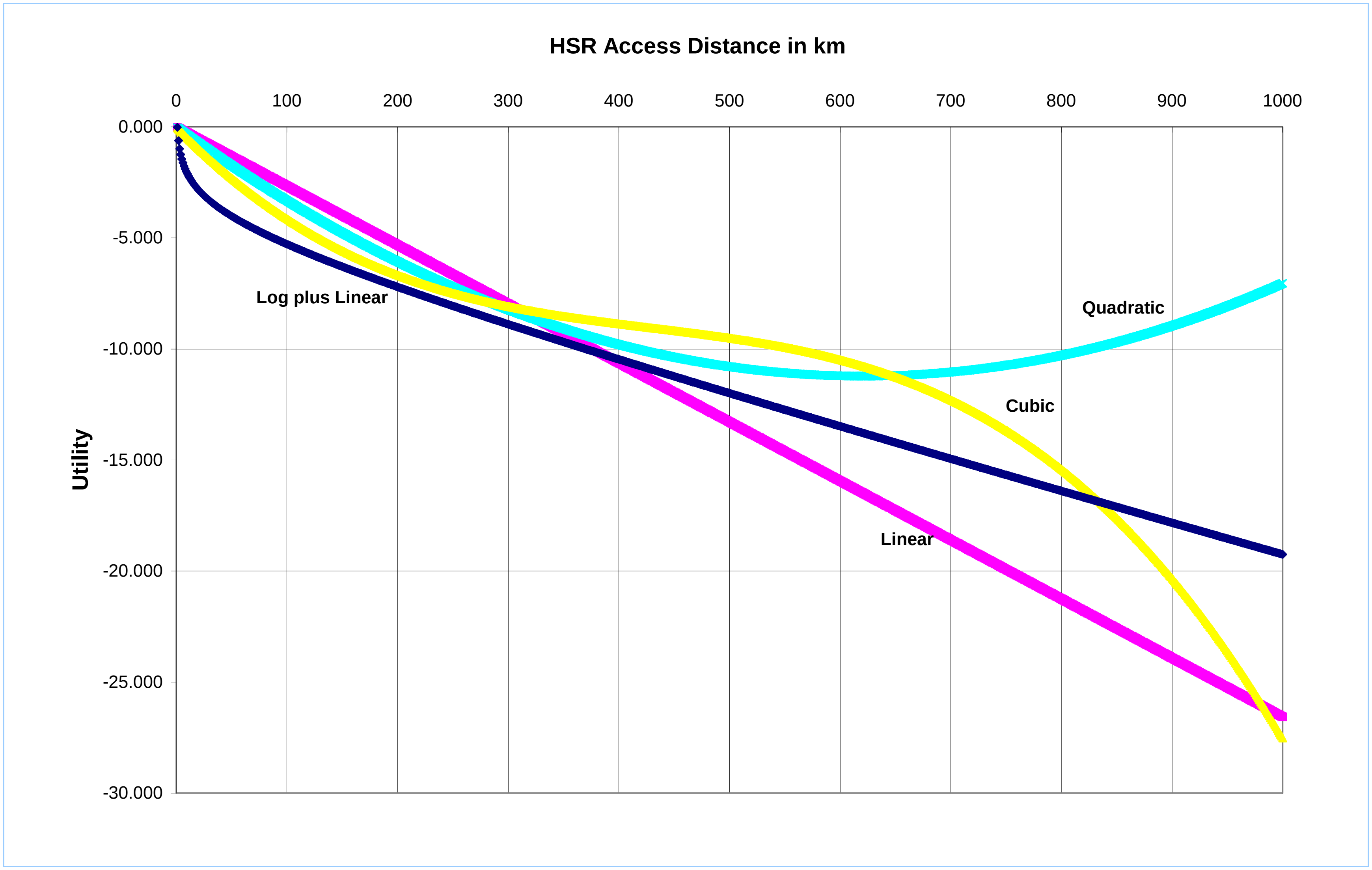


Figure 25 - Comparison of Different Specifications for HSR Access Distance

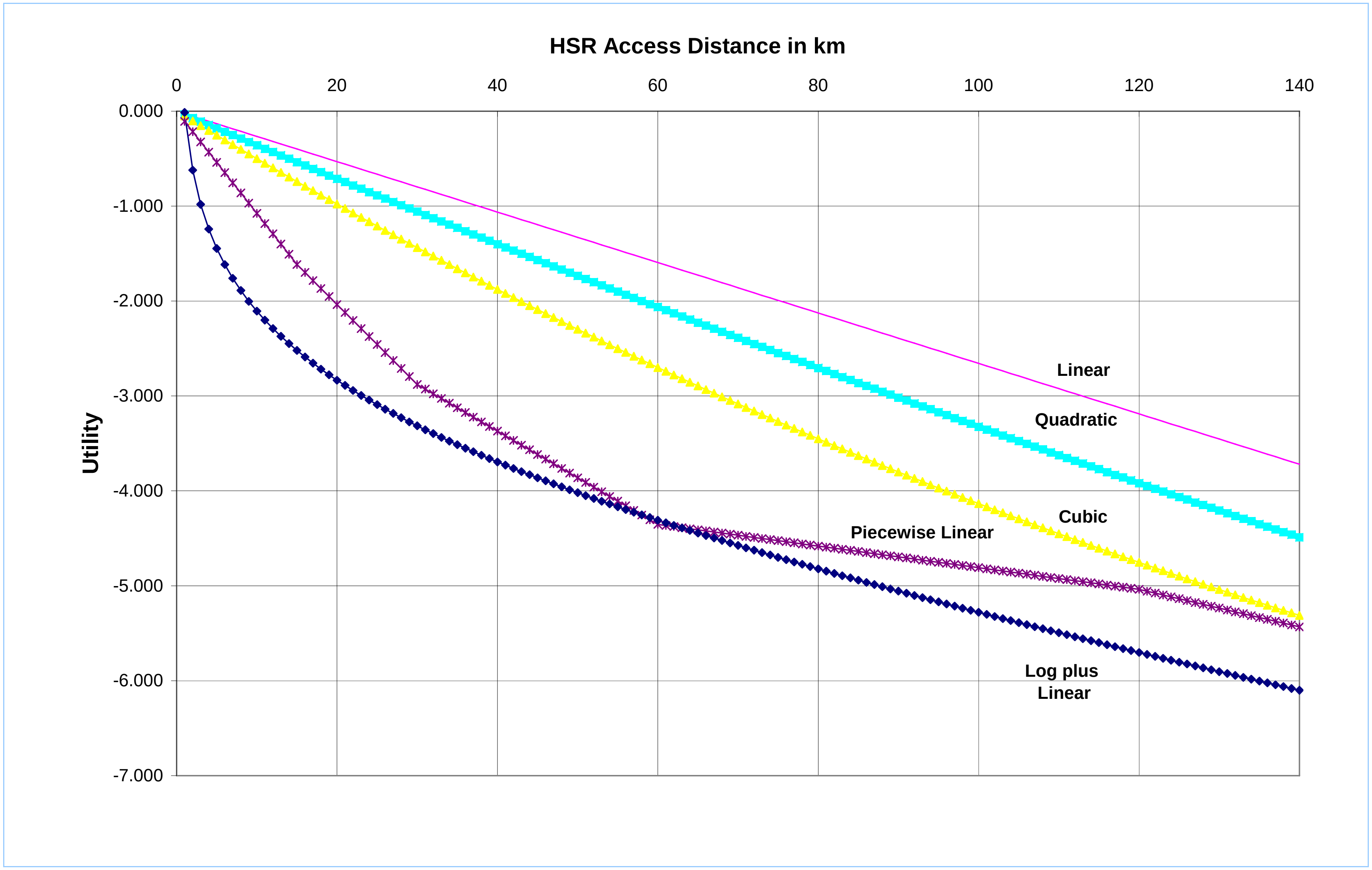


Figure 26 - Comparison of Different Specifications for HSR Egress Distance

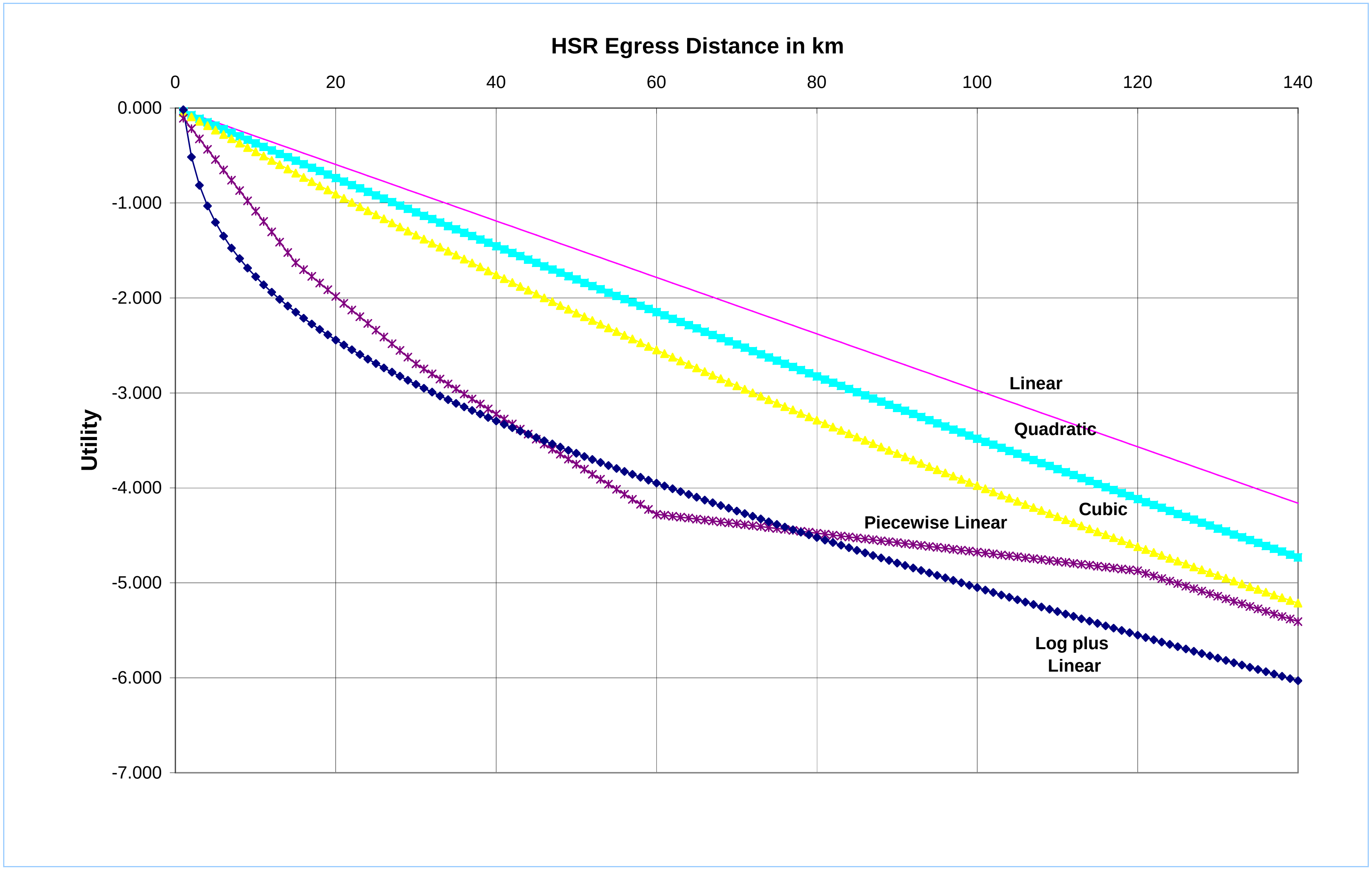
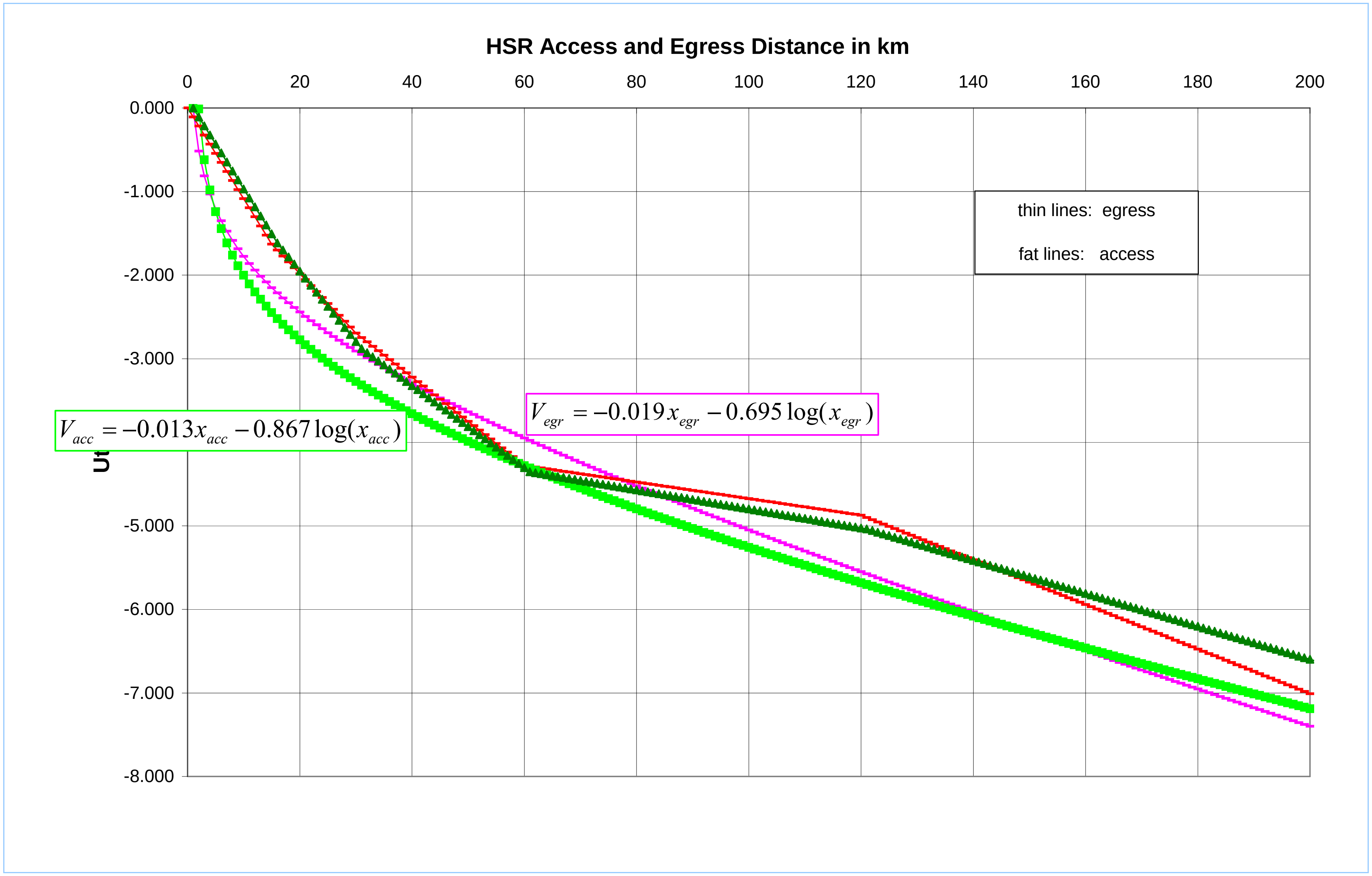


