CalSpeed
California High Speed Rail Series

Intercity Rail Ridership Forecasting and the Implementation of High-Speed Rail in California

Erin Vaca

Working Paper
UCTC No. 182

The University of California
Transportation Center
University of California
Berkeley, CA 94720

The University of California Transportation Center

The University of California Transportation Center (UCTC) is one of ten regional units mandated by Congress and established in Fall 1988 to support research, education, and training in surface transportation. The UC Center serves federal Region IX and is supported by matching grants from the U.S. Department of Transportation, the California Department of Transportation (Caltrans), and the University.

Based on the Berkeley Campus, UCTC draws upon existing capabilities and resources of the Institutes of Transportation Studies at Berkeley, Davis, Irvine, and Los Angeles; the Institute of Urban and Regional Development at Berkeley; and several academic departments at the Berkeley, Davis, Irvine, and Los Angeles campuses. Faculty and students on other University of California campuses may participate in Center activities. Researchers at other universities within the region also have opportunities to collaborate with UC faculty on selected studies.

UCTC's educational and research programs are focused on strategic planning for improving metropolitan accessibility, with emphasis on the special conditions in Region IX. Particular attention is directed to strategies for using transportation as an instrument of economic development, while also accommodating to the region's persistent expansion and while maintaining and enhancing the quality of life there.

The Center distributes reports on its research in working papers, monographs, and in reprints of published articles. It also publishes *Access*, a magazine presenting summaries of selected studies. For a list of publications in print, write to the address below.

University of California Transportation Center

108 Naval Architecture Building
Berkeley, California 94720
Tel: 510/643-7378
FAX: 510/643-5456

The contents of this report reflect the views of the author who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the U.S. Department of Transportation. This report does not constitute a standard, specification, or regulation.

Intercity Rail Ridership Forecasting and the Implementation of High-Speed Rail in California

Erin Vaca

Institute of Urban and Regional Development
University of California at Berkeley
Berkeley, CA 94720

CALIFORNIA HIGH SPEED RAIL SERIES

Working Paper
May 1993

UCTC No. 182

The University of California Transportation Center
University of California at Berkeley

Table of Contents

List of Figures

List of Tables

PREFACE

This is one of a series of reports now being published as the output of IURD's study of the potential for a high-speed passenger train service in California. The present series includes twelve studies. This is the twelfth of twelve studies, eleven of which have already been published.

We gratefully acknowledge the support provided by the United States Department of Transportation and the California Department of Transportation [CALTRANS] through the University of California Transportation Center. Of course, any errors of fact or interpretation should be assigned to us and not to our sponsors.

PETER HALL

Principal Investigator

AUTHOR'S NOTE

The CalSpeed project is a study of the potential for a high-speed passenger rail service in California. The work published so far in the CalSpeed series includes seven assessments of high-speed ground transportation technologies along with *High-Speed Trains for California; Strategic Choice: Comparison of Technologies and Choice of Route* (UCTC No. 104), which treats choice of technology and route. A follow-up working paper on the likely capital costs of a high-speed rail project, *The Cost Escalation of Rail Projects: Using Previous Experience to Reevaluate the CalSpeed Estimates*, was published in April 1993.

The second phase of the project will primarily examine the market potential for high-speed rail service in California. This paper critically examines past rail ridership forecasting exercises as preparation for approaching the CalSpeed demand forecasting project. Also now published as part of the second phase of the Calspeed project is *The Development Effects of High-Speed Rail Stations and Implications for California*. The final paper in the series will contain an analysis of the California intercity travel market and estimate the potential ridership for high-speed trains in California.

The support provided by the California Department of Transportation (CALTRANS) and the University of California Transportation Center is gratefully acknowledged.

INTRODUCTION AND BACKGROUND

> . . . transit systems, power plants, hospitals and airports are constructed only after forecasts have demonstrated that a "need" exists for their services, and that their costs are justified by expected benefits.[1]

This working paper is the first in the CalSpeed series to address the question of the market potential for high-speed passenger rail service in California. Along with construction cost, forecast ridership is the second significant question asked about any proposed high-speed rail system. Indeed, the decision to construct such systems is often portrayed as contingent on such forecasts. For this reason, an examination of the rail ridership forecasting experience is a worthwhile preliminary to the CalSpeed ridership estimates.

The first part of the paper focuses on the high-speed rail systems in Europe and Japan and the role that ridership forecasts played in the decision to build the original high-speed rail lines. The European and Japanese situation is then compared to the situation California faces today. The next part draws an analogy between the decision to construct the second-generation rail transit systems in the United States and the decision to construct high-speed rail in California. This part reviews criticism of rail transit ridership forecasting and examines the original bay Area Rapid Transit (BART) forecasts as an illustration of the difficulties involved in forecasting for new modes.

Throughout the paper, questions of accuracy and sources of error are addressed as thoroughly as possible. What does the accuracy of past ridership forecasts say about forecasting as a decision-making tool in the public or private arenas? How might the rail forecasting experience benefit a California study? The paper will address these questions and others, and concludes with a discussion of the role that ridership forecasting might play in the decision to build a California high-speed rail system.

PART I: RIDERSHIP FORECASTING AND THE DECISION TO BUILD IN JAPAN AND EUROPE

JAPANESE SHINKANSEN

Tokaido/Sanyo

The impetus for construction of the Shinkansen (Figure 1) or "New Trunk Line" between Tokyo and Osaka was a need for increased capacity.[2] In its first fiscal year of operation (1965), the Tokaido Shinkansen carried over 30 million passengers and today the line is close to capacity. The future demand for travel on this corridor (based on population trends) was thought to be so large that Japan National Railways (JNR) simply concentrated on building the highest capacity system possible. The population, density, and geographical configuration along the Tokaido line make this a unique transport corridor in the world, one where a high-capacity transport mode needed little prior analytical justification.

Tohoku and Joetsu Lines and Future Extensions

The "Law for the Construction of Nation-wide Shinkansen Railway Network," enacted in 1970, was ". . . intended to expand the nation-wide railway network by means of the Shinkansen railway system and thus serve to develop national economy . . . for the purpose of a comprehensive and universal development of the national land."[3] The Sanyo Shinkansen, extending from Osaka to Hakata, was completed in 1975 and the Tohoku and Joetsu lines were completed between 1982 and 1985.

