FORECAST BASED ON DIFFERENT DATA TYPES: A BEFORE AND AFTER STUDY

M Vrtic*, KW Axhausen
IVT, ETH, CH – 8093 Zürich
vrtic@ivt.baug.ethz.ch, axhausen@ivt.baug.ethz.ch

Abstract

The estimation of route- and mode choice models is of central importance for traffic forecasts. Models which are estimated from revealed preference (RP) data are often subject to statisti- cal problems and might be of limited use. It is therefore standard practise today, to use stated preference (SP) data to supplement or even replace the RP estimates. The main objective of this research project was to predict the changes in traffic demand due to planned infrastructure projects and service changes, and afterwards to verify the results by an before- and after analysis. The most important change was the introduction of tilting trains for the Zürich to Lausanne line via Neuchatel in the Swiss *Mittelland*. For the mode choice forecasts, three different parameter sets were used: Parameters estimated with revealed preference survey data; Parameters estimated with stated preference survey data, which in turn were based on RP data used for the estimation of the previous parameter set; Demand elasticities based on aggregate data. With a before- and after traffic survey (traffic counts and origin-destination surveys), it was possible to compare the actual and the predicted changes of traffic demand, as well as to test the performance of the different parameter sets. The results showed that the forecasts using the SP-parameters performed best, while the RP and elasticity-based forecasts did not perform much worse, they had some consistent problems, which made their application difficult.

Keywords: Route and mode choice; Traffic forecast; Stated preference; Switzerland; Tilting train; ETH Zürich; Institute for Transport Planning and Systems (IVT)
Topic Area: D1 Passenger Transport Demand Modeling

1. Introduction

Next to a suitable data base, knowledge of the links between traffic demand and traffic supply is central to traffic forecasting. The models describing these links are estimated from differrent data sources: aggregate/disaggregate; revealed or stated preference. Most of the literature compares the results derived from the sources partially or only on small spatial scales. For this reason previous Swiss research had recommended to compare the results of forecasts performed with different parameter sets (RP, SP and aggregate elasticities) for a major infrastructure project in the framework of a before-and-after study (Vrtic, Meyer-Rühle, Rommerskirchen, Cervenka und Stobbe, 2000).

The Swiss Federal Railways (Schweizerische Bundesbahnen; SBB) introduced in 2001 a new high-quality tilting train to improve their service along the main east-west corridor between Geneva and St. Gallen via Lausanne, Bern or Neuchatel and Zürich. This ICN (Intercity Neigezug – Intercity tilting train) (ICN) operates on the northern of the two main lines in the corridor, i.e. via Biel and Neuchatel (Jurasüdfuss), and equalizes the travel time between Zürich and Geneve on the two routes (See the northern, red corridor in). Additional improvements of the transport supply for train and road traffic between 1999 and

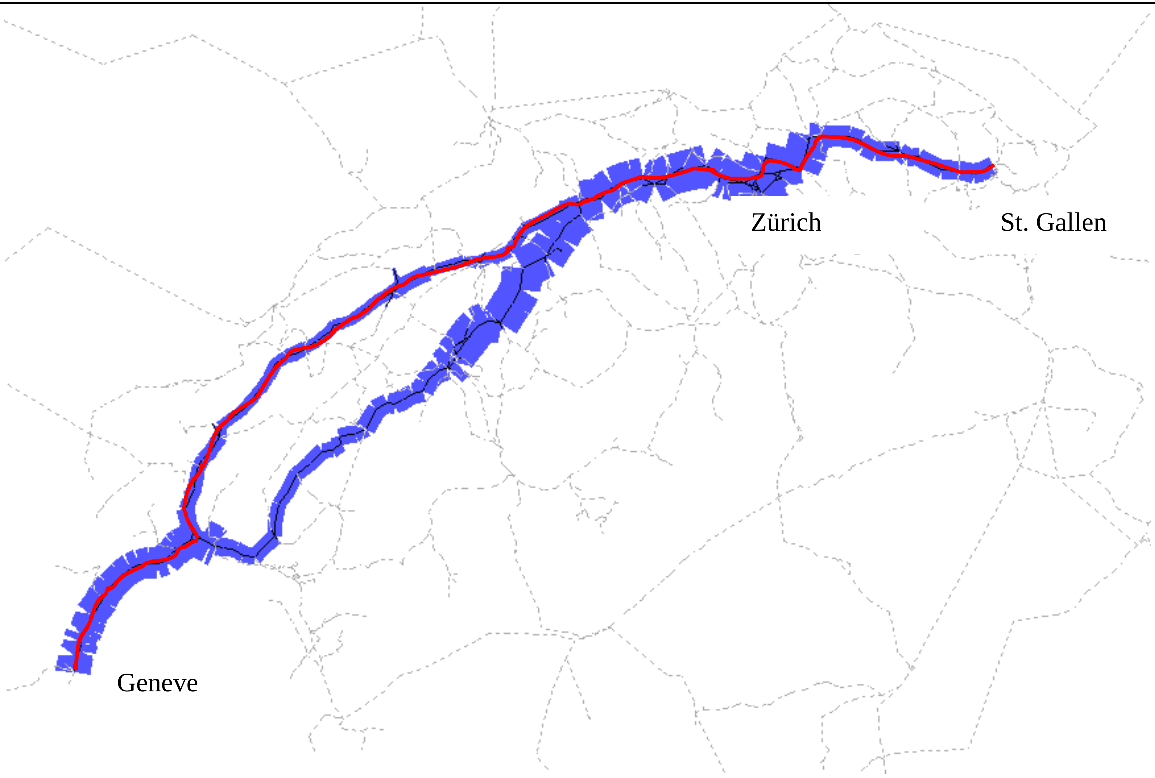
* Corresponding Author: Tel: +41 1 633 31 07, Fax: +41 1 633 10 57

2001 gave the opportunity to conduct the recommended before and after study, which was commissioned by the SBB and the Bundesamt für Raumentwicklung (ARE) (Federal Office for Spatial Development) (Vrtic, Axhausen, Maggi and Rossera, 2002).

While the changes are substantial in the corridor, they are less weighty on a national scale. In the short term, it was felt that any change would be limited to those taking place at the level of mode and route choice. The mode choice forecasts were based on three different parameter sets:

- Parameters estimated with revealed preference survey data
- Parameters estimated with stated preference survey data, which in turn