Figure 27 - Comparison of HSR Access and Egress Distance



In order to rank the four specifications, two types of statistical tests were employed. Likelihood ratio tests, which evaluate nested hypotheses, confirmed basic expectations: any model was better than the linear one, and the cubic was better than the quadratic specification. In order to test non-nested hypotheses we compared the adjusted likelihood ratio indices $\overline{\rho}^2$. Ben-Akiva and Swait (1984) using a test developed by Horowitz (1983) derived a formal statistical test to appraise the difference of $\overline{\rho}^2$ between two models. Under the null hypothesis that model 1 is the true model, the following inequality holds asymptotically (M. E. Ben-Akiva & Lerman, 1985, pp. 171-172):

$$P\left(\overline{\rho}_2^2 - \overline{\rho}_1^2 > z\right) \leq \Phi\left(-\sqrt{-2z\,\ell\ell(0) + \left(K_2 - K_1\right)}\right) \tag{100}$$

where

$\ell\ell(0)$ is the value of the log likelihood function when all parameters are zero.

$\ell\ell(\hat{\beta})$ is the value of the log likelihood function at its maximum.

$\overline{\rho}_l^2 = 1 - \frac{\ell\ell(\hat{\beta}) - \mathrm{df}}{\ell\ell(0)}$ is the adjusted likelihood ratio index for model $l = 1, 2,$.

K_l is the number of parameters in model l.

Φ is the standard normal cumulative distribution function.

The $\overline{\rho}^2$ of the log plus linear model was 0.0123 greater than the $\overline{\rho}^2$ of the cubic model. The probability that this difference is 0.0123 or greater, given that the cubic model is the correct one, is bound by the right hand side of equation (100). That probability turned out to be less than 0.001. The log plus linear specification was the clear winner. It also uses two fewer degrees of freedom than the cubic model and does not require checking the first and second derivatives, because it gives intuitive results for the full range of distances.

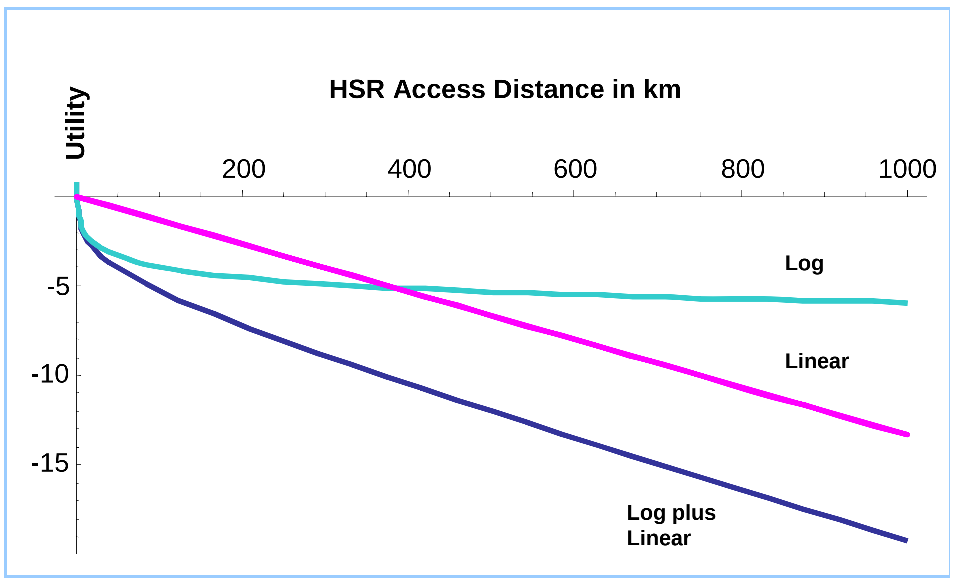
Piecewise linear models were estimated to obtain an even more accurate reading of the non-linearly decreasing utility. Four breakpoints were chosen at 15 km (the median), 30 km, 60 km, and 120 km resulting in five ranges. Results are shown in Figure 25 for HSR access and Figure 26 for HSR egress. The piecewise linear proved to be better than the log plus linear specification at the same significance level found when comparing the log plus linear to the cubic model. The probability that the higher $\bar{\rho}^2$ of the piecewise linear model was due to chance fluctuation was less than 0.001.

The distance range in Figure 25 and Figure 26, 1 km – 140 km, covers about 95% of the observed data. There is surprisingly little difference between the linear, quadratic, and cubic specification. Only the logarithmic line reflects the pattern observed through piecewise linear approximation. For almost half of the datapoints, between 20 km and 140 km, the slope of all functional forms is very similar, resulting in approximately the same marginal disutility. But for the first half of the datapoints, the linear, quadratic, and cubic forms depict the underlying data inaccurately and vastly underestimate the disutility of access distance. The problem is that if those functions tried to match the slope of the piecewise linear or logarithmic curves for the first half of the data points, the negative utility for long access distances would be so high, that the model would predict they'd never be chosen. And, as with all regressions based on minimizing least squares, outliers have a disproportionate influence on the final estimate.

Logarithms make small numbers big and big numbers small. This adds to their ability to pick up gradually decreasing slopes in a relatively small interval at the beginning of the data series. It also provides a counterbalance to the impact of large outliers in Ordinary Least Squares (OLS) regressions and the resulting difficulty of many functional forms to reflect important patterns in intervals of small numbers, as we have seen in this example.

The log plus linear function composition allows the linear element to minimize squared deviations for the longer distances while the logarithmic component can match the data in the small but densely populated segment of short access distances (Figure 28).

Figure 28 - Graphical Addition of Logarithmic and Linear Component for HSR Access Distance



The range of coefficients for the logarithmic term in the four models shown in Table 34 and Table 35 is –0.7 to –0.9. The rail model based on 10 alternatives, R 10-2, resulted in a coefficient smaller than –1. Since this range is close to –1 it is helpful to recall the first and second derivatives of the function $-\log(x)$.

$$y = -\log(x) \tag{101}$$

$$y' = -\frac{1}{x} \tag{102}$$

$$y'' = \frac{1}{x^2} \tag{103}$$

At 10 km the first derivative is –0.1 and each additional km noticeably reduces utility. But at 100 km the first derivative is –0.01 and the second derivative 0.0001. Both first and second derivatives approach 0 very fast and the logarithmic function behaves more and more like a constant function. This is desirable when modeling air frequency. There is not much gained from an increase in daily frequencies from 30 to 60. But utility continues to decrease with rising feeder distance. Therefore, the linear term has been highly significant in every estimation.

The flexibility of the log plus linear specification is illustrated by the difference in estimation between HSR access and egress distance (Figure 27). The logarithmic coefficient is larger for access, but the linear parameter is greater for egress distance. Because of the larger logarithmic element, the utility of access distance falls faster at the beginning of the range where the logarithmic component dominates the linear. As distance increases the linear term becomes more important and the access curve falls more slowly than the egress line.

Our estimation finds that for smaller distances access has a higher influence on utility, while for longer distances egress becomes more onerous. The log plus linear function lines cross at 157 km. The access and egress lines of the piecewise linear specification intersect at 144 km. Assuming that the piecewise linear estimation is the most accurate, we can conclude that the log plus linear function composition replicates this crossover effect very faithfully.

Only the cubic specification comes close to this performance. It estimates a crossover point of 175 km. Both quadratic and linear functions incorrectly portray egress distance to be more onerous for the full range of common values.

It might appear from Figure 29 and Figure 30 that the log plus linear function composition overestimates the disutility of access distance. However, since the number of range combinations for a piecewise linear specification is infinite, and the results are based on a single dataset and one random sample of alternatives, there is not enough evidence to reach that conclusion.

On the other hand, the facts that piecewise linear and log plus linear specifications follow the same general pattern, and that this pattern is the one expected from experience and common sense, in addition to the high significance levels of the statistical tests, lead us to conclude that polynomial expansions up to degree three should be excluded from further consideration for modeling high speed rail access and egress distance in our data.

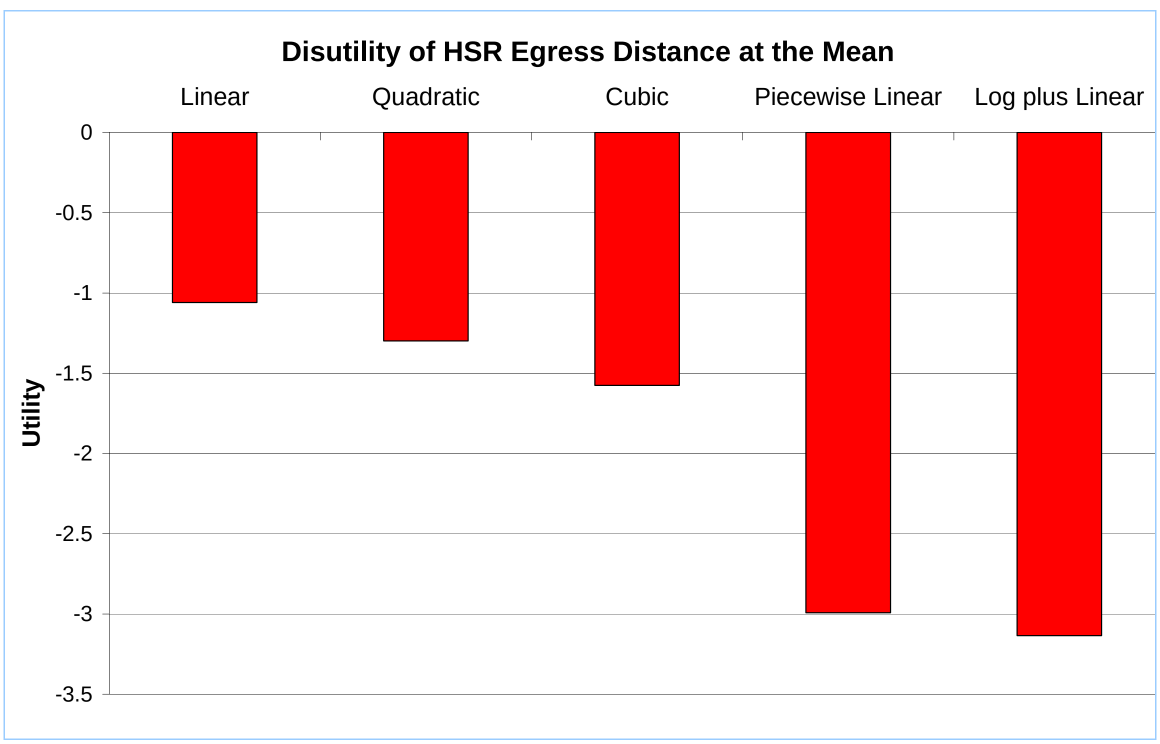
The five models discussed so far in this section were estimated using the same sample of 299 randomly selected station pair alternatives plus the selected station pair. Any differences in the parameter estimates are only due to specification differences.

There was no a priori expectation that airport pair choice models would deviate noticeably from the results obtained for HSR. We can see from Table 41 on page 224 that, except for the tails, the distributions of air and rail are very similar.

Somewhat surprisingly, for air access and egress distance the cubic specification fitted the data more closely than the log plus linear function composition. The next subsection may shed some light on why this happened.

Figure 29 - Estimation Variations at Median Caused by Different Functional Forms

Figure 30 - Estimation Variations at Mean Caused by Different Functional Forms



6.3.2 Piecewise Linear Specification and Threshold Effects

Only focusing at first on model R 10A in Table 45 on page 246, we notice a steady drop in the marginal disutility of access distance. This is what we would expect. However as the number of choices increases the coefficients of the 60 – 120 km range become more and more distinct. Travelers appear to be fairly indifferent to increases in access distance within this range.

This effect is more pronounced with egress distance. Even with only 10 choices the 60 – 120 km range stands out. As the number of alternatives increases the parameters of that range diverge more and more. We may suspect that the piecewise linear specification is detecting a threshold effect. Numerically, the symptom is a "trough," an unusually small number in a series of parameter estimates for a piecewise linear function. Geometrically, a threshold effect may reveal itself as a "ledge." Figure 25 graphs the piecewise linear function of model R 300A for access distance, Figure 26 for egress, and Figure 27 for both access and egress.