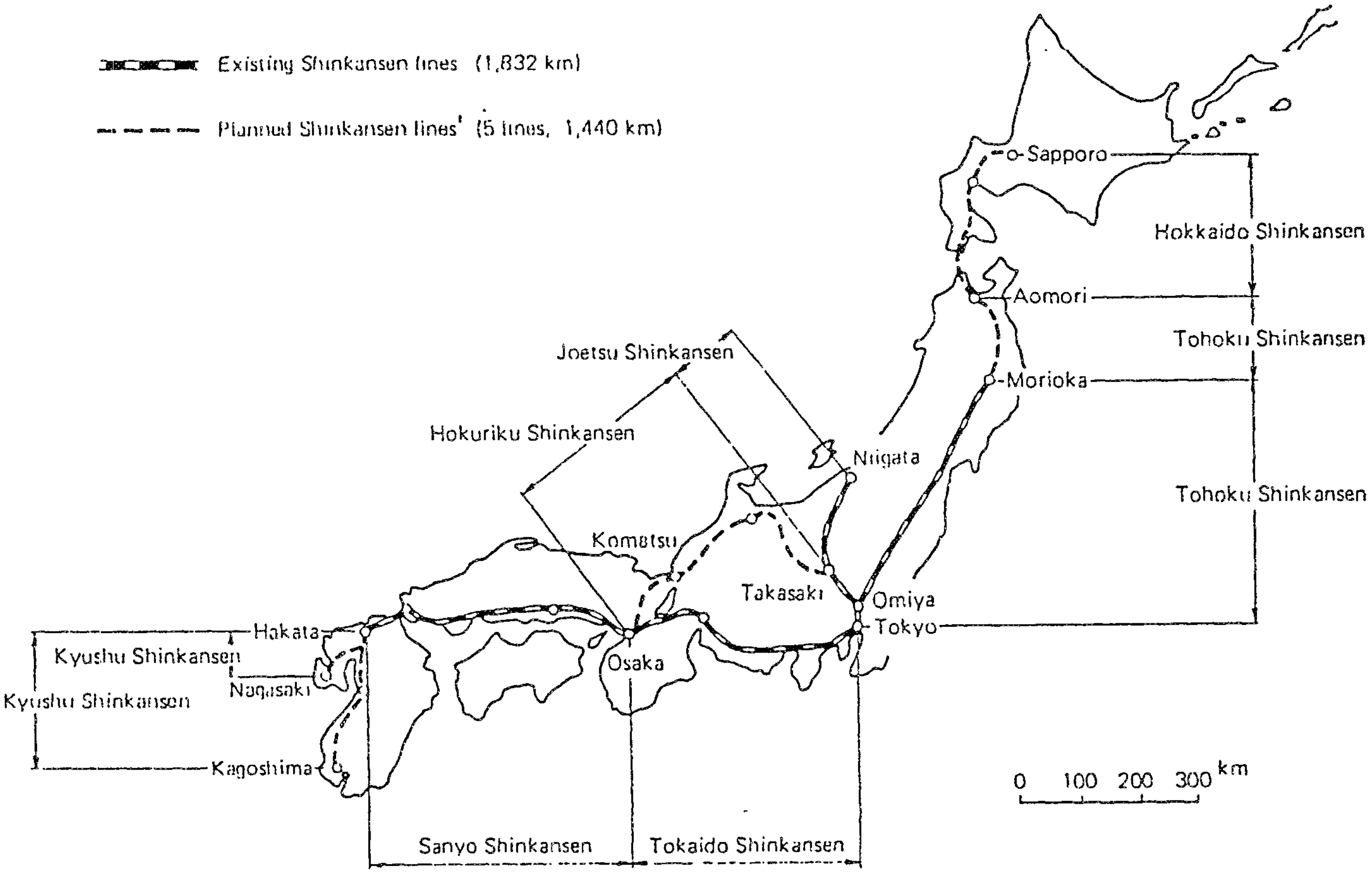
JNR was privatized in 1987 and was reorganized into six separate companies. Because these new companies did not keep records from the time before their reorganization, information on the ridership forecasts made before construction of these lines is difficult to find. However, it is clear that the newer lines served relatively fewer passengers than did the original Tokaido line (Table 1). Thus, the decision to build these lines may have been predicated more on such national policy goals as stimulating the economy and spreading development more evenly over the country than on a clear-cut transportation demand need.

The case for further extensions of the Shinkansen system appears to have been even less certain. Table 1a shows the forecast ridership for each of four proposed extensions and their respective lengths. Although there was found mention of a "Modal Demand Model" capable of reflecting a big change in level of service and the resulting induced travel, no details on this model were provided.

An article in *Japanese Railway Engineering* discusses the importance of reducing capital costs on these lines because, "The planned Shinkansen lines cannot expect a large volume of ridership."[4] Given that the newly organized Japanese railway companies do not wish to destroy their balance sheets with commercially unprofitable lines, why and how would construction of these lines come about? The answer is that the railway companies will build the lines only if the infra

Figure 1

The Japanese Shinkansen



Source: Minemoto, 1987.

Table 1
Ridership and Length of Shinkansen Lines

	Tokaido (1964)	Sanyo (1972-75)	Tohoku (1984)	Joetsu (1982)
1985 Ridership (daily round trip passengers)	152,000	58,000	42,000	29,000
Route Length (km)	515	398	646	275
Passengers/km	295	145	65	105

Source: Minemoto and Taniguchi.

Table 1a
Extension of Shinkansen

	Kyushu (Hakata-Hokuriku	Kyushu (Tosu-Kagoshima)	Nagasaki)	Hokkaido
Forecast 1995 (daily round trip passengers)	25,200	12,300	11,000	10,800
Route Length (km)	594	249	140	282
Passengers/km	42	49	78	38

Source: Minemoto and Taniguchi.

structure costs are publicly subsidized. With this subsidy, the railway companies will find their overall profit greater than it would have been without construction of the new lines.

A Japanese research institute calculated an overall economic benefit to Japan exceeding the cost of servicing the Shinkansen bond debt that would be incurred. Even so, construction of these lines has yet to begin, and "the planned Shinkansen lines include extremely political elements . . . which cannot be measured merely by economic rationality."[5] In all probability, ridership forecasts will not be the decisive factor.

THE FRENCH TRAIN à GRANDE VITESSE (TGV)

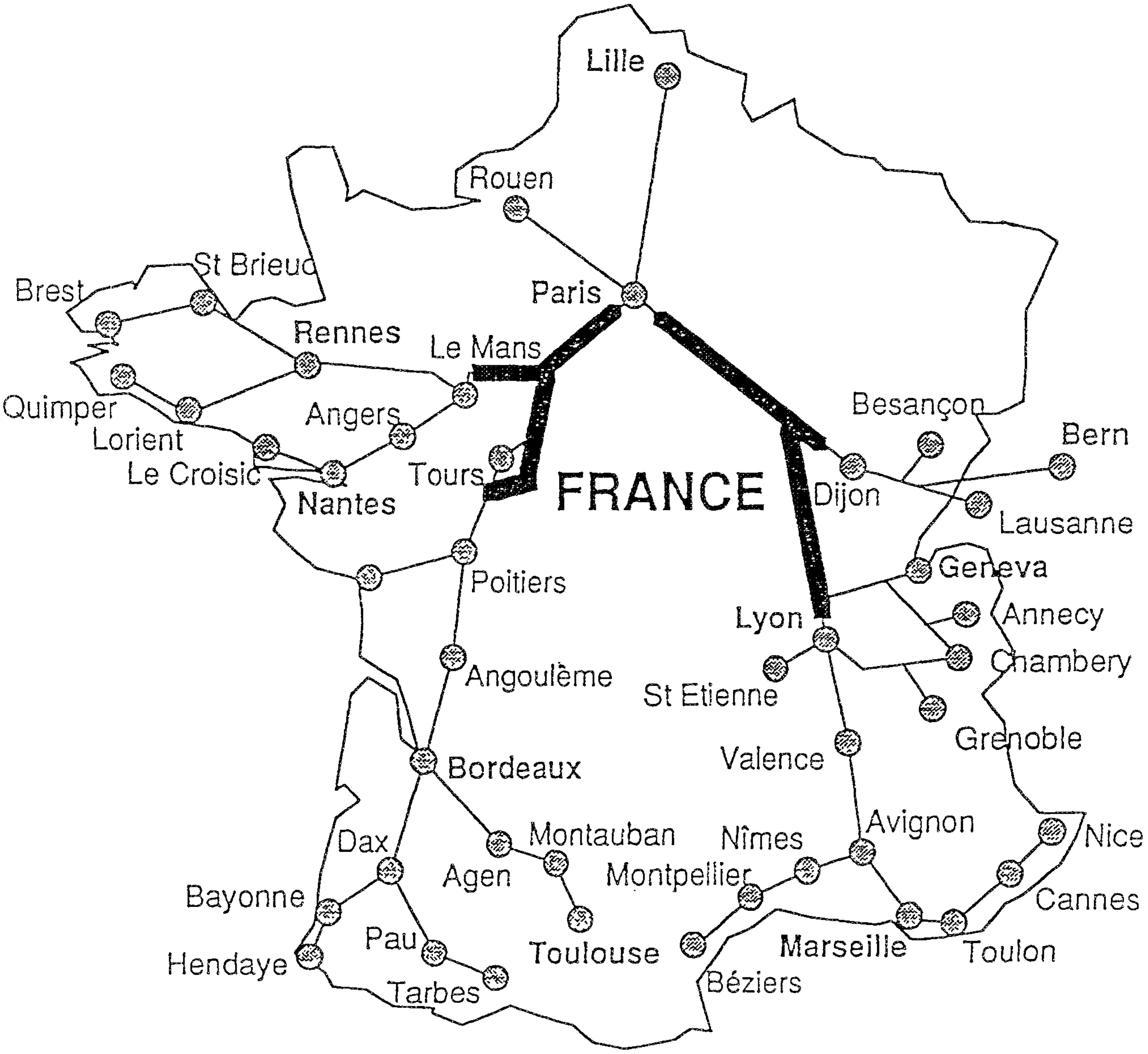
TGV Sud Est (TGV-SE)