The asymptotic t-test confirms that the change in slope at the 120 km breakpoint is significant in every instance. The test statistic for the null hypothesis that the slopes are equal is given by (M. E. Ben-Akiva & Lerman, 1985, p. 161):

$$\frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\operatorname{var}\left(\hat{\beta}_1 - \hat{\beta}_2\right)}} = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\operatorname{var}\left(\hat{\beta}_1\right) + \operatorname{var}\left(\hat{\beta}_2\right) - 2\operatorname{cov}\left(\hat{\beta}_1, \hat{\beta}_2\right)}} \tag{104}$$

The pattern in the test statistics elucidates that as the number of alternatives in the sample increases the effect comes more sharply into focus.

Estimation results of airport pair choice models (Table 46 and Table 47) show a comparable pattern for a different range, 15 – 30 km. We notice a "trough" for access distance, however this trough is small (and insignificant) for model A 1A. Again, the effect is much stronger with egress distance. It is so extreme, that, except for model A 1A, the coefficients for airport egress distance between 15 km and 30 km are positive. The *t*-statistics for the null hypothesis that the slopes are equal at the 30 km breakpoint in model A 1A are 1.4 for access and 7.9 for egress.

15 – 30 km is still within the range where mostly the logarithmic term determines the precipitousness of the decline in utility. It would be very hard for a logarithmic function to reproduce a "flat spot" in the utility curve this close to the origin. Log plus linear function compositions are not optimal when threshold effects can be detected near the beginning of the data series.

Experience in the San Francisco Bay Area may shed some light on these estimation results. Travelers in Northern Alameda County are usually fairly indifferent between Oakland and San Francisco International Airport. The former has a shorter access time (half hour, versus one hour to San Francisco) but less nonstop service. Airport choice between Oakland and San Francisco seems to be mostly affected by the availability of nonstop service and fare levels. On the other hand, San Jose International Airport is only seldomly considered. It is further away and has less service than San Francisco International. It is seen to be almost in "Southern California." The additional half hour access time (One and one half hours to San Jose versus one hour to San Francisco) would be perceived as very onerous by most air travelers from Northern Alameda County.

Japanese intercity rail travelers may find 100 km access distance, or about one hour access time at typical speeds of conventional trains, "reasonable." However, they consider access times above one hour "unreasonable" and start looking for other options. What access time is concerned, Japanese intercity rail travelers appear to have about the same threshold of pain as Northern California air travelers.

The reader might find it interesting that a 1992 study (our dataset is from 1995) points to 25 km/h as the average speed in urban areas in Japan (Schafer, 2000). This places the 15 to 30 km range for airport access distance at the outside edge of the one hour window for access time.

An important "universal" law, which emerged after over 40 years of HSR operation in Japan and almost 25 years in Europe, was discussed earlier (3.3.4 Linehaul Time, page 90). For high speed rail to achieve a dominant share of the combined air/rail market, HSR's linehaul time must be kept below three (3) hours. If HSR's linehaul time falls inside of this three hour window, HSR is considered an alternative to air by most travelers. If the linehaul time falls outside of this range, it isn't. The desire to employ this three hour time window in longer and longer corridors drives the technological development of HSR in Europe and Japan (Jackson, 2004b).

These examples illustrate the concept of the threshold of pain. It is akin to the medical phenomenon of *therapeutic dose*. Many drugs, when administered below a minimum dosage, which varies from individual to individual, have almost no effect. Human responses are often best modeled by step functions. Power series expansions and logarithmic specifications obscure important thresholds resulting from human perceptions. When studying mode choice determinants in detail the researcher has no alternative to experimenting with piecewise linear functions.

Travelers appear to be more sensitive to "pain" during the egress portion of the trip. Carefully studying Table 45 one notices that for all HSR models the egress coefficient of the 60 – 120 km range is smaller than the corresponding access coefficient. However, for the range both immediately preceding and following, the egress parameters are markedly larger than the equivalent access coefficients. The evidence is not as convincing for the air models because of incorrect signs, but we may suspect the same phenomenon.

One might wonder whether the larger drop followed by a steeper rise in the magnitude of the egress parameters indicates a much steeper threshold of pain for the egress portion of the trip, while the increase in pain during access may be more gradual. We might suspect that an egress distance greater than the (arbitrarily chosen) breakpoint of 120 km becomes very onerous.

This finding is consistent with the earlier one: for smaller distances egress seems less important than access distance. But for distances greater than approximately 150 km, egress distance is more onerous than access distance.

Table 45 - Piecewise Linear Estimation for HSR Access Distance Using Various Size Choice Sets

	R 10A-2 **10 Choices** *Parameter Estimate*	**R 50A-2** **50 Choices** *Parameter Estimate*	**R 100A-2** **100 Choices** *Parameter Estimate*	**R 300A-2** **300 Choices** *Parameter Estimate*
Origin Sta = Tokyo Central	0.4850	0.6380	0.6695	0.6600
Dest Sta = Tokyo Central	0.5589	0.5357	0.6330	0.6014
Origin Sta = Shin Osaka	0.3162	0.4214	0.5244	0.4756
Dest Sta = Shin Osaka	0.2967	0.5013	0.5235	0.5581
Nozomi Origin Sta	1.5811	1.6098	1.5654	1.5523
Nozomi Dest Sta	1.5433	1.4915	1.4851	1.4503
Access Distance ≤15 km	-0.0998	-0.0977	-0.1043	-0.1077
15 km < Acc Dist ≤30 km	-0.0856	-0.0909	-0.0903	-0.0843
30 km < Acc Dist ≤60 km	-0.0519	-0.0493	-0.0514	-0.0491
60 km < Acc Dist ≤120 km	-0.0170	-0.0159	-0.0146	-0.0114
Access Distance > 120 km	-0.0139	-0.0187	-0.0207	-0.0198
t-stat (preceding 2 coeff eq)	5.0	11.1	13.8	14.9
Acc Dist * Co Business	*0.0014*	*0.0016*	0.0019	0.0022
Acc Dist * Veh Ownership	-0.0000440	-0.0000374	-0.0000333	-0.0000651
Acc Dist * Linehaul Time	0.0000256	0.0000356	0.0000369	0.0000412
Acc Dist * Origin Pop Dens	0.0000003	0.0000004	0.0000004	0.0000004
Acc Dist * Vacation	*0.0031*	0.0041	0.0041	0.0043
Acc Dist * (Age < 30)	*-0.0011*	*0.0000*	*0.0003*	*0.0006*
Egress Distance ≤15 km	-0.0999	-0.0891	-0.0976	-0.1087
15 km < Egr Dist ≤30 km	-0.0725	-0.0783	-0.0707	-0.0710
30 km < Egr Dist ≤60 km	-0.0560	-0.0564	-0.0556	-0.0528
60 km < Egr Dist ≤120 km	-0.0146	-0.0112	-0.0118	-0.0099
Egress Distance > 120 km	-0.0221	-0.0242	-0.0258	-0.0267
t-stat (preceding 2 coeff eq)	8.2	14.3	17.4	20.7
Egr Dist * Co Business	*0.0031*	*0.0016*	0.0021	0.0027
Egr Dist * Male	0.0031	0.0020	0.0019	0.0014
Egr Dist * Veh Ownership	-0.0000373	*-0.0000084*	*-0.0000106*	*-0.0000183*
Egr Dist * Linehaul Time	0.0000200	0.0000214	0.0000269	0.0000278
Egr Dist * Dest Pop Dens	0.0000002	0.0000003	0.0000003	0.0000003
Egr Dist * Vacation	0.0049	*0.0014*	*0.0016*	*0.0016*
Egr Dist * (Age < 30)	*0.0003*	0.0013	0.0014	0.0014
Nozomi Time	-0.0142	-0.0156	-0.0162	-0.0175
Nozomi Transfers	-0.6036	-0.4870	-0.4652	-0.4015

All parameter estimates are significant at the 5% level, except when in *italic*.

Table 46 - Piecewise Linear Estimation for Airport Access Distance Using Various Specifications

	A 2-2 Parameter Estimate	A 3-2 Parameter Estimate	A 4-2 Parameter Estimate	A 1A-2 Parameter Estimate
Access Distance ≤ 15 km	-0.1152	-0.1032	-0.1036	-0.1007
15 km < Acc Dist ≤ 30 km	-0.0398	-0.0254	-0.0251	-0.0629
30 km < Acc Dist ≤ 60 km	-0.0724	-0.0582	-0.0582	-0.0716
60 km < Acc Dist ≤ 120 km	-0.0618	-0.0460	-0.0459	-0.0666
Access Distance > 120 km	-0.0329	-0.0199	-0.0198	-0.0356
Acc Dist * Origin Pop Dens	0.0000006	0.0000003	0.0000003	0.0000007
Acc Dist * Co Business		0.0092	0.0092	
Acc Dist * Vacation	0.0129	0.0209	0.0209	0.0127
Acc Dist * Veh Ownership		-0.0000620	-0.0000622	
Acc Dist * Linehaul Time		-0.0002594	-0.0002617	
Acc Dist * (Age < 30)		*-0.0015*	*-0.0015*	
Egress Distance ≤ 15 km	-0.1661	-0.1486	-0.1482	-0.1473
15 km < Egr Dist ≤ 30 km	*0.0003*	0.0135	0.0132	-0.0200
30 km < Egr Dist ≤ 60 km	-0.0685	-0.0519	-0.0519	-0.0705
60 km < Egr Dist ≤ 120 km	-0.0564	-0.0349	-0.0350	-0.0598
Egress Distance > 120 km	-0.0165	-0.0056	-0.0058	-0.0208
Egr Dist * Dest Pop Dens		0.0000002	0.0000002	
Egr Dist * Co Business	-0.0071	0.0044	0.0044	-0.0081
Egr Dist * Vacation		0.0119	0.0119	
Egr Dist * Male		-0.0030	-0.0030	
Egr Dist * Veh Ownership		-0.0000852	-0.0000853	
Egr Dist * Linehaul Time		-0.0002768	-0.0002738	
Egr Dist * (Age < 30)		*0.0009*	*0.0009*	
log (Air Frequency)	1.2784	1.2954	1.2947	0.9666
Air Linehaul Time	-0.0572	-0.0330	-0.0335	-0.0458
Fare / Avg Zonal Income	-0.0002	-0.0002	-0.0002	-0.0002
Tokyo Narita				-0.5710
Fukuoka Itazuke			0.4695	-1.6341
Osaka Itami				0.6283
Kagoshima				-2.0181
Miyazaki				-1.6063
Nagoya				*0.0990*
Nagasaki				-2.3208
Oita				-2.4742
Kumamoto				-2.6170
Hiroshima				*-0.1211*
Osaka Kansai				1.0270
Yamaguchi Ube				-3.2647
Okayama				-0.7240
Fukuoka Kita-Kyushu				-3.6792
Tokyo Haneda			0.0316	.
Origin = Tokyo Haneda	-0.1443	-0.0464		
Destination = Tokyo Haneda	-0.0370	0.0897		
Origin = Fukuoka Itazuke	0.4542	0.4582		
Dest = Fukuoka Itazuke	0.3735	0.4843		

All parameter estimates are significant at the 5% level, except when in *italic*.

Table 47 - Global Fit Statistics for Airport Pair Choice Models with Piecewise Linear Specification

	A 2-2	**A 3-2**	**A 4-2**	**A 1A-2**
Number of Individuals	30,126	30,126	30,126	30,126
Number of Choices per Individual	82	82	82	82
Number of Records	2,470,332	2,470,332	2,470,332	2,470,332
Degrees of Freedom (df)	20	30	28	30
$\ell\ell(0)$	-133,837.870	-133,837.870	-133,837.870	-133,837.870
$\ell\ell(\hat{\beta})$	-13,472.222	-13,341.393	-13,341.793	-12,973.224
Likelihood Ratio Test Statistic	240,731.296	240,992.955	240,992.155	241,729.292
ρ^2	0.8993	0.9003	0.9003	0.9031
$\bar{\rho}^2$	0.8992	0.9001	0.9001	0.9028

	H_0 : A 2-2 better than A 1A-2	**H_0 : A 3-2 better than A 1A-2**	**H_0 : A 4-2 better than A 1A-2**
z	0.003653659	0.002750854	0.0027389
N	30,126	30,126	30,126
J	82	82	82
K_1	20	30	28
K_2	30	30	30
Test Statistic	-31.3065	-27.0257	-27.0040
Upper Bound of Probability (H_0 = TRUE)	0.0000	0.0000	0.0000

For an explanation of symbols please see page 236.

6.4 Other Access and Egress Related Variables

The terminal pair choice models estimated are conditional logit models as explained in Chapter 4. We cannot enter socio-economic variables directly into a conditional logit model. Individual characteristics have to be interacted with, in our case, km of access or egress distance, e.g. "egress distance × company business." The parameter estimate then increases or decreases the slope of egress distance with respect to utility for all business travelers.

6.4.1 Trip Purpose

Recall that the data distinguishes between three trip purposes:

1. Company Business
2. Personal Business, and
3. Vacation

The trip purposes of company business and vacation are explicitly included in all models making people who travel on personal business the reference group.

Due to a lower level of complexity symmetries are more easily detected in the airport pair choice model (Table 34 on page 213). For vacationers the incremental disutility for an additional km of access distance is reduced by about one third from 0.050 to 0.034. That brings it below the level of egress distance which is 0.037. On the other hand, for business travelers the disutility of egress distance is increased to about the level of access distance. So for 91% of all air travelers (Table 37 on page 221) access is about as onerous as egress since the logarithmic terms are not substantially different. The similarity between access and egress for the “typical” air traveler, i.e. the business air traveler, is reflected in Table 43 on page 230. Contrast this to the different access and egress related utilities for HSR business passengers (Table 44).

Vacation as trip purpose also reduces the disutility of access distance for rail travelers by about one third. Vacation interacted with *egress* distance is not included in the air model because of its lack of explanatory power. It is also insignificant in two of the three HSR models in Table 35 and in three out of the four rail models with piecewise linear specification in Table 45 on page 246.

With vacation trips we can see why the access portion of the trip would reduce the disutility of feeder distance much more than the egress segment. “Getting there is half the fun.” The trip to the airport or the train station is seen as part of their holiday and vacationers are relaxed and are looking forward to their time off. This makes the access portion of the trip considerably less onerous than when traveling for personal business. On the other hand, when vacationers reach their destination station, they are usually tired and just want to get to their final destination as fast as possible. The transfer to the hotel in a very unfamiliar environment while already exhausted by the trip is seen as less pleasurable than the trip from home to the airport or railway station.

The coefficients for business travel are more difficult to explain since the interaction with egress distance has a negative sign for air, but is positive for HSR. The term is positive in all seven HSR models, significant at the 5% level in five of them, and fairly stable at around 0.002 to 0.003. The difference in signs is not likely due to chance fluctuation.

Business travel interacted with access distance does not seem to be important for air models, is insignificant in three rail models, and of a lower magnitude (generally below 0.002).

The HSR coefficients are the easiest to explain. When people travel on company business, the egress portion of the trip is customarily arranged and paid for by his or her employer, making it extremely convenient for the trip maker. So it is easy to see why egress would be less onerous when traveling on company business. On the other hand, many times the business traveler him or herself has to arrange and pay for the access part. So a slightly lower coefficient value for the business trip dummy interacted with access distance could be explained.

What has been learned about the difference between air and HSR concerning access and egress in the last few sections might lead us to at least speculate about the negative sign for airport egress. Airport access and egress distance appear to be perceived as being more onerous than access and egress to and from a HSR station. The linear component of both access and egress distance in log plus linear models was found to be more dominant in airport pair choice models, leading to a rapid decline in utility as distance increases (Figure 22 versus Figure 23). The range in which air travelers appear to be the most indifferent to changes in access or egress distance is between 15 km and 30 km (Table 46). For HSR travelers the equivalent range is between 60 km and 120 km (Table 45). The distribution of access and egress distance reveals that only HSR passengers are willing to travel long distances to and from the linehaul terminal (Table 41).

HSR passengers can easily and conveniently continue their journey on the narrow gauge (1 067 mm) railway. Except for a lower speed, the travel experience on the conventional railroad is likely to be very comparable to the high speed rail segment. Most business travelers would probably transfer from the high speed train's First Class to the conventional train's First Class and continue their work. Except for the inconvenience of a transfer, the marginal disutility of travel time might not change very markedly.

While the distinction between access, egress, and linehaul can be somewhat blurred for HSR passengers, air travelers may perceive those transitions as very abrupt. Egress from the airport can seldom be used productively and most business travelers might think of it as a "waste of time." That makes egress from the airport more onerous for the business segment than for people traveling on a different purpose. Chapter 3 introduced the notion that the distinction between access and linehaul is really the distinction between productive and unproductive time (3.2.2 The Competitive Triangle on page 70). It is against this background that the opposite sign of the company business indicator for air and HSR might be understood.

6.4.2 Sex

Sex being an important variable for station pair choice does not come as a surprise. Recall from the literature review that Harvey, in a study of airport access mode choice, had found non-business females evaluating the concept of "accessibility" to an intercity terminal completely differently than males (Harvey, 1986a, 1986b). Additionally, two studies in the United States observed that female travelers show a strong preference for modes that provide escorted door to door service, like being dropped off at the airport, or more secure modes like taxis and limousines (Harvey, 1988; Sobieniak et al., 1979).

Our models found sex to be a significant variable only for the egress segment of rail travel. Here, the new environment offers new challenges and new stimuli, reducing the onerousness of egress distance for males. Females, on the other hand tend to be more safety conscious and feel more insecure than males in unfamiliar territory.

6.4.3 Age

Young people tend to be better risk takers and also seem to possess a higher level of curiosity. Just like with males, this turns a new environment into an advantage. The new challenges and stimuli are more likely to be experienced with egress than with access. So it was not surprising that the same age variable interacted with access distance was not significant in any of the choice sets (Table 35) and for neither specification of access and egress distance (Table 45). Age was insignificant in all air models.

6.4.4 Population Density of Origin and Destination Zone

The interaction of access distance and population density measures the change in utility per km of access distance when population density is increases by 1 person per km^2. It is a proxy for the level of service that can be expected from the public transportation system.

The coefficient estimate is always positive, meaning a high population density in the origin or destination zone reduces the onerousness of access distance. The positive coefficient is easy to understand in the context of HSR station access, because the public transportation network of a region converges at major downtown rail stations.

Public transportation is less likely to be used for the egress portion of the trip, mainly due to unfamiliarity with the area. Destination zone population density is not an important explanatory variable for airport pair choice models and produces a coefficient of a slightly lower magnitude for HSR station pair choice models.

As we have seen from Figure 22 and Figure 23, this terminal pair choice determinant has little influence on total utility. However its coefficient estimates are significant in all of our models and do not fluctuate greatly. When the information is available it should be included in choice models.

6.4.5 Home Zone Automobile Ownership

As a proxy for automobile ownership of the trip maker, which we do not have, we use the average household automobile ownership percent in the traveler's home zone. This yields a variable that is related to the probability of vehicle ownership. The higher the probability, the more onerous egress and especially access distance are to the station. That is understandable. Vehicle ownership would indicate a person being used to the comfort of one's own car. Having to take public transportation to the HSR station of the region would be perceived as a step down. Another point to take into consideration is the inverse relationship between car ownership and transit service in a given zone. For the access portion the estimate is unstable but consistently significantly different from zero.

The disadvantage would not be felt as strongly on the egress portion of the trip. Travelers do not have their own cars available to them, and the variable mainly measures a certain discomfort that a car owner may feel when being more dependent on shared forms of transportation. The parameter estimate is of a lower magnitude and inconsistently significant.

The likelihood of the trip maker owning an automobile does not appear to influence the disutility of access to and egress from airports.

Together with Nozomi time and Nozomi transfers, access distance interacted with home zone automobile ownership is one of the only three variables whose R 100-2 estimate does not fall within the 95% confidence interval of model R 300-2. That means we cannot be very confident in the relative magnitude of those parameter estimates.

However if those coefficients are:

1. consistently significantly different from zero and
2. do not change signs

general conclusions like the ones reached in this subsection would still be appropriate.

6.4.6 Linehaul Time and Linehaul Transfers

Of the 25 variables selected for the HSR station pair choice model, twelve are significant at the one hundredth of one percent level for every single choice set (Table 31 on page 206). They are, as expected, linehaul time and linehaul transfers, both the linear and logarithmic component of access and egress distance, and most of the large station effect representatives. But we also find the interaction between both access and egress distance and linehaul time in this group. The coefficients are positive meaning that the relative importance of feeder distance decreases as linehaul time increases. The result seems intuitive, since as linehaul time increases the proportion of total travel time spent on access and egress decreases and therefore becomes less and less of a consideration.

This effect seems to be a little stronger for access distance. The egress distance coefficient is about two thirds the magnitude of that for access.

The important relationship between feeder distance and linehaul time may be more complicated than what the simple linear coefficient would suggest. At this point the author does not wish to postulate final conclusions about this interaction term.

6.5 Fixed Effects

Coefficients of fixed effects dummies represent the relative preference for a particular airport or HSR station, either as origin or destination, after differences in the other attributes (variables) are taken into account. As discussed at the beginning of this chapter, every critical terminal pair choice determinant with regards to the linehaul portion of the trip is included in both air and rail models. All key LOS (level of service) attributes for linehaul are accounted for, but we have them for neither access nor egress. In many instances distance to the terminal alone will not explain a particular choice, everything else being equal.

A home bound trip maker may disembark at a local Kodama station, only 5 km from his home, or at the Nozomi station in the central city, increasing egress distance to 15 km. But if travel via the local Kodama station would require taking a bus that only operates infrequently, while excellent subway service is available from the main station to a point close to home, the trip maker would probably choose the more distant central station as the destination terminal. The better level of service would be absorbed in the fixed effects coefficient of the main station.

A similar scenario might be envisioned for an air traveler whose home is located closer to Narita than Haneda. Transportation options are almost certainly better from central Tokyo and thereby Haneda.

We could expect that parking is generally more convenient at the smaller suburban and rural stations as opposed to central city terminals. All HSR fixed effects are positive. That means the portion of the coefficient/utility attributable to accessibility must only refer to public transportation. The magnitude of these coefficients and their relative large contribution to overall utility (Table 44) make it very unlikely that only accessibility issues by public transportation are represented by these effects.

When using these stations one passes through the core of urban activity. In many instances this is both convenient and pleasurable due to the high concentration of shops, services and entertainment. As gas stations in the United States have found out a long time ago, people are more likely to shop where they have to interrupt their journey anyway. A travel break in a central location is often far more productive than at a small out of the way railroad station.

Of the 36 stations on the Tokaido, Sanyo, and northern Shinkansen lines within metropolitan Tokyo, only eight stations were served by every Nozomi in 1995. They are Tokyo, Nagoya, Kyoto, Shin Osaka, Okayama, Hiroshima, Kokura, and Fukuoka. Even after controlling for shorter linehaul time and lack of transfers (all train categories serve Nozomi stations) Nozomi origin and or destination station make the largest contribution to station pair choice. Their coefficients are directly translated into utility by multiplication with one (Table 35 and Table 45).

Among all Nozomi stations, the size of Tokyo Central's fixed effect especially as a destination station stands out. Tokyo's coefficient is the Nozomi station coefficient plus 53% in the log plus linear (Table 35), and plus 42% in the piecewise linear model (Table 45). Tokyo is perhaps a special case inasmuch as it is the largest conurbation, but has nevertheless only one Nozomi station. Osaka has Kyoto (a Nozomi station) to the east and Kobe (a Hikari station) to the west. Tokyo only has Shin Yokohama (a Hikari station) to the west. Access from all other directions require a transfer at Tokyo Central.

On October 1st , 2003, a new station 7 km west of Tokyo Central was opened: Shinagawa. It is served by every train on the Tokaido line. Running the same models on the 2005 survey would most likely show a lower fixed effect for Tokyo Central as origin and destination station.

The airport pair choice model reveals a somewhat striking disparity of preferences for different airports. Ceteris paribus, the odds of choosing Fukuoka-Itazuke, the commercial airport nearest to the business district of Fukuoka, are only one fifth the odds of choosing Tokyo-Haneda, also the airport closest to downtown. On the other hand, the odds of selecting Itami (closest to the center of Osaka) are more than 3 times greater than the odds of deciding on Tokyo-Haneda.

Besides the access and egress related differences discussed in the context of HSR stations, there are other specific attributes that vary from airport to airport, like membership in an exclusive Airport Club or Lounge, the amount of walking required to and from the gates, the spaciousness or crowdedness of a particular airport, etc. This does not even include subjective valuations based on certain personal experiences, good or bad, which could overshadow all the other elements in deciding on an airport pair. These and more choice variables are reflected in the fixed effects.

Without an intimate knowledge of Japanese airports and air service one can only speculate why the two Osaka airports are rated so highly, while the two Fukuoka ones are valued so poorly. Fukuoka's hinterland, Kyushu, is the best served region by air in Japan. This is probably due to 1) its higher population density compared to areas of equal distance north of Tokyo, and 2) its lack of high speed rail service up to this point. Because of this intense competition between airports, the odds of choosing any of the two Fukuoka airports are low compared to the other major airports.

Travelers almost always prefer airports near the city to airports in outlying areas. The odds of choosing Narita for example are only 60% of the odds of Haneda. Why Osaka-Kansai even bests the already highly rated Osaka-Itami may be due to the fact, that Kansai is a new airport with first-rate facilities and outstanding express train service from Osaka. But it may also be because the Osaka Kansai coefficient is inaccurate.

The main conclusion of this section is that the relative attractiveness of a terminal has a far greater impact on HSR station pair choice than on airport choice.

6.6 Nested Models

As we know from Table 19 all observations with an origin / destination zone pair indicating that the competing mode is not available have to be excluded from the upper level mode choice model. Those are mainly HSR passengers between cities where air service has been abandoned (8 001 trips) and travel solely on the island of Kyushu.

Weights have to be normalized again for both air and HSR to add up to the new sample sizes. In the dataset used for upper level nest estimation the weights for air observations now add up to 28 678, and for HSR to 9 577. This would suggest that air was chosen three times more frequently than rail. The sum of the respective daily extension coefficients, however, implies the opposite. Among all the passengers represented by this combined air/HSR sample 69.5% chose rail and 30.5% air. In the fifth normalization all air weights are multiplied by 0.40653 and the rail weights by 2.77712. Weights now add up to 11 659 for air and to 26 596 for HSR (both are rounded to the nearest integer).

Table 48 shows the parameter estimates for the upper level nest. The lower level nest air and rail models were introduced at the beginning of this chapter as base models in Table 34 and Table 35 respectively. Coefficient estimates of inclusive values are reported using standard conventions. The actual results obtained were based on a different sign convention. For more information please see Appendix 4 on page 290.

Table 48 - Parameter Estimates of the Upper Nest Mode Choice Model

	Parameter Estimate	Standard Error	*p*-Value
HSR Intercept	2.9299	0.04170	<.0001
Inclusive Value for Air	0.0638	0.00631	<.0001
Inclusive Value for HSR	0.7271	0.01160	<.0001

The small magnitude of the air coefficient suggests high commonality between the air alternatives, i.e.: adding another airport pair would mostly draw passengers away from other airport pairs. Or, expressing the same idea slightly differently: the only airport pair that really matters is the best one.

The coefficient for the inclusive value of HSR is considerably larger. Adding a choice set to the HSR alternatives would be more sensible, since adding more stations would have a higher impact on other modes.