As with the Japanese Shinkansen, the fundamental impetus for developing the first TGV[6] line linking Paris to Lyon (see Figure 2) was a shortage of capacity on the existing conventional rail lines. While the demand for travel in this corridor was apparent from the increasing automobile and air traffic, the French National Railroads (SNCF) foresaw that the maximum possible frequency of trains on the line would soon be reached. At the same time, the improvement and expansion of the French airport and highway network made these modes increasingly competitive. The SNCF studied various options for resolving these problems, including increasing capacity on the existing existing conventional line. After much analysis, construction of a new, high-capacity, high-speed passenger line was found to best satisfy the technical, operational, and economic criteria.[7]

Figure 2

TGV Network in 1992

436 miles of new High Speed Line
(Total *TGV* network: 2,920 miles)



Source: Minemoto, 1987.

Projecting existing trends, SNCF predicted that ridership in the southeast corridor would stagnate at about 12.2 million annual revenue passengers without construction of the TGV line (Table 2). Then, for the period 1981 to 1984, SNCF designated the difference between this expected baseline and actual traffic (both TGV and conventional) as "new" rail traffic generated by the TGV line. Service on the TGV-SE began in 1981 with 5.9 million new southeast corridor passengers forecast for 1985. As Table 2 shows, forecasts for TGV-induced rail ridership in the southeast corridor were considerably *under* the ridership actually attained. Comparisons of city pair forecasts to actual ridership show that ridership was underpredicted in 11 of 14 city pairs (Table 2a).

SNCF produced these forecasts by first modelling the "reference situation" (defined as the southeast corridor rail traffic without any intervention such as introduction of the TGV) as a Cobb-Douglas production function[8] of consumption, population, and an index of wealth in the region (Equation 1, Figure 3). New rail traffic was then assumed to come from three sources: traffic diverted from the airlines, traffic diverted from the highways, and traffic induced by the introduction of the TGV.

The diversion from air travel to rail was forecast with a "price-time" model, which relies on the assumption that people will choose the mode that presents the lowest generalized cost. Generalized cost (Equation 2) is defined as the sum of journey fare and travel time multiplied by the individual's value of time. The values of time within the population were assumed to be distributed log-normally and the travel market was thus divided between rail and air.

Because of uncertainty about the origins and destinations of road traffic, traffic diverted from the highways was not modelled explicitly but accounted for in estimates of induced traffic (new trips generated by the availability of the TGV). A gravity model (Equation 3), with population indicators of wealth as the attraction factors and generalized cost as the impeding factor, was used to predict traffic induced by the TGV.

The introduction of the TGV would change the generalized cost of travel, and the resulting change in traffic is shown in Equation 4. Adding the induced traffic to that diverted from airlines produced the overall estimate of new rail traffic attributed to the TGV.

Some shortcomings are apparent in the approach taken by SNCF to predict rail traffic attributable to the TGV. For example, competition was assumed between only first-class rail travel and air. In addition, diversion from auto travel was not adequately treated. The actual TGV traffic attained in 1985 was 7.0 million, a result "even more remarkable," according to SNCF, because of the poor economic conditions at the time.[9] Had the poorer economic conditions been used in the models, the forecasts would likely have been even further off the mark.

TGV Atlantique

The TGV-A service (Figure 2), serving the west and southwest of France from Paris, began in 1989 on the west branch. Exceeding the forecast of 1.5 million passengers, more than 1.6 million

Table 2
TGV-SE and Conventional Southeast Corridor Traffic
(millions of passengers)

	1981	1982	1983	1984	1985
Predicted SE Corridor Traffic Without TGV	12.2	12.2	12.2	12.2	12.2
Actual SE Corridor Traffic	12.7	14.3	15.7	18.4	19.2
Actual TGV Traffic	1.2	6.0	9.0	13.3	14.7
"New" Traffic	0.5	2.1	3.5	6.2	7.0
Forecast of New Ridership Made in 1980	-	-	-	-	5.9

Source: Michel Leboeuf, SNCF

Table 2a
Actual and Forecast Southeast Corridor Rail Ridership by City Pair

Connections from Paris to	**Actual** winter 1983-84/ winter 1980-81 ratio	**Forecast** winter 1983-84/ winter 1980-81 ratio
Annency	1.6	1.3
Avignon	2.0	1.4
Besançon	1.9	1.1
Bourg	2.0	1.4
Chambéry	1.6	1.6
Dijon	1.2	1.1
Grenoble	1.4	1.4
Lyon	2.4	2.2
Mâcon	1.7	1.3
Marseille	1.4	1.2
Montpellier	2.0	1.4
Nîmes	2.9	1.5
Saint Etienne	1.9	1.5
Valence	1.9	1.7

Source: Arduin

Figure 3

SNCF Models for the SE Corridor

Southeast Corridor Reference Demand

$\text{Traffic}_{Mt} = KC_t^{c}F_t^{f}P_t^{p}W_t^{w}$ *(1)*

where T is traffic on mode M,
t = time subscript,
K = a constant,
C = consumption of families, c = elasticity of traffic to consumption,
F = fare, f = elasticity of traffic to fare,
P = population, p = elasticity of traffic to population,
W = index of wealth, w = elasticity of traffic to wealth.

Air to Rail Diversion

$C_{gm}^{k} = F_m + h_k T_m$ *(2)*

where C_{gm}^{k} is the generalized cost for passenger k on mode m,
F_m = fare on mode m,
h_k = passenger k.s value of time,
T_m = journey time on mode m.

Induced Traffic

$$T_{ij} = \frac{kP_iP_jW_iW_j}{Cg_{ij}^{\gamma}}$$ *(3)*

where T_{ij} is the traffic between two regions, i and j,
W_i and W_j are indicators of wealth in the respective regions,
Cg = the generalized cost travelling by the mode in question,
) = elasticity of traffic to generalized cost,
k = parameter of adjustment

Change in Traffic

$$\frac{\partial T_{ij}}{T_{ij}} = \frac{\partial Cg_{ij}}{Cg_{ij}}$$ *(4)*

where an average price and value of time are used in Cg_{ij}

Source: Arduin.

passengers rode the line between September 24 and December 31, 1989. The south-west branch of the Atlantique opened in 1990 and 3.8 billion passenger kilometers were travelled on the TGV-A network that year.[10] Because there are no published forecasts of ridership for these early phases of TGV-A service, however, it is difficult to test the assertion that the line has been a "smashing success."[11]

An early forecast for the TGV-A service called for a ridership of 10.2 million passengers in 1992,[12] a ridership that was attained by 1991 with 10.492 million revenue passengers. Note that 1992 was the first fully operational year of TGV-A service. New services have been added, new stations brought on line, and electrification extended continually over this period, so the line has not yet reached its equilibrium ridership.