The model fit is somewhat disappointing (Table 49). Ideally, the off diagonal elements would be zero, resulting in the row totals (predicted mode) matching the column totals. The model correctly predicts HSR use for 95% of all travelers who actually chose HSR, which is good. However it breaks down for air passengers, more than half of whom are predicted to use rail.

On a more positive note the sum of the diagonal elements, the choices predicted correctly, is 30 854. That means 81% of all mode choices are mirrored accurately by the model. The model underpredicts the number of air passengers and overpredicts travel on the dominant mode, HSR.

Table 49 - Chosen Mode versus Predicted Mode in the Upper Nest

			Chosen: Air	Chosen: HSR	Chosen: Total
			11,659	26,596	38,255
Predicted	**Air**	7,171	**5,714**	1,457	
			49%	5%	
	HSR	31,084	5,945	**25,140**	
			51%	**95%**	
	Total	38,255	100%	100%	

The less than perfect fit is reflected in the global fit statistics, especially when compared to the lower level models (Table 50). The upper nest model has a $\bar{\rho}^2$ of only 0.21 versus 0.90 for the airport pair choice and 0.77 for the HSR terminal pair choice model.

Table 50 - Global Fit Statistics of Upper versus Lower Level Models

	Upper Level	Lower Level	
	Nested Model	**Airport Pair Choice A 1-2**	**HSR Station Pair Choice R 300-2**
Number of Individuals	38,255	18,395	18,395
Number of Choices per Individual	2	82	300
Number of Records	38,255	1,839,500	5,518,500
Degrees of Freedom (df)	2	24	25
$\ell\ell(0)$	-23,520.793	-133,837.870	-107,618.180
$\ell\ell(\hat{\beta})$	-18,684.408	-13,298.322	-24,760.941
Likelihood Ratio Test Statistic	9,672.771	241,079.097	165,714.479
ρ^2	0.2056	0.9006	0.7699
$\bar{\rho}^2$	0.2055	0.9005	0.7697

$\ell\ell(0)$ is the value of the log likelihood function when all parameters are zero.

$\ell\ell(\hat{\beta})$ is the value of the loglikelihood function at its maximum.

Likelihood Ratio Test Statistic (H_0: all β are 0) = $-2[\ell\ell(0)-\ell\ell(\hat{\beta})]$

$\rho^2 = 1 - \frac{\ell\ell(\hat{\beta})}{\ell\ell(0)}$

$\bar{\rho}^2 = 1 - \frac{\ell\ell(\hat{\beta})-\mathrm{df}}{\ell\ell(0)}$

7 Conclusions

The conclusions are based on the estimation results of multinomial logit models for terminal pair choice and nested logit models for mode choice. The empirical results are valid for the time period the survey was taken, the Fall of 1995, for the Tokaido and Sanyo Corridors (Tokyo – Osaka and Osaka – Fukuoka, respectively), and given the assumptions outlined in the last section of Chapter 4 on page 161.

Lessons learned from these empirical results that might be applicable to the planning of high speed rail systems in the United States are examined in the section on policy implications.

We believe that the methodological findings are sample independent.

7.1 Methodological Conclusions

7.1.1 Randomized Choice Sets

It has been shown that estimation of a multinomial logit model with just a subset of the alternatives, including the chosen one, yields consistent estimates. Section 5.6.2 A Consistent Estimator starting on page 199 points out that consistency is the weakest of the desirable properties of an estimator, since it only guarantees unbiasedness with an infinite sample size. How many observations are needed in practice to obtain useful approximations is not known.

We have found that the estimator for simple random sampling of alternatives exhibits the same properties with an increasing number of alternatives that we could expect with an increasing number of observations: some estimates appear to converge on a true value (see Figure 19 – Figure 21, pp. 204 – 205). The speed of convergence appears to be highly dependent on the correct specification of the model. We have also found that, as the number of alternatives increases, standard errors become more reliable.

Researchers are cautioned not to rely on a single model, blindly trusting confidence intervals, when analyzing estimation results obtained with the help of randomized choice sets. One can only properly understand the significance of individual mode choice determinants and become aware of the magnitude of their coefficients by looking at a *set* of models.

It would appear self-evident that as the number of sampled choices is increased, estimates become more precise. Our research also indicates that important effects might go undetected with limited sample sizes. The coefficients of the piecewise linear estimation of access distance for Model R 10A-2 in Table 45 on page 246 do not reveal any unusual pattern. Only when looking at models based on a larger number of random choices does a threshold effect start to become apparent. The systematic increase of the *t*-statistics for both access and egress distance illustrates the diagnostic value of a set of models based on a different number of alternatives.

7.1.2 Utility Impact Analysis of Access and Egress

The literature review (Chapter 2) reveals that many intercity mode choice studies, even up to the beginning of the 21st century, do not include access or egress related variables. A utility impact analysis for a "typical" air traveler and HSR passenger shows access and egress related variables contributing as much as 43% to total absolute utility (Table 44 on page 230). Please also see Figure 22 and Figure 23 on pages 228 – 229.

Our research would suggest at a minimum to include access and egress distance (or access plus egress distance) and fixed effects. The latter would reflect the relative attractiveness of a particular terminal compared to other stations or airports, and how important that terminal's attraction is in comparison to other mode or terminal pair choice determinants.

7.1.3 The Correct Specification of Access and Egress Distance

Common sense suggests that an increase in access or egress distance from 5 km to 6 km would be perceived as more onerous than from 105 km to 106 km. Sharply decreasing marginal disutilities at the beginning of the range of common access and egress distances are best modeled by a log plus linear function combination. This specification follows the piecewise linear line most closely (see e.g. Figure 26 on page 234).

Linear, quadratic, and cubic specifications cannot match the log plus linear or the piecewise linear function for the short distances where most of the observations occur. If they did, they would predict that the alternatives requiring longer feeder distances would never be chosen. And, as with all regressions based on minimizing least squares, outliers have a disproportionate influence on the final estimate.

As a result, linear, quadratic, and cubic specifications substantially underestimate the onerousness of access and egress distance in the range where most observations occur (Figure 29 and Figure 30 on page 241). For our data the linear specification underestimates the negative utility of HSR egress distance at the median and at the mean by about two thirds. The quadratic and cubic functions still undervalue that same disutility by about one half.

A careful analysis of a *set* of HSR station pair choice models with a piecewise linear specification for access and egress distance (Table 45 on page 246) reveals strong evidence of threshold effects around 100 km. The airport model also exposes a similar effect, but for a much shorter access and egress distance (Table 46 on page 247). It is very difficult for a function composition which includes a logarithmic term to model a "flat spot" on the utility curve close to the origin. This may be the reason why the cubic specification fits airport access and egress data better than the log plus linear combination. The airport example demonstrates that a "one size fits all solution" for the modeling of access and egress distance with a continuously differentiable function for researchers whose main focus is not access or egress does not appear to exist.

Because it would be difficult for the researcher to know in advance whether strong threshold effects exist, and if they do, where they would occur, the only recommendation that can be given at this point is to estimate access and egress distance, or the sum of access and egress distance (feeder distance) using both a log plus linear and a cubic specification and then select the one that fits the data best. Our research results would discourage the consideration of linear or quadratic specifications.

The evidence of threshold effects also points to the necessity of experimenting with piecewise linear functions when a deeper understanding of the importance and the behavior of access and egress related variables is desired.

7.2 Empirical Conclusions

7.2.1 Feeder Distance to Airports versus Feeder Distance to HSR Stations

Airport access and egress distance appear to be perceived as being more onerous than access and egress to and from a HSR station. The linear component of both access and egress distance in log plus linear models was found to be more dominant in airport pair choice models, leading to a rapid decline in utility as distance increases (Figure 22 on page 228 versus Figure 23 on page 229). The range in which air travelers appear to be the most indifferent to changes in access or egress distance is between 15 km and 30 km (Table 46 on page 247). For HSR travelers the equivalent range is between 60 km and 120 km (Table 45 on page 246). The distribution of access and egress distance reveals that only HSR passengers are willing to travel long distances to and from the linehaul terminal (Table 41 on page 224).

For egress many HSR riders are able to continue on the conventional railroad. Their travel experience during the egress portion is likely to be comparable to that on the high speed rail segment. The same would apply to access by conventional rail. While the distinction between access, egress, and linehaul can be somewhat blurred for HSR passengers, air travelers may perceive those transitions as being very abrupt. Access to or egress from the airport can seldom be used productively.

This may help explain why egress distance interacted with an indicator variable for company business travel is negative for air, but positive for HSR.

7.2.2 The Difference between Access and Egress

At first sight, access to the airport may appear to be more onerous than egress. The magnitude of the linear coefficient for airport egress distance is 26% smaller than the one for access distance, while the logarithmic terms are not substantially different (Table 34 on page 213). However for business travelers the negative interaction term with egress distance about equalizes both disutilities. The same is true for vacation travelers, except that in this case the positive interaction term with access distance brings the disutility of access down to the level of egress. Since business people and vacationers constitute about 91% of the air clientele, access and egress does not appear to be dissimilar for air travelers in general. This is reflected in Table 43 on page 230.

HSR travelers do exhibit different preferences concerning access and egress (Table 44 on page 230). For the most common values of access and egress distance (experienced by about 95% of all HSR travelers) access seems to be perceived as being more onerous. However HSR passengers show a stronger dislike for long egress distances (about 120 km to 140 km and above) than for the equivalent length access distance. We might call this a "scissors effect." Please see Figure 27 on page 235. This finding is confirmed by the fact that the threshold effect is much stronger for egress. The larger drop followed by a steeper rise in the magnitude of the egress coefficients may indicate a much steeper threshold of pain for the egress portion of the trip, while the increase in pain during access might be more gradual (Table 45 on page 246). The stronger threshold effect for egress is also visible in Figure 27 on page 235.

Whether HSR travelers really consider access to be more onerous than egress, or whether they give more weight to access considerations because they do not have as much information about the egress portion of the trip (the behavioral equivalent to weighted least squares), is not clear at this time. What we do strongly suspect though is that once egress distance has increased beyond a certain point, travelers tend to become very aware of it.

7.2.3 The Importance of Fixed Effects

The relative attractiveness of a terminal has a far greater impact on HSR station pair choice than on airport pair choice. Table 44 on page 230 shows that fixed effects can contribute as much as 44% of the total absolute utility for a HSR station pair. By contrast, fixed effects only explain about 10% of the air utility (Table 43).

All key LOS (level of service) attributes for the linehaul portion of the trip are accounted for. That means fixed effects parameters for HSR mainly reflect the accessibility of the particular terminal by public transportation (large station fixed effects are all positive and central stations in the main business district are probably not known for ample free parking). But they also mirror the attractiveness of the central districts in which they are located. As hub for the business and entertainment activity of a particular region they are in some way the ideal place to stop for a break in the journey.

7.2.4 Nested Model Results

Which airport I use does not matter that much for terminal pair choice as long as frequent nonstop service is available and access and egress is reasonably short. The equivalent HSR level of service attributes, linehaul time and number of transfers, in addition to access and egress distance, do matter as well for HSR station pair choices, but not to the same extent as they do for airport pair choice. How important and how centrally located the HSR terminal is has a disproportionately large influence on station pair choice. That sums up what we might be able to conclude from the significant impact that large station fixed effects seem to have on overall utility.

The nested model results can be interpreted similarly. The inclusive value coefficient for HSR is of a much greater magnitude than the one for air. In addition to that the intercept increases the difference between air and HSR even more (Table 48 on page 261). What that means is that any improvements in the HSR nest will draw ridership out of the air nest, while improvements in the air nest, e.g. the addition of a new airport, will take travelers mostly away from other airport pairs. The addition of a new HSR terminal would have a beneficial effect on the rail mode. A new airport on the other hand would do very little for the competitiveness of the air mode as a whole vis-à-vis HSR.

7.3 Policy Implication – A Paradigm Shift to Convergence

Of great concern to transportation planners in the United States would be the degree to which HSR utility in Japan appears to depend on the "pull" that large rail stations exert on travelers. The public transportation system of the whole metropolitan area converges on these stations making them the transportation hub of the region. In the United States only Grand Central Station in New York and to a lesser extent Union Station in Washington, DC would be able to duplicate this draw.

Transbay Terminal in San Francisco is the proposed Northern terminus of the California high speed rail system. This station is so dilapidated and has such a bad reputation that even Amtrak moved out and now picks up its passengers at the Ferry Building. Transbay Terminal so far only has repulsive properties. There are plans to rebuild this station, but the challenge to move its image in the public's eye from extremely negative to very positive is formidable.

If the utility of HSR travel in the United States would depend on the magnetism of large downtown stations as much as it appears to in Japan, innovative solutions must be found to replicate this effect.

One innovative solution, which is particularly suited to relatively low density and vastly dispersed US metropolitan areas, was presented in Chapter 1 and illustrated by Figure 1 on page 9.

To visualize another innovative solution one might look to Toronto, Canada. The downtown area near Union Station is interlaced with a labyrinth of underground passage ways that not only keep pedestrians insulated from the harsh weather, but which also convey the impression that the downtown area in the immediate vicinity of Union Station is just like one big “airport” terminal combining transportation functions with shopping, dining, and entertainment.

Protection from harsh weather is not as big of an issue in San Francisco, so the equivalent scenario might consider aerial walkways on the second floor level. A shopping street might be similar in appearance to a multi-story shopping mall, except that cars and delivery vehicles occupy the ground floor.

Both scenarios, underground passage ways and aerial walkways, achieve the same objective: pedestrian flows unimpeded by other modes, creating the atmosphere of a big shopping mall or airport terminal.

Passengers using Detroit’s Metro Airport were required to walk up to 1500 m before the opening of the people mover system. If HSR passengers were willing to walk only half that distance from any of the four downtown San Francisco Market Street subway stops, aided by underground passageways or aerial walkways, the center of San Francisco from City Hall to the Embarcadero including parts of Chinatown and the South of Market Convention Center district, would be one single HSR terminal (Map 7). Anybody originating in or destined to this area would only be subject to the terminal time and terminal distance that most airline passengers experience as a routine part of their journey. Staying in a hotel in downtown San Francisco would be akin to staying in a hotel located inside of the central terminal area of an airport.

To briefly address some of the doubts the reader might have, the technical feasibility of operating high speed trains and light rail vehicles on the same track was proven by the well known Karlsruhe Model in Germany. Having street cars and long distance trains share the same infrastructure was also common practice in the United States prior to World War II. Rail vehicles serving the five Muni Metro lines that use the Market Street subway can be coupled together and operated as a single consist before branching out to different destinations. To some extent, this already is standard operating procedure. The tunnel is not capacity constrained as long as light rail vehicles only require one train path every five minutes.

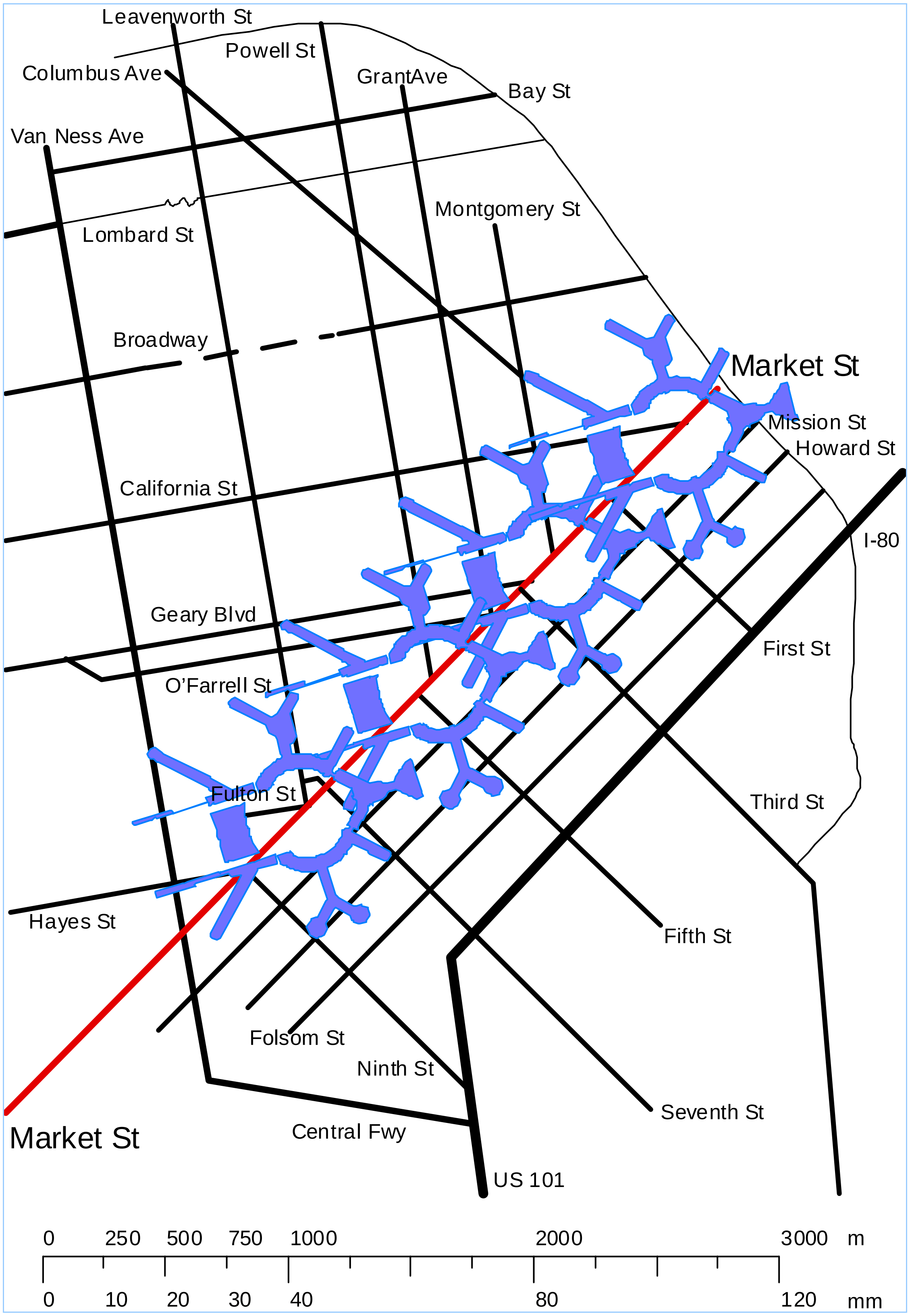
If all of this still sounds too futuristic, it may be because of conscious and unconscious restrictions imposed by an old, and now largely obsolete, paradigm. Intercity rail in general and high speed rail in particular has traditionally been viewed as a completely different mode than urban rail transportation systems. The basic paradigm has always been to optimize rail systems for one particular application like airport people movers, urban transit like MUNI, regional rapid transit like BART, and intercity transportation like high speed rail. If regional rapid transit systems like BART and some urban transit systems like the Los Angeles Red Line would be seen as a completely integrated part of an inter-metropolitan high speed rail system, HSR would be much more competitive.

The incompatibilities that have been designed into modern urban transit systems like non-standard gauges for the Bay Area Rapid Transit (BART) system, the Washington metro, and the Toronto subway, are the most ostentatious symptoms of the old paradigm. Going hand in hand with this old paradigm, high speed rail systems proposed for the United States have been mostly point to point systems, copying the European and Japanese models, even though urban densities and land use patterns are vastly different in North America. The new paradigm sees urban, regional, and high speed rail as one coherent system. It is the author's belief that this much more sophisticated approach is what is needed to make high speed rail work and work well in the North American market.

*

There is one more policy implication that can be derived from this research, which focused on the impact of access and egress related issues on terminal pair and mode choice. An analysis of high speed rail that only looks at a Los Angeles Union Station to San Francisco Transbay Terminal connection or equivalent cannot come to the conclusion that HSR is infeasible.

A feasibility analysis must take all competitive advantages of HSR into account or present good reasons for excluding them. The internal distribution advantage of HSR, summarized and defined on pages 3 and 4, is an integral part of this mode's positioning in the market place.



Map 7 - Downtown San Francisco with SF Int'l Airport Terminal Overlay, Scale 1 : 25000

Source: USGS San Francisco North (1995) and Montara Mountain, CA (1997) Quadrangles

Appendix

Appendix 1 – HSR Survey Instrument for 1995 Intercity Travel Survey

Government Registered Survey No. 19756
Effective until 95-11-30

Intercity Travel Survey

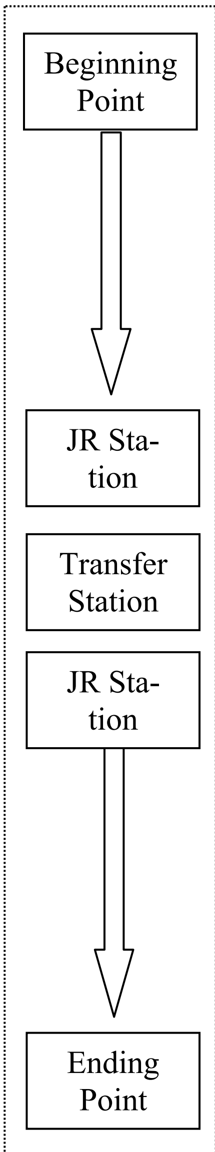
Ministry of Transportation

The Ministry of Transportation is conducting a Survey of Intercity Travelers. We appreciate your cooperation. Our staff will come to collect your completed surveys later.

Note There are some answers to be circled and others to be written in the box, such as area or station names. Please fill it out again even if you have already been asked the same question before.

Q1 Is this a one-day trip?
1 Yes 2 No

Q2 Please describe your travel schedule.



Where did you begin your journey?
1 Home 2 Other … Prefecture ☐ City ☐
(country if you came from abroad)

Please enter the modes of transportation from the beginning to first JR station.

Beginning Point ☐ ☐ ☐ ☐ ☐ First JR Station

1 Bicycle, Motorbike, Walk	2 Domestic Car
3 Taxi, Limousine	4 Private Railway, Subway
5 Inner-City Bus	6 Express Bus
7 Airplane	8 Other ( )

What was your first JR station? ☐ Station

Which are your major transfer stations? ☐ ☐ ☐ ☐

What will be your final JR station? ☐ Station

Please enter the modes of transportation from last JR station to ending point.

Final JR Station ☐ ☐ ☐ ☐ ☐ Ending Point

1 Bicycle, Motorbike, Walk	2 Domestic Car
3 Taxi, Limousine	4 Private Railway, Subway
5 Inner-City Bus	6 Express Bus
7 Airplane	8 Other ( )

Where will you end your journey?
1 Home 2 Other … Prefecture ☐ City ☐
(country if you are going abroad)

Q3 If a portion of the trip described above was by air, please list the airports.

From [ ] To [ ] Airport

Q4 What is the purpose of your trip?

1 Business 2 Vacation 3 Personal Business 4 Commuting 5 Other

Q5 How many people are traveling (including yourself)?………… ( ) person

Q6 Please describe the type of ticket you purchased.

1 General/Express Ticket 2 Multi-Ride Ticket
3 Other Discount Tickets 4 Monthly Pass

Q7 Please describe yourself.

Where do you live? Prefecture [ ] City [ ]

What is your sex? 1 Male 2 Female

What is your age? 1 Under 19 2 20s 3 30s 4 40s
5 50s 6 60s 7 70s

What is your occupation?

1 Management/Board Member 2 Businessman
3 Civil Servant 4 Agricultural Worker
5 Commerce/Own Business 6 No Full-time Occupation
6 Student 7 No Occupation/Other

Thank you for your cooperation.

Ministry of Transportation

Appendix 2 – Database Record Layout of the 1995 Intercity Travel Survey

Traffic Mode	Line-haul Mode	Beginning Point								Ending Point								Purpose of Trip	Length of Trip	Overseas Trip
		5 Digit Code					207 Code			5 Digit Code					207 Code					
1	1	5					3			5					3			1	1	1
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

Route																												
Air																												
Departure Airport								Transfer Airport 1								Transfer Airport 2								Arrival Airport				
Airport Code					207 Code			Airport Code					207 Code			Airport Code					207 Code			Airport Code				
5					3			5					3			5					3			5				
22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50

Route																				
Air (cont'd)			Rail																	
Arrival Airport (cont'd			Departure Station									Transfer Station 1								
207 Code			Station Code						207 Code			Station Code						207 Code		
3			6						3			6						3		
51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71

Route																												
Rail																												
Transfer Station 2									Transfer Station 3									Transfer Station 4									Arrival Station	
Station Code						207 Code			Station Code						207 Code			Station Code						207 Code			Station Code	
6						3			6						3			6						3			6	
72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Route																												
Rail (cont'd)							Bus																					
Arrival Station (cont'd)							Departure Station											Arrival Station										
Station Code (cont'd)				207 code			Bus Station Code								207 Code			Bus Station Code								207 Code		
6				3			8								3			8								3		
101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129

Route																				
Ferry																				
Departure Port											Arrival Port									
Port Code								207 Code			Port Code								207 Code	
8								3			8								3	
130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150

Route															Transfer Information										
Ferry (cont'd)	Automobile														Trans-fer Flag	Mode		Transfer Location		Transfer Location		Mode		Others	
Arrival Port	Departure Interchange							Arrival Interchange								From	To	Prefecture Code		Prefecture Code		From	To	Acc-ess	Egr-ess
207 Code	Interchange Code				207 Code			Interchange Code				207 Code													
3	4				3			4				3			1	1	1	2		2		1	1	1	1
151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176

Personal Attributes										Expansion Coefficient													
Sex	Age	Home Address								Daily Expansion Coefficient 1										Daily Expansion Coef 2			
		5 Digit Code					207 Code			Integer					Decimal					Integer			
1	1	5					3			5					5					5			
177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200

Expansion Coefficient (cont'd)																				
Daily Expansion Coefficient 2 (cont'd)						Annual Expansion Coefficient 1														
Integer (cont'd)	Decimal					Integer												Decimal		
5	5					12												3		
201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221

Expansion Coefficient (cont'd)															ID Code 1	ID Code 2	ID Code 3	ID Code 4	ID Code 5
Annual Expansion Coefficient 2																			
Integer												Decimal							
12												3			1	1	1	1	1
222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241

blank								
242	243	244	245	246	247	248	249	250

Appendix 3 – Randomizing Algorithm

There are at least two different methods of creating a randomized choice set. The first one is the one most often used. Its disadvantage is that the number of choices selected for each individual vary greatly.

```
data myFirstSample10;
      set Japan.HSR_los12 (where=(stationPairChoice=2));
      if ranuni(-1) < 0.00714853057982;                     * = 9/1259;
run;
```

The second method, shown on the next page, will guarantee that each individual will have exactly the same number of choices. It was developed by SAS Corp. and is taught as part of a computing class at UC Berkeley (Spector, 2002).