Because of their proprietary nature, it is not possible to examine the models used to predict TGV-A ridership. Nevertheless, it is important to note that nearly a third of the French population lives in the regions that were to be served by the TGV-A, and that the SNCF could count on the well-established conventional train ridership of 17 million passengers in 1985 as a base for TGV-A ridership. In a 1986 article, SNCF estimated that traffic on the TGV-A would grow to 21 million annual passengers by the early 1990s with 68 percent coming from conventional trains and 32 percent new ridership diverted from air and auto modes.[13]

For many years the marketing strategy of SNCF has been to maintain its existing rail ridership base and to attract new riders.[14] Without construction of the TGV-A, SNCF would have continued to lose passengers to air and highway in the Atlantique corridors. Even so, the TGV-A is unlikely to see the phenomenal growth in ridership found on the southeast line both because of less robust economic conditions and because the Atlantique corridors were not suffering the same degree of saturation problems.

Since ". . . only clearly profitable projects may be borne by the company alone,"[15] the TGV-A received a 30 percent infrastructure grant from the French government (in contrast to the TGV-SE line, which was entirely self-financed by SNCF with funds from the international capital market). Thus, the decision to construct the Atlantique lines seems to have been less clear-cut and may have involved consideration of social and economic impacts, as well as an element of regional prestige.

THE GERMAN INTERCITY EXPRESS (ICE)

The development of Germany's high-speed Intercity Express (ICE) service came in response to the declining share of the transportation market that rail faced in the 1960s, combined with capacity problems on the nation's highway, air, *and* rail networks. Part of the response was a national policy decision to shift as much passenger and freight traffic as possible onto the rail mode.

Germany's high-speed rail infrastructure was first conceived as a way to increase rail capacity and improve comfort levels. Only later, with the achievements of Japan and France apparent and with increasing intermodal competition, did the development of a high-speed rolling stock begin.

Currently, ICE service operates on two routes between Hamburg and Munich as well as service to Basel via Mannheim (Figure 4). Future service plans include international connections such as the Paris-Brussels-Köln-Amsterdam service expected by 1997.[16]

Significantly, the decision to build the new infrastructure was not based upon any feasibility study or demand projection showing that the service and capacity were needed. Rather, it was more of a policy decision to meet an obviously growing transportation demand with rail. To shift traffic, additional rail capacity was needed. Later, higher speeds and greater comfort were recognized as being necessary to implement this policy.

Since the ICE service replaced certain trains of a previously existing premium service, there was no need to verify the existence of demand for these trains. Deutsche Bundesbahn (DB), the German national railway, roughly estimated that the increased speed and comfort would cause rail ridership to rise by about 30 percent while decreasing the number of air passengers. Lufthansa did, in fact, experience a 50 percent drop in ridership on the Hannover-Frankfurt corridor once the ICE was introduced, and the rail ridership in fact increased by about 30 percent.[17]

CONTRASTS BETWEEN CALIFORNIA AND THE FOREIGN EXPERIENCE

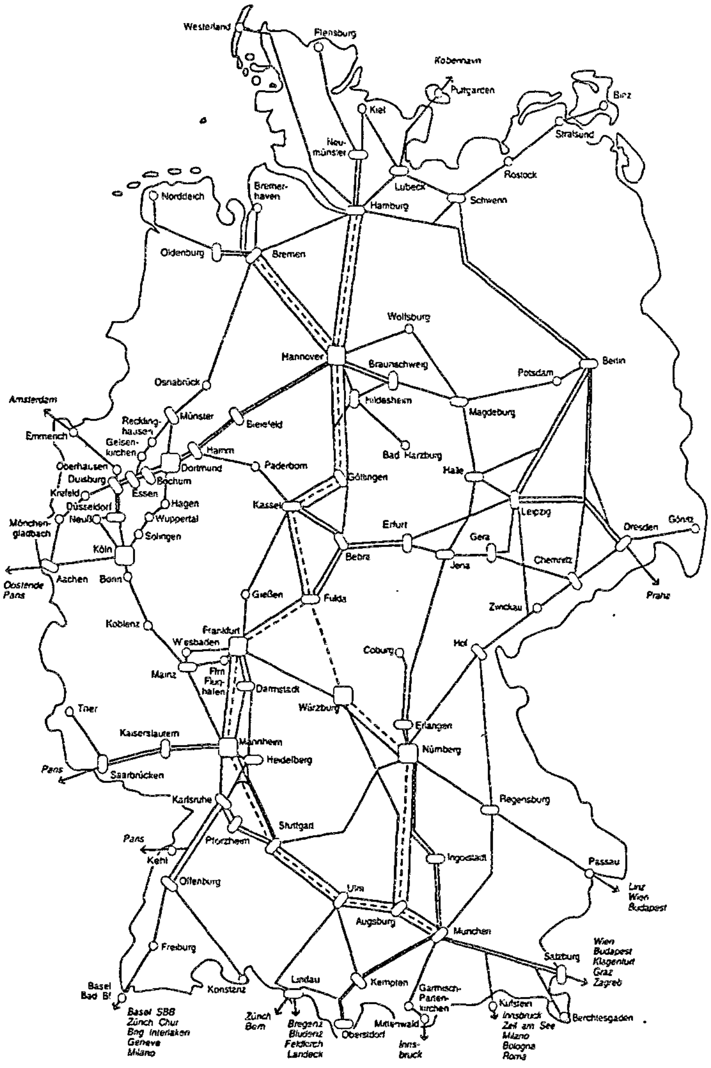
Although the European and Japanese high-speed rail systems have been in operation for as long as 20 years, relatively little information is available on the ridership forecasting techniques employed before their construction. In the case where the most information is available, the forecasts under-predicted ridership (although, to be fair, such forecasts are often purposely conservative). Apparently, many of the European and Japanese investments relied less on ridership forecasts than on policy goals for justification.

Partly because the demand for travel in these first high-speed rail corridors was so obviously high, and partly because of an already existing base of rail travellers, ridership forecasting was not critical to establish a need or to justify the public investment. The situation in California today is very different in that there is no obvious need for increased passenger rail capacity; indeed, rail barely exists as an intercity transport mode. California is thus faced with the much more difficult problem of forecasting demand for an entirely new mode and a more radical transformation of its transportation market.

In addition, the European and Japanese systems were largely publicly financed or self-financed by national railway companies. By contrast, the U.S. National Passenger Railway Corporation (Amtrak) has no mandate to undertake an intrastate project such as a California high-speed rail network and does not enjoy a position of influence on overall national transportation policy. Since the U.S. federal government currently looks to private investment as the main source of high-speed rail financing, accurate ridership estimates at the early decision stage will be more critical in California.

Figure 4

The German Inter-City Express



Streckennetz Fernverkehr

Zeichenerklärung

ICE InterCityExpress

EC EuroCity Europäischer Qualitätszug .

IC InterCity Nationaler Qualitätszug

IR InterRegio, überregionaler Zug mit gehobenem Komfort

Wichtiger Hinweis
Diese Übersichtskarte zeigt nur die großen Bahnhöfe

Informationen über die weiteren Haltebahnhöfe und Fahrpläne erhalten
Sie bei allen Fahrkartenausgaben. Oder unter Telefon 1 94 19

Stand Januar 1992

- - - - ICE Linie

══ EC/IC Linie

—— IR Linie

Part II: Rail Ridership Forecasting in the United States.

THE NORTHEAST CORRIDOR

In 1976, the U.S. Congress mandated what eventually amounted to a $2.19 billion investment in the nation's northeast rail corridor, called the Northeast Corridor Improvement Project (NECIP) (Figure 5). The Railroad Revitalization and Regulatory Reform Act of 1976 was spurred by factors similar to those found in Europe: growing congestion on a transportation corridor and an inability to continue competing with other intercity modes. Although the conditions in the U.S. were more extreme than in Europe —Amtrak had inherited a seriously deteriorated capital plant from the Pennsylvania Railroad —the response to these conditions was less decisive.

In the early 1970s, NECIP planners set ambitious service goals of 150 mph top speed and non-stop travel times of 2 hours between New York City and Washington, D.C., and 2 hours, 45 minutes, between New York and Boston. A Baseline Implementation Master Plan was drawn up in 1977 to meet these goals, along with ridership estimates. Because of budgetary pressures, however, the Northeast Corridor Improvement Project was scaled back and its initial service goals never met.

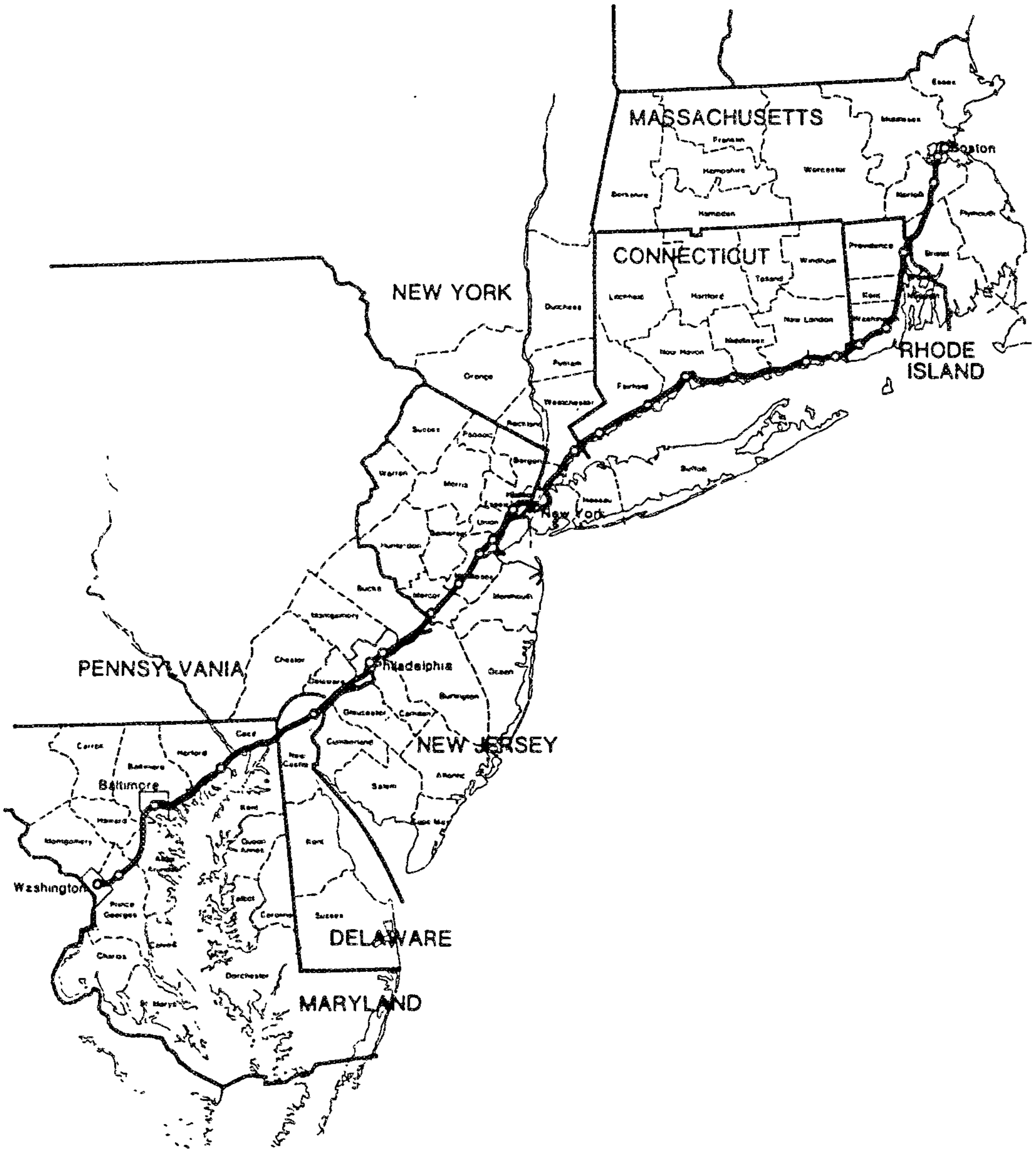
Initial justification for the project was found in forecasts of increasing travel demand, growing congestion in the corridor, and the potential environmental benefits of rail. By the mid-1980s, however, the Northeast Corridor had not yet attracted the ridership foreseen in the 1960s and 1970s. For example, 1977 models developed for the Federal Railroad Administration (FRA) predicted 14.7 million passengers over the completed mainline in 1982. Actual ridership on a line under construction was 9.5 million (see Table 3). By 1984, with trip time improvements still not fully operational, ridership grew to 9.8 million (two-thirds of the original FRA projection for 1982).[18] (See Table 3.)

Consultants working for the FRA employed a two-step procedure to forecast the ridership for an improved Northeast Corridor. First, gravity models calibrated with historical patronage data from 1960 and 1972 predicted total intercity travel demand between city pairs (Figure 6). A mode split model calibrated on the 1972 Census of Transportation survey data was then applied to predict the share of unimproved and improved rail in 1982 and 1990.[19] The two models were linked by using a measure of transportation system quality generated by the mode split model in the total demand model.

Intercity travel mode choice was treated disaggregately by generating individual travellers from probability distributions of key attributes. These attributes included city of residence, trip purpose, income, trip duration, local origin and destination zones, travel period, "time value" (a function of the traveller's income and trip purpose), party size, "preference factor" for each alternative mode, auto availability, and sensitivity to mode frequency. The probability of a trip originating or terminating in a specific zone was a function of the population, income, employment, and hotel

Figure 5

The Northeast Corridor



Source: FRA (1978).

Table 3
NEC Forecast and Actual Ridership (millions of person trips)

1982 Forecast	1984 Actual
14.7[1]	9.8[2]

Sources: 1) FRA (1978).
2) FRA (1976).

Figure 6

Northeast Corridor Total Demand Model

$$TD = CF(13.4541 \times IP^{1.3872} \times ECF^{1.3843} \times ADP^{.4910})$$

where TD = City pair total demand (annual person-trips in thousands)
IP = City pair income product, a proxy income variable
ECF = City pair economic cost function — total money and time cost of a trip computed as a composite of all modes
ADP = Alternate destination population
CF = City pair correction factor.

rooms in each zone. After the individual traveller's attributes were generated, the generalized costs of the trip by each possible mode and path were calculated and the individual was assigned to the mode with the lowest "effective perceived cost." This effective perceived cost included out-of-pocket cost, door-to-door trip time, and traveller preferences.

Notably, the estimate of total market demand was revised downward between the draft and final Environmental Impact Statements —from a total of 131 million trips to 117 million trips —to reflect changes in population and income trends. This raises the possibility that the revised estimate was still too high and that part of the forecasting error was due to an overestimating of the total market.

Actual ridership fell short of expectations for several additional reasons. One was that the delays and changes in scope of the NECIP resulted in a less attractive service than had originally been planned. Reliability and service suffered dramatically during the construction years and a negative public opinion took years to turn around. Also, the 1977 models did not anticipate a deregulated air industry with cut-throat low fares and cut-rate carriers like People Express. For these reasons, a 1986 assessment of the NECIP published by the FRA concludes that, "The

Table 4

Demand Model Trip-Time Assumptions Versus Reality

City Pair	Assumed for 1982	Actual Best 1984
Washington-New York City	2:40	2:49
New York City-Boston	3:40	4:09

Source: FRA (1986).

traditional line of reasoning for further NEC investment has not yet had the opportunity to be either validated or disproved by the facts. In effect, therefore, it remains to be seen whether rail can ultimately attract the sizeable patronage foreseen during the lengthy studies preceding the NECIP. . . ."[20]

THE RAIL TRANSIT EXPERIENCE

The urban rail transit systems built since the 1960s are relevant to this discussion because they provide the only opportunity to examine ridership forecasting attempts for new modes in the United States. These systems can be considered new modes since most of the original rail transit systems in American cities were dismantled by the 1950s. Between the demise of the old systems and the construction of new ones, the public lost the habit and memory of travelling by rail and became used to the automobile. Some of the new systems (the San Francisco Bay Area Rapid Transit [BART] and the Washington, D.C., Metro, for example) were intended to provide a much higher level of service than the previous generation of rail transit.

Significantly, ridership forecasts were offered publicly as objective justifications for building the systems or choosing among alternatives. Most of the ridership estimates made in planning the "second generation" of rail transit systems, however, were less than successful for a variety of reasons. Unrealistic expectations about system performance, errors in forecasting socio-economic and demographic variables, and model misspecification and misuse are some of the sources of error.

Nonetheless, there is an analogue between the introduction of the second generation rapid transit systems and high-speed intercity rail. High-speed rail in California would be a re-introduction of a mode that has, for all practical purposes, ceased to exist. Its expected attributes (such as speed and comfort) would be a radical improvement over the existing intercity rail mode. Just as Californians lost the habit of making urban trips by rail transit, they have lost the habit of making intercity trips by rail. Finally, high-speed rail would again put decision-makers in the position of deciding on a large-scale public investment since even an entirely privately financed high-speed rail system would entail enormous public sector cooperation.

Forecast vs. Actual Ridership of Rail Transit in the United States

One of the main sources of information on rail transit forecasting and estimating is research conducted by Don Pickrell for the Office of Grants Management of the Urban Mass Transportation Administration. Pickrell's work has received considerable attention from both advocates and adversaries of mass transit because he asserts that the decisions to undertake the new urban rail transit projects were based on faulty information. In particular, according to Pickrell, ridership estimates were consistently too high while the construction cost estimates were too low. Had better forecasts been available, he argues, many rail transit projects would never have been built.

Pickrell's work is not concerned with ridership forecasts made *after* the decision to build. In most cases these estimates have been more accurate because they can be based on observed rail ridership trends and people's revealed preferences. Comparing a variety of rail transit projects, Pickrell found actual ridership between 28 percent and 85 percent lower than forecasts made prior to the decision to build (Table 5).

Table 5

Forecast vs. Actual Rail Transit Ridership

	Washington	Atlanta	Baltimore	Miami	Buffalo	Pittsburgh	Portland	Sacramento	Miami	Detroit
	Year to Which Data Reported in this Table Apply									
Forecast	1977	1978	1980	1985	1995	1985	1990	2000	1985	1985
Actual	1986	1987	1987	1988	1989	1989	1989	1989	1988	1988
	Weekday Rail Passengers (thousands)									
Forecast	569.6	NF	103.0	239.9	92.0	90.5	42.5	50.0	41.0	67.7
Actual	411.6	184.5	42.6	35.4	29.2	30.6	19.7	14.4	10.8	11.3
% Diff.	-28%	--	-59%	-85%	-68%	-66%	-54%	-71%	-74%	-83%

Source: Pickrell, 1990.

Less than half the error in the ridership forecasts was typically attributed to errors in model inputs (for example, demographic variables, transit service levels, and auto costs). One significant source of error was overly optimistic assumptions about the achievable frequency and speed of the rail services. In addition to errors attributable to model inputs, however, Pickrell claims a substantial amount of unexplained error, a finding that would suggest problems with model structure, model application, or interpretation of results.

Some problems with the comparisons of forecast versus actual ridership presented above are apparent. In some cases, the difference between planned and actual completion dates complicates the assessments. In three cases, forecasts for horizon years not yet reached were compared with actual ridership data (1989 ridership data compared to a 1995 forecast, for example). Pickrell cautions comparisons made for these three systems since a transit "equilibrium" has not yet been reached. In other cases, actual ridership did not begin until well after the forecast year so changes in demographic and economic variables cloud the issue.

The measures compared were average weekday passengers, total transit ridership, and change in total transit ridership. Since, ". . . a new facility's forecast and actual effects on total transit ridership are both difficult to isolate," and ". . . it is impossible to develop a precisely comparable 'with versus without' measure of the actual impact of each rail project. . .,"[21] it is perhaps unfair to criticize the forecasts on the second measure as Pickrell does.

In attempting to assess the overall impact of input errors (such as demographic, economic, transit level-of-service, and auto costs), Pickrell assumed a constant elasticity of demand value for each variable and applied these values across all the models. The percentage error of each forecast variable was multiplied by these elasticities. Perhaps a more useful analysis would have been to study how each variable entered each model and how input errors would affect the output on a case-by-case basis. All the rail projects may have employed similar modelling procedures; no explicit information is given in the report. Thus, because ". . . transit ridership elasticities implied by the models employed were generally not reported explicitly . . . the procedure used here [in Pickrell's report] was forced to rely on published estimates of transit demand elasticities."[22] Although the elasticity values were chosen at the upper range of commonly published values so as to maximize their contribution to forecasting error, without studying each model structure in detail one cannot definitely attribute a certain percentage of error to inputs and a certain percentage to model structure, application, or interpretation.

Given the ambiguities in the analysis in *Urban Rail Transit Projects*,[23] the basic accuracy of rail transit ridership forecasting remains an open question. Still, some lessons may be drawn from this study, and Pickrell recommends several steps to improve the accuracy of ridership forecasts. Most of these steps are common practice in the transportation planning field today and are quite similar to recommendations published by the High Speed Rail Association. They bear repeating for emphasis:

- Use as near as possible "horizon" year in ridership forecasts. The closer the year, the less uncertainty there is over economic and demographic variables.
- Conduct sensitivity analyses to examine the effects of alternative assumptions about inputs.
- Subject forecasts to expert review.
- Acknowledge uncertainty in ridership forecasts.

Urban Rail Transit Projects also illustrates the difficulty in predicting the impacts of a new mode on the overall transportation market before its introduction and the difficulty in sorting out such effects after its construction. Pickrell is correct in pointing out that the success and value of a mass transit system depends not only on the sheer magnitude of its ridership but also on its effect across the *total* transportation market. Will a new mode simply divert existing travellers or will it bring about an increase in the total amount of travel? Would this be a desirable outcome? Unfortunately, questions such as these are among the most troublesome to answer.

A California Example: The BART System

In 1976, the San Francisco Bay Area Rapid Transit (BART) system (Figure 7) carried only 51 percent of the passengers that were forecast when the decision was taken to build the system. A 1962 forecast predicted 258,496 average weekday passengers by 1975. Actual ridership in 1975 was 131,370 average weekday passengers. Sixty-five percent of BART's ridership was expected to divert from cars; in actuality, only 35 percent did. In fact, the construction of BART actually induced an increase in total transbay travel.[24]

The ridership estimates made in 1962 assumed that travellers would divert from other modes to BART solely on the basis of travel times. Existing travel patterns based on 1947 and 1954 surveys were updated to 1959 levels using traffic growth factors developed from cordon counts. A 1975 travel projection was made using trip attraction and generation analysis. Finally, diversion curves from other modes to BART were developed as a ratio of the door-to-door travel times. However, whether or not the analysts had accurately measured travel patterns in the Bay Area, at least two factors threw their forecasts off track.

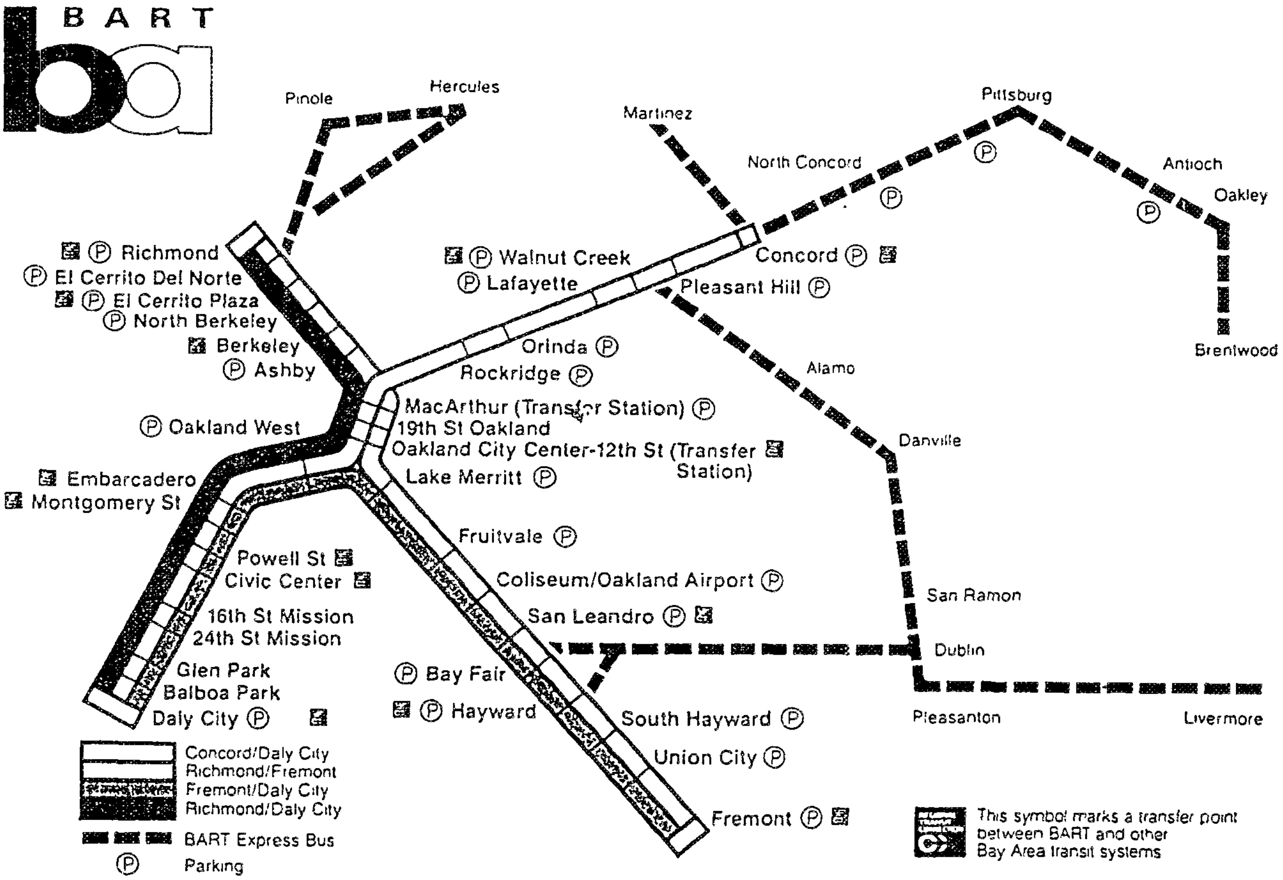
First, BART's planners had envisioned a space-age system of automatic train control with average speeds of 50 miles per hour and headways between trains of 90 seconds. BART was untried technology, however, in the sense that its components had never before all been used together, its systems engineering was neglected, and contracts were given to inexperienced firms. After a series of mechanical problems, BART opened with an archaic manual block control system in 1972 and was prone to breakdowns in its early years. Frequency and reliability suffered. Not until recent years has BART performance and reliability come close to the system originally envisioned (see Table 6). Undoubtedly, the negative image caused by BART's poor early performance curtailed ridership not only in early years but in the years after some of the problems had been corrected.

A more fundamental error in the 1962 forecast was the failure to account for other factors that influence travel decisions. For instance, the out-of-pocket nature of BART fares (versus automobile costs, which are not as apparent to most travellers), the convenience of having one's car in the middle of the day, and the nuisance of walking or taking a bus to BART stations are all factors that would tend to work against rapid transit usage but were not considered. "From the beginning, BART's planners were handicapped, because the state of transport-choice theory was so inadequate that it was impossible to simulate accurately what would happen if BART were built. They did not even have adequate descriptive data showing how people choose among travel modes. . . . However cautious the disclaimers that were attached to the forecasts, once in print the numbers somehow became reified, then accepted as facts by political leaders, voters, and bond buyers."[25]

Transportation researchers have since learned more about the factors that influence people's travel decisions, yet forecasting for new modes remains problematic. An academic study[26] carried out in the early 1970s on BART ridership forecasting illustrates the prime difficulty in forecasting

Figure 7

Bay Area Rapid Transit (BART) System



Source: *California Transit Guide.*

Table 6

Forecast and Actual Peak Hour Travel Times on BART

Between and	Concord Powell St. (S.F.)	Rockridge Powell St. (S.F.)	Richmond 12th St. (Oakland)	Berkeley 12th St. (Oakland)	Fremont 12th St. (Oakland)	Daly City Montgomery St. (S.F.)
Forecast (minutes)	34	15	18	8	27	12
Actual 1992	45	22	23	10	34	17

Sources: Parsons Brinckerhoff (1962) and published BART schedules.

for new modes. A multinomial logit mode share model was estimated on a pre-BART sample of travel data. After BART went into operation, the model was tested to see how closely it predicted actual shares of different modes for the work trip (Table 7), and it was found to overestimate total transit usage by 37 percent. The largest errors were made in predicting the walk-to-BART share and the auto-to-BART share.

Table 7

Actual and Predicted Work Trip Mode Shares for BART Study

	Actual Share (%)	Predicted Share (%)
1) Auto alone	59.53	53.19
2) Bus/Walk	10.71	11.37
3) Bus/auto	1.42	2.20
4) BART/walk	0.63	7.53
5) BART/bus	0.94	0.82
6) BART/auto	5.20	3.94
7) Carpool	21.57	20.95

The model was then re-estimated using the same structure with post-BART data to see whether the coefficients would be the same. The model estimated with post-BART data had a higher value placed on the walk-time coefficient and a different "mode-specific"[27] dummy variable for the walk-to-BART alternative (the first model had used the "walk-to-bus" mode specific dummy variable as a substitute). The authors explored several possible sources of error, including failure of the IIA[28] property, non-genericity of attributes of BART with walk access, and incorrect walk time data. The last two factors were thought to contribute to most of the model's misprediction.

The BART case illustrates two of the most common pitfalls in forecasting demand for a new transportation modes. First, technical performance is not guaranteed in new modes. Second, existing behavior patterns may not adequately model behavior with respect to a new mode.

Some Lessons from Rail Transit

What can we learn from the rail transit experience? While ridership forecasting has become more theoretically sophisticated and probably more accurate since the 1960s, there is still a tendency for non-technicians to accept forecasts as irrefutable facts. Progress has been made in understanding the factors which influence people's travel decisions, yet the estimation of "mode-specific" effects for new modes remains one of the more difficult aspects of demand forecasting.

The rail transit experience also reveals that there is sometimes an inherent conflict of interest since engineering consulting firms often provide both the demand forecasting and engineering-construction services. Estimating large, unique civil engineering projects and forecasting demand for new modes are notoriously difficult tasks. Though firms may strive to remain objective, they often receive subtle pressure from their public sector clients to shade a forecast or estimate one way or the other. This shading can occur without compromising professional or technical standards because of the value judgements inherent in the forecasting process.

Once large-scale public projects are begun, they are rarely abandoned no matter how far off the original estimates and forecasts turn out to be. To quote New York's Robert Moses, "Once you drive the first stake, they will never make you pull it up."[29] This observation may or may not apply to high-speed rail in the United States, depending upon its eventual funding mechanism and institutional framework. High-speed rail may require at least some private sector participation in its finance or operation, in which case the private sector will presumably assure itself of the system's viability prior to investment.

CONCLUSION

Travel demand forecasting is a process full of opportunities for error. As we have seen, dependence upon outside forecasts, value judgements, lack of knowledge about the future, uncertainty about macroeconomic trends, and political pressures may all drive a forecast far off course. The evidence gathered on past rail ridership forecasting exercises suggests that, on the whole, ridership forecasting has *not* been successful and casts doubts on the validity of ridership forecasting as a purely objective decision-making tool. As one official of a transportation research organization recently declared, ". . . ridership forecasting is a waste of money but necessary. Those responsible for committing the money want to see numbers and tend to be more comfortable with results coming out of a computer."

Even though the field of transportation demand forecasting has made significant advances in recent years, from a theoretical standpoint, many of the same obstacles that stood in the way of accuracy from the 1960s to the mid-1980s remain in place today. Transportation demand models still rely upon projections of future demographic and economic trends. Planners are still faced with uncertainty over the future of key variables (the Northeast Corridor forecasters did not foresee a deregulated air industry competing with Amtrak trains; similarly, planners today cannot predict the level of gasoline taxes in the future). Finally, and perhaps most importantly, policy objectives will still influence any attempt at an "objective" analysis.

Although past experience with forecasting is most discouraging, one cannot seriously propose a multi-billion project without at least attempting to understand its likely ridership and market potential. Whether the proposed system is publicly or privately financed, ridership estimates are needed as a basis for public discussion and as a way to move the issue of high-speed rail investment forward. Perhaps the most important point taken from the experience with rail forecasting is that ridership estimates must be presented as a set of likely scenarios, with a full discussion of all the assumptions upon which they depend.

A recognition that ridership estimates should not be relied upon in and of themselves as a decision-making criterion would also improve the use of demand forecasting. In his perceptive analysis of the TGV system, Roth outlines the various possible reasons for developing high-speed rail lines: saturation of capacity on the existing transportation network in a corridor; the need or desire to improve rail service to retain market share; stimulation of local or regional economies; strengthening of the national economy through trickle-down effects; and improvement of national transportation infrastructure. Such national and regional objectives should be balanced with the results of "objective" demand studies.

The CalSpeed study will certainly endeavor to meet the above goals in its presentation of a market study. Analysis of the sensitivity of the forecasts to key input assumptions will be an important feature in the study. Additional work in the CalSpeed series will explore possible benefits such as environmental impact, potential infrastructure savings from other modes, and economic development. It is hoped that the ranges of both ridership and revenue estimates, as well as social and economic benefits generated, will provide useful tools for the decision-makers of California.

NOTES

[1] Wachs, p. 246.

[2] Taniguchi, p. 2.

[3] Minemoto, p. 7.

[4] Minemoto, p. 9.

[5] Minemoto, p. 11.

[6] "Train à Grande Vitesse" or "high-speed train."

[7] Roth.

[8] The forecasting methodology is described in Berlioz and Arduin.

[9] SNCF data from Mr. Denis Douté of Rail Transportation Systems, Inc.

[10] Streeter, p. 33.

[11] Streeter, p. 33.

[12] Roth, p. 87.

[13] Halaunbrenner, 1986, p. 752, referenced in Streeter.

[14] Ibid.

[15] SNCF, 1992.

[16] This service will operate with TGV trainsets bought by Deutsche Bundesbahn.

[17] Henn.

[18] FRA (1986), pp. 6-1.

[19] Modelling procedure described in FRA (1978) and Aerospace Corp. (1979).

[20] FRA (1986), pp. 6-7.

[21] Pickrell, p. 19.

[22] Ibid . p. 28.

[23] Pickrell (1990).

[24] Webber (1976), p. 10.

[25] Webber (1976), p. 38.

[26] Train, et. al 1977.

[27] Mode-specific variables are used to capture the effect of modal attributes not accounted for by the generic variables. For instance, one might model the mode split between auto and bus as a function of travel time, cost, and personal income, but a significant degree of predictive power would be found in such hard-to-measure attributes such as uncomfortable bus seats, or the feeling of control in a personal automobile. Dummy variables are introduced to account for such "mode-specific" attributes. Before a mode actually exists, it is very difficult to predict what these mode specific effects might be. In the case of the BART study, behavior observed for people who walked to bus stops turned out not to be a good model for predicting when people would walk to BART stations.

[28] Independence of Irrelevant Alternatives.

[29] Wachs, 1985.

REFERENCES

Aerospace Corporation. 1979. *Demand Forecasting Methodology for the Northeast Corridor*, for the NEC Project Office, FRA, March.

Berlioz, C. C. Vilmart, and J. Arduin. 1985. "An example of response to demand for an improved service offer: the Paris — South-East TGV (high-speed train)." *Transport Policy and Decision Making* 3: 205-281.

Federal Railroad Administration (U.S.). 1986. *Northeast Corridor: Achievement and Potential.* United States Department of Transportation, November.

_____. 1978. *Northeast Corridor Improvement Project Final Programmatic Environmental Impact Statement*, Volume I. U.S. Department of Transportation, June.

Hall, P., D. Leavitt, and E. Vaca. 1992. *High-Speed Trains for California Strategic Choice: Comparison of Technologies and Choice of Route.* Working Paper No. 564, Institute of Urban and Regional Development, University of California at Berkeley, June.

Henn, W. 1992. Interview with Dr.-Ing. Wolfgang D. Henn of Deutsche Eisenbahn-Consulting GmbH, September.

Pickrell, D. 1990. *Urban Rail Transit Projects: Forecat Versus Actual Ridership and Costs.* Office of Grants Management, UMTA, USDOT, October.

Richmond. J. 1991. *Transport of Delight — The Mythical Conception of Rail Transit in Los Angeles.* Massachusetts Institute of Technology, June.

Roth, D. 1990. *The TGV System.* Philadelphia: University of Pennsylvania.

Streeter, W. 1992. *The French Train à Grande Vitesse: Focusing on the TGV-Atlantique.* Working Paper No. 558, Institute of Urban and Regional Development, University of California at Berkeley, April.

Taniguchi, M. 1992. *High Speed Rail in Japan: A Review and Evaluation of the Shinkansen Train.* Working Paper No. 561, Institute of Urban and Regional Development, University of California at Berkeley, April.

Texas TGV. 1991. Section 4.4 in *Texas TGV Franchise Application to Construct, Operate, Maintain, and Finance a High-Speed Rail Facility.* Texas High-Speed Rail Authority, Austin, January.

Thompson, C., and T. Bawden. 1992. "What Are the Potential Economic Development Impacts of High-Speed Rail?" *Economic Development Quarterly* 6(3), August.

Train, K. *Qualitative Choice Analysis. Theory, Econometrics, and an Application to Automobile Demand.* Cambridge: The MIT Press.

Wachs, M. 1985. "Ethical Dilemmas in Forecasting for Public Policy." In *Ethics in Planning*, M. Wachs, ed.. New Brunswick, NJ: Center for Urban Policy Research, 246-258.

Webber, Melvin. 1976. *The BART Experience — What Have We Learned?* Monograph No. 26. Berkeley, California: The Institute of Urban and Regional Development.