Both methods were used and yielded comparable results.

```
 * RANDOMIZED CHOICE SET GENERATOR - guarantees that each individual will have the same number of choices;

%let K=9;                                                                * total # of unselected choices desired;
%let file = Japan.sample10;
%let needed = (keep = id DEC stationPairChoice accessDistance egressDistance nozomiTime2 nozomiTransfers
                              HSRorigin HSRdestination companyBusiness male VehOwnershipPerMille vacation
                              youngAge originZonePopulation originZoneSq_km
                              destinationZonePopulation destinationZoneSq_km);

proc sort data=Japan.HSR_los12;
      by id descending stationPairChoice;            *   2 = not chosen 1 = chosen last.id is chosen pair;
run;
data &file;
      retain k n;                                    * retain the value of these vars between obsrvations;
      drop   k n;                                   * temporary variables for the duration of the session;
      set Japan.HSR_los12 &needed;
            by id;

      if last.id then do;
            output;                                                        * outputs the chosen alternative;
            return;
      end;
      if first.id then do;                                                           * initializes k and n;
            k=&K;                                                  * k = # of non-chosen alternatives desired;
            n=1259;                                        * n = # of non-chosen alternatives still available;
      end;

      if k=0 then return;
      if ranuni(-1) < k/n then do;
            output;
            k=k-1;
      end;
      n=n-1;
run;

* The data step in SAS is an implied FOR Loop.  The data step is executed for each observation.
```

Appendix 4 – Sign Convention

There are at least three different ways in which we can summarize the nested logit model results shown in Table 48 on page 261.

$$V_{HSR} = 2.9299 + 0.7271\ \text{InclValue}_{HSR} \tag{105}$$

$$V_{AIR} = 0.0638\ \text{InclValue}_{AIR} \tag{106}$$

$$V_{HSR} - V_{AIR} = 2.9299 + 0.7271\ \text{InclValue}_{HSR} - 0.0638\ \text{InclValue}_{AIR} \tag{107}$$

While the individual utilities only have positive coefficients, the utility of HSR relative to that of air has a negative coefficient. It makes sense, because as the utility of air increases the utility of HSR relative to that of air decreases. (107) is what binary mode choice models estimate. Since our nested logit model was estimated sequentially, the upper nest coefficients were estimated as a simple binary mode choice model.

SAS output:

```
                  Analysis of Maximum Likelihood Estimates

                                         Standard          Wald
Parameter             DF     Estimate       Error    Chi-Square    Pr > ChiSq

Intercept              1       2.9299      0.0417     4938.1897        <.0001
inclusiveValue_Air     1      -0.0638     0.00631      102.0374        <.0001
inclusiveValue_HSR     1       0.7271      0.0116     3959.2048        <.0001
```

Limdep/N-Logit output:

```
+---------+-------------+---------------+--------+---------+---------
 Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean (X)
+---------+-------------+---------------+--------+---------+---------
          Characteristics in numerator of Prob[Y = 1]
Constant      2.92993073      .04169399    70.272   .0000
INCL_AIR      -.06377305      .00631332   -10.101   .0000   -8.73342834
INCL_HSR       .72714195      .01155620    62.922   .0000   -3.40868256
```

Back References

Agresti, A. (1990). *Categorical Data Analysis*. New York: Wiley.
Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. New York: Wiley.
Agresti, A., & Finlay, B. (1986). *Statistical Methods for the Social Sciences* (2nd ed.). London: Collier Macmillan.
Airport Capacity Benchmark Report 2001. (2001). Washington, D.C.: Department of Transportation - Federal Aviation Administration.
Airport Characteristics Data Bank. (1992). Montréal: International Civil Aviation Authority (ICAO).
Anton, H. (1992). *Calculus with Analytic Geometry* (4th ed., brief ed.). New York: John Wiley & Sons.
Bakker, J. J., Calkin, J., & Sylvester, S. (1988). Multi Centered Time Transfer System for Capital Metro, Austin, Texas. *Transportation Research Record,* **1202**, 22-28.
Barwise, P., & Meehan, S. (2004). *Simply Better*. Boston: Harvard Business School Press.
Beimborn, E. A. (1968). *An Examination of the Effects of the Access Portion of a Trip upon an Intercity Mode Choice.* Dissertation, Northwestern University, Evanston, Ill.
Beimborn, E. A. (1969). Terminal Access and the Choice of Intercity Modes, *Transportation Engineering Journal, Proceedings*. Reston, VA: American Society of Civil Engineers.
Ben-Akiva, M., & Swait, J. (1984). *The Akaike Likelihood Ratio Index* (Working Paper). Cambridge, Mass.: Department of Civil Engineering - Massachusetts Institute of Technology.
Ben-Akiva, M. E. (1977). *Choice Models with Simple Choice Set Generating Processes* (Working Paper). Cambridge, Mass.: Department of Civil Engineering, MIT.
Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, Mass.: MIT Press.
Ben-Akiva, M. E., & Morikawa, T. (1990a). Estimation of switching models from revealed preferences and stated intentions. *Transportation research,* **24A** (6), 485-495.
Ben-Akiva, M. E., & Morikawa, T. (1990b). Estimation of Travel Demand Models from Multiple Data Sources. In M. Koshi (Ed.), *Transportation and Traffic Theory* (461-476). Oxford: Elsevier Science Ltd.
Bernick, M., & Cervero, R. (1997). *Transit Villages in the 21st Century*. New York: McGraw-Hill.
Bhat, C. R. (1995). A Heteroskedastic Extreme Value Model of Intercity Travel Mode Choice. *Transportation Research,* **29B** (6), 471-483.
Bowman, J. L. (1998). *The day activity schedule approach to travel demand analysis.* Dissertation, Massachusetts Institute of Technology, Cambridge, Mass.
Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society, Series B,* **26**, 211-214.
Boyce, W. E., & DiPrima, R. C. (1992). *Elementary Differential Equations and Boundary Value Problems*. New York: John Wiley & Sons.

Briginshaw, D. (2005). Safety Must Take Priority. *International Railway Journal,* **25** (6), 1.
Brög, W. (1982). Application of the Situational Approach to Depict a Model of Personal Long Distance Travel. *Transportation Research Record,* **890**, 24-33.
Brög, W. (1988). *Zukunftsperspektiven des ÖPNV: Szenarien und Handlungsalternativen (Future Perspectives of Public Transit: Alternative Courses of Action - in German).* Paper presented at the VÖV-Jahrestagung '88.
Brown, D. A. (1968). *An Intercity Passenger Transportation Demand Model.* Dissertation, Stanford University, Stanford, Calif.
von Bülow, A. (2003). *Die CIA und der 11. September*. München, Zürich: Piper Verlag GmbH.
California Intercity High Speed Rail Commission, & Charles River Associates. (1996). *Independent Ridership and Passenger Revenue Projections for High Speed Rail Alternatives in California: Final Report*. Boston, Mass.: Charles River Associates.
Central Transportation Planning Staff. (1997). *Transfer Penalties in Urban Mode Choice Modeling* (Report DOT-T-97-18). Washington, D.C.: US Department of Transportation, Federal Transit Administration.
Cervero, R. (1994). Rail Oriented Office Development in California: How Successful? *Transportation Quarterly,* **48** (1), 33-44.
Cervero, R., & Duncan, M. (2002). *Transit's Value-Added: Effects of Light and Commuter Rail Services on Commercial Land Values.* Paper presented at the Transportation Research Board 81st Annual Meeting, January 13-17, 2002, Washington, D.C.
Cervero, R., & Landis, J. (1993). Assessing the Impacts of Urban Rail Transit on Local Real Estate Markets Using Quasi Experimental Comparisons. *Transportation Research,* **27A** (1), 13-27.
Chiiki Keizai Soran (Comprehensive Databank of Regional Economics - in Japanese). (2000). Tokyo: Toyo Keizai Shinposha.
Clever, R. (1994). Eine Zukunftvision des InterCity Systems: notwendige Änderungen in der Fahrplan- und Tarifgestaltung. *Zeitschrift für Verkehrswissenschaft,* **65** (2), 121-147.
Clever, R. (1996). Schnelligkeit oder Häufigkeit: Überlegungen zur Einführung des Integralen Taktfahrplans im Fernverkehr der Eisenbahn. *Zeitschrift für Verkehrswissenschaft,* **67** (2), 138-182.
Clever, R. (1997). Integrated Timed Transfer: a European Perspective. *Transportation Research Record,* **1571**, 109-115.
Clever, R. (2005). Eine Vision für Berlin und den Transrapid. *Zeitschrift für Verkehrswissenschaft,* **76** (1), 69-89.
Cohen, G. S., Erlbaum, N. S., & Hartgen, D. T. (1978). Intercity Rail Travel Models. *Transportation Research Record,* **673**, 21-25.
DeFao, J. (2000-11-2). `Transit Village' Planned in Oakland Project - Planned for MacArthur BART Area. *San Francisco Chronicle,* p. A22.
Dobson, A. J. (2001). *An Introduction to Generalized Linear Models* (2nd ed.). Boca Raton, Florida: Chapman & Hall/CRC.

Domencich, T. A., & McFadden, D. (1975). *Urban Travel Demand : a Behavioral Analysis : a Charles River Associates Research Study*. Amsterdam: North-Holland Pub. Co.

Ellis, R. H., Bennett, J. C., & Rassam, P. R. (1974). Approaches for Improving Airport Access. *Transportation Engineering Journal,* **100** (TE3).

First N700 Shinkansen Train Starts Testing. (2005). *International Railway Journal,* **45** (5), 40-41.

Forinash, C. V. (1992). *A Comparison of Model Structures for Intercity Travel Mode Choice.* Masters Thesis, Northwestern University.

Forinash, C. V., & Koppelman, F. S. (1993). Application and Interpretation of Nested Logit Models of Intercity Mode Choice. *Transportation Research Record,* **1413**, 98-106.

G-7 Countries: Transportation Highlights. (BTS99-01)(1999). U.S. Department of Transportation, Bureau of Transportation Statistics.

Gaudry, M. J. I. (1991). Three Families of Mode Choice Models Applicable to Intercity Travel Demand with Aggregate Data, *Canadian Transport Policy* (43-50). Kingston, Ont.: John Deutsch Institute for Economic Policy.

Gaudry, M. J. I. (1995). *Three Families of Mode Choice Models Applicable to Intercity Travel Demand with Aggregate Data*. Montréal: Université de Montréal, Centre de recherche sur les transports.

Gaudry, M. J. I., & Dagenais, M. G. (1979). The DOGIT Model. *Transportation Research,* **13B**, 105-112.

Gaudry, M. J. I., & Wills, M. J. (1978). Estimating the Functional Form of Travel Demand Models. *Transportation Research,* **12** (4), 257-289.

Gemeindeverzeichnis. (2002). Wiesbaden: Statistisches Bundesamt Deutschland.

Godin, S. (1999). *Permission Marketing*. New York: Simon & Schuster.

Goldberger, A. S. (1991). *A Course in Econometrics*. Cambridge, Mass.: Harvard University Press.

Gosling, G. D. (1984). *An Airport Ground Access Mode Choice Model* (Technical Document UCB-ITS-TD-84-6). Berkeley, Calif.: Institute of Transportation Studies, University of California.

Goulias, K. G., Brög, W., & Erl, E. (1998). Perceptions in Mode Choice Using the Situational Approach : a Trip by Trip Multivariate Analysis for Public Transportation. *Transportation Research Record,* **1645**, 82-93.

Grayson, A. (1981). A Disaggregate Model of Mode Choice in Intercity Travel. *Transportation Research Record,* **835**, 36-42.

Green, P. E., & Srinivasan, V. (1978). Conjoint Analysis in Consumer Research: Issues and Outlook. *Journal of Consumer Research,* **5**, 103-123.

Gujarati, D. N. (1995). *Basic Econometrics* (3rd ed.). New York: McGraw-Hill.

Hansen, M. (1990). Airline Competition in a Hub Dominated Environment : an Application of Noncooperative Game Theory. *Transportation Research,* **24B** (1), 27-43.

Hansen, M. M. (1988). *A Model of Airline Hub Competition.* Dissertation, Institute of Transportation Studies University of California Berkeley, Berkeley, Calif.

Harvey, G. (1986a). *Analyzing Airport Access in a Multiple Airport Region*: National Research Council, Transportation Research Board.

Harvey, G. (1986b). A Study of Airport Access Mode Choice. *Journal of Transportation Engineering,* **112** (5).
Harvey, G. (1987). Airport Choice in a Multiple Airport Region. *Transportation Research,* **21A** (6).
Harvey, G. (1988). *ACCESS : Models of Airport Access and Airport Choice for the San Francisco Bay Region* (Report). Berkeley, Calif.: Deakin Harvey Skabardonis.
Hensher, D. A. (1997). A Practical Approach to Identifying the Market Potential for High Speed Rail: a Case Study in the Sydney-Canberra Corridor. *Transportation research,* **31A** (6), 431-446.
Hensher, D. A. (1998). Intercity Rail Services: a Nested Logit Stated Choice Analysis of Pricing Options. *Journal of Advanced Transportation,* **32** (2), 130-151.
Hensher, D. A., Brotchie, J., & Gunn, H. F. (1989). A Methodology for Investigating the Passenger Demand for High Speed Rail, *14th Australasian Transport Research Forum* (Vol. 1, 459-476).
Hope, R. (2003). Tokaido Upgrade to Beat back Air Competition. *Railway Gazette International,* **159** (4), 199-201.
Horowitz, A. J. (1981). Subjective Value of Time in Bus Transit Travel. *Transportation,* **10**, 149-164.
Horowitz, A. J. (1983). Statistical Comparison of Non-Nested Probabilistic Choice Models. *Transportation Science,* **17**.
Hughes, M. (2004). On Track to Hit Market Share Target. *Railway Gazette International,* **160** (10), 691-694.
Infratest Sozialforschung. (1992). *Umsteigewiderstände als Bestimmungsfaktor der Nachfrage im Schienenpersonenfernverkehr (Transfer Resistance as Decision Variable of Demand in Intercity Rail Transport - in German)* (Report). München.
Investigation Into Japanese Crash Begins. (2005, 2005-04-29). International Railway Journal. Available: http://www.railjournal.com/latenews_archive.html [2005, 2005-05-31].
Jackson, C. (2004a). Pioneer High Speed Line Is still at The Cutting Edge. *Railway Gazette International,* **160** (10), 677-680.
Jackson, C. (2004b). Rival Concepts Advance Train Designs. *Railway Gazette International,* **160** (10), 682-684.
Jackson, C. (2004c). Shinkansen Recovery Boosts Growth. *Railway Gazette International,* **160** (10), 686-688.
Jackson, C. (2005). Effective & Speedy Wins the Traffic. *Railway Gazette International,* **161** (5), 221-223.
Jänsch, E. (1994). Hochgeschwindigkeitsverkehr in Europa Wächst Zügig (High Speed Rail Service Is Growing Fast in Europe - German). *Eisenbahntechnische Rundschau* (5).
Jänsch, E., & Rump, R. (1990). Hochgeschwindigkeitsverkehr international. In P. Münchschwander (Ed.), *Schienenschnellverkehr* (Vol. 4, 250). Heidelberg: R. v. Decker's Verlag, G. Schenck GmbH.
Jinbunsha Editorial Staff (Ed.). (2000). *Shin Zenkoku Chimei Yomigana Jiiten (A New Dictionary and Pronunciation Guide of Place Names for the Whole Country - in Japanese)*. Tokyo: Jinbunsha.

Jomard, J. M. (2000). Eurostar Fête son Cinquième Anniversaire. *Revue Générale des Chemins de Fer* (4), 45-46.
Kanafani, A. K., & Youssef, W. (1993). *Roles and Market Share: the Interaction between Air and High Speed Rail - Transportation in California: Phase one Final Report*. Berkeley, Calif.: Institute of Transportation Studies, University of California.
Kasai, Y. (2000). Tokaido Shinkansen Reaches Technical Perfection. *International Railway Journal,* **40** (1), 30-33.
Kennedy, P. (1992). *A Guide to Econometrics* (3rd ed.). Cambridge, Mass.: MIT Press.
Knutton, M. (2004). Vertical Integration Proves to Be a Winner. *International Railway Journal,* **19** (10).
Koppelman, F. S. (1989). Multidimensional model system for intercity travel choice behavior. *Transportation Research Record,* **1241**, 1-8.
Koppelman, F. S., & Hirsh, M. (1986). Intercity Passenger Decision Making: Conceptual Structure and Data Implications. *Transportation Research Record,* **1085**, 70-75.
Koppelman, F. S., Kuah, G. K., & Hirsh, M. (1984). *Review of Intercity Passenger Demand Modeling : Mid-60's to the Mid 80's*. Evanston, Ill.: The Transportation Center and Department of Civil Engineering, Northwestern University.
Kraft, J., & Kraft, A. (1976). Mode Choice Characteristics as Determinants of Interurban Transport Demand. *Transportation Research,* **10** (1), 31-35.
Krista, M., & Oltrogge, C. (undated). *Linienplanung Fernverkehr DB 2000 Abschlussbericht. (Long Distance Route Planning for German Rail in the Year 2000 - Final Report - in German)*. Braunschweig: Ingenieurgesellschaft für Verkehrsplanung und Verkehrssicherheit.
Kuhfeld, W. F. (2003). *Marketing Research Methods in SAS* (SAS 9.1 ed. Vol. TS-689). Cary, NC: SAS Institute Inc.
Leake, G. R., & Underwood, J. R. (1977). An Intercity Terminal Access Modal Choice Model. *Transportation Planning and Technology,* **4**, 11-21.
Lee, J.-H., & Chang, J. S. (2006). *The Impact of High Speed Rail Service on Mode Share of Intercity Passenger Travel in Korea.* Paper presented at the 85th Annual Meeting of the Transportation Research Board, Washington, D.C.
Lunsford, M. E., & Gosling, G. D. (1994). *Airport Choice and Ground Access Mode Choice Models: A Review and Analysis of Selected Literature* (Working Paper UCB-ITS-WP-94-5). Berkeley, Calif.: Institute of Transportation Studies, University of California.
Lyles, R. W., & Mallick, S. R. (1990). An Intercity Modal Split Considering four Modes. *Journal of Advanced Transportation,* **24** (3), 191-212.
Machefert-Tassin, Y. (1999). La ligne nouvelle à Grande Vitesse de l'Eurotunnel à Londres (The new high speed line from the Eurotunnel to London - in French). *Revue Générale des Chemins de Fer* (3).
Mandel, B. N., Gaudry, M. J. I., & Rothengatter, W. (1993). *A Disaggregate Box-Cox Logit Mode Choice Model of Intercity Passenger Travel in Germany*. Montréal: Université de Montréal, Centre de Recherche sur les Transports.
Maxwell, R. R. (2003). Converting a large region to a multimodal pulsed-hub public transport network. *Transportation Research Record,* **1835**, 128-136.
Mayrzedt, H. (2001). Japans Autobahnen-Mautstraßen. *Internationales Verkehrswesen,* **53** (4), 137-141.

McCarthy, P. S. (1997). The Role of Captivity in Aggregate Share Models of Intercity Passenger Travel. *Journal of Transport Economics and Policy,* **31** (3), 293-308.
McCullagh, P., & Nelder, J. A. (1998). *Generalized Linear Models*. Boca Raton, Florida: Chapman & Hall/CRC.
McFadden, D. (1972). *Conditional Logit Analysis of Qualitative Choice Behavior* (Working Paper 199). Berkeley, Calif.: Institute of Urban & Regional Development, University of California.
McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (105-142). New York: Academic Press.
McFadden, D. (1978a). Modelling the Choice of Residential Location. *Transportation Research Record,* **673**, 72-78.
McFadden, D. (1978b). Modelling the Choice of Residential Location. In A. Karlquist (Ed.), *Spatial Interaction Theory and Residential Location* (75-96). Amsterdam: North-Holland Pub. Co.
Merriam Webster's Collegiate Dictionary 10th Edition. (1993). Springfield, Mass.: Merriam-Webster, Incorporated.
Mohammadian, A., & Miller, E. J. (2002). *A Comparison of the Performance of Nested Logit Models and Artificial Neural Networks for Predicting Household Automobile Choices.* Paper presented at the Transportation Research Board 81st Annual Meeting, January 13-17, 2002, Washington, D.C.
Morikawa, T. (1989). *Incorporating Stated Preference Data in Travel Demand Analysis.* Dissertation, MIT, Cambridge, Mass.
Morikawa, T., Ben-Akiva, M. E., & Yamada, K. (1991). Forecasting Intercity Rail Ridership Using Revealed Preference and Stated Preference Data. *Transportation Research Record,* **1328**, 30-35.
Morrison, S. A., & Winston, C. (1985). An Econometric Analysis of the Demand for Intercity Passenger Transportation. *Research in Transportation Economics,* **2**, 213-237.
Munizaga, M. A., Heydecker, B. G., & Ortúzar S., J. d. D. (2000). Representation of Heteroskedasticity in Discrete Choice Models. *Transportation Research,* **34B** (3), 219-240.
Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalised Linear Models. *Journal of the Royal Statistical Society, Series A,* **135**, 370-384.
Nicholson, W. (1992). *Microeconomic Theory: Basic Principles and Extensions* (5th ed.). Fort Worth, TX: The Dryden Press.
OAG Desktop Flight Guide - Worldwide Edition. (Vol. 19 No. 3)(1994). Oak Brook, IL: Official Airline Guides, a division of Reed Elsevier Inc.
Oetting, A. (1995). *Verkehrsverlagerung und Induktion im Personenfernverkehr* (Diplomarbeit). Aachen: Technical University of Aachen, Institute of Transportation Studies.
Overseas Timetable. (2000). Peterborough, UK: Thomas Cook Publishing.
Peers, J. B., & Bevilacqua, M. (1976). Structural Travel Demand Models: An Intercity Application. *Transportation Research Record,* **569**, 124-135.

Pels, E., Nijkamp, P., & Rietveld, P. (2003). Access to and competition between airports : a case study for the San Francisco Bay area. *Transportation research,* **37A** (1), 71-83.
Quandt, R. E., & Baumol, W. J. (1966). The Demand for Abstract Transport Modes: Theory and Measurement. *Journal of Regional Science,* **6** (2), 13-26.
Rassam, P. R., Ellis, R. H., & Bennett, J. C. (1970). *The N-dimensional Logit Model: Development and Application*: Peat, Marwick, Mitchell & Co.
Rassam, P. R., Ellis, R. H., & Bennett, J. C. (1971). The N-dimensional Logit Model : Development and Application. *Highway Research Record,* **369**, 135-147.
Recker, W. W. (2001). A bridge between travel demand modeling and activity-based travel analysis. *Transportation research,* **35B** (5), 481-506.
Ruud, P. A. (2000). *An Introduction to Classical Econometric Theory*. New York: Oxford University Press.
Schafer, A. (2000). Regularities in Travel Demand: An International Perspective. *Journal of Transportation and Statistics,* **3** (3), 1-31.
Schneider, J. B. (1993). Selecting and Evaluating Intermodal Stations for Intercity High Speed Ground Transportation. *Transportation Quarterly,* **47** (2), 221-245.
Sharp, A. (2005). Which Way to the Airport? *Railway Gazette International,* **161** (4), 217-219.
Sobieniak, N. D., Westin, R. B., Rosapep, T., & Shin, T. (1979). Choice of Access Mode to Intercity Terminal. *Transportation Research Record,* **728**, 47-53.
Spear, B. D. (1984). *An Analysis of the Demand for Airport Bus Services at Washington National and Dulles International Airports* (Report DOT-TSC-FAA-84-2). Cambridge, Mass.: US Department of Transportation, Transportation Systems Center.
Spector, P. (2002). *Course Notes for S100*. Berkeley, Calif.: University of California.
Stephanedes, Y. J., & Kumar, V. (1984). Comparative Evaluation of Intercity Demand Models, *Discrete Choice Models in Regional Science* (114-121). London: Pion, Ltd.
Stephanedes, Y. J., Kumar, V., & Padmanabhan, B. (1984). A Fully Disaggregate Mode Choice Model for Business Intercity Travel. *Transportation Planning and Technology,* **9** (1), 13-23.
Stopher, P. R., Metcalf, H. M. A., Wilmot, C. G., Catalina, A. J., & Schimpeler, C. C. (1999). Estimating Patronage for a Feasibility Study of High Speed Rail in Thailand. *Transportation Research Record,* **1691**, 12-18.
Stopher, P. R., & Prashker, J. N. (1976). Intercity Passenger Forecasting: The Use of Current Travel Forecasting Procedures. *Transportation Research Forum,* **XVII** (1), 66-75.
Tokaido Line to be Improved. (2003). *International Railway Journal,* **18** (5), 2.
Train, K. (1986). *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*. Cambridge, Mass.: MIT Press.
Train, K. (2003). *Discrete choice methods with simulation*. Cambridge, U.K. ; New York: Cambridge University Press.
Transit Fact Book. (1996). American Public Transit Association (APTA).
Vilmart, C., & Paix, J.-F. (1994). Les Effets des Ruptures de Charge sur la Concurrence entre l'Avion et le Train. *Revue Générale des Chemins de Fer* (10), 17-21.

Vranich, J. (1997). *Derailed : What Went Wrong and what to Do about America's Passenger Trains* (1st ed.). New York: St. Martin's Press.
Vrtic, M., & Axhausen, K. W. (2003). *The impact of tilting trains in Switzerland: A route choice model of regional- and long distance public transport trips.* Paper presented at the Transportation Research Board 82nd Annual Meeting, January 12-16, 2003, Washington, D.C.
Walker, J. L. (2001). *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables.* Dissertation, Massachusetts Institute of Technology, Cambridge, Mass.
Walker, J. L. (2002). *The Logit Kernel (or Continuous Mixed Logit) Model: Dispelling Misconceptions of Identification.* Paper presented at the Transportation Research Board 81st Annual Meeting, January 13-17, 2002, Washington, D.C.
Walther, K. (1991). *Maßnahmenreagibler Modal-Split für den städtischen Personenverkehr* (Heft 45). Aachen: Technical University of Aachen, Institute of Transportation Studies.
Wardman, M., Whelan, G. A., & Toner, J. P. (1994). *Diverting Interurban Car Users to Rail : Results from a Revealed Preference Mode Choice Model* (Working Paper 423). Leeds, England: University of Leeds, Institute for Transport Studies.
Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2003). *Statistical and Econometric Methods for Transportation Data Analysis*. Boca Raton, Florida: Chapman & Hall/CRC.
Watson, P. L. (1974). Comparison of the Model Structure and Predictive Power of Aggregate and Disaggregate Models of Intercity Mode Choice. *Transportation Research Record,* **527**, 59-65.
Wen, C.-H., & Koppelman, F. S. (2001). The Generalized Nested Logit Model. *Transportation Research,* **35B** (7), 627-641.
Westin, R. B., & Manski, C. F. (1979). Theoretical and Conceptual Developments in Demand Modeling. In D. A. Hensher & P. R. Stopher (Eds.), *Behavioural Travel Modeling*. London: Croom Helm.
Widlert, S. (1994). *Stated Preference Studies: the Design Affects the Results.* Paper presented at the 7th International Conference on Travel Behaviour, Valle Nevado, Santiago, Chile.
Widlert, S. (1998). Stated Preference Studies: the Design Affects the Results. In J. d. D. Ortúzar & D. A. Hensher & S. Jara-Díaz (Eds.), *Travel Behaviour Research : Updating the State of Play* (105-121). Oxford: Elsevier Science Ltd.
Wolinsky, J. (2000). California's Passenger Rail Success Story. *Railway Age,* **201** (9), 95-101.
Wonnacott, R. J., & Wonnacott, T. H. (1979). *Econometrics*. New York: John Wiley & Sons.
Xie, J.-R. (2000). *A neural network approach to modeling and predicting intercity passenger flows.* Dissertation, Indiana University.

Forward References

To obtain a list of papers, theses, or books which were published *after* this dissertation and cite this thesis, please go to:

thinkMetric.com/diss/fwd

If you have cited this dissertation, please send a brief email to the author, so your publication can be added to this Forward References section:

Reinhard.Clever@thinkMetric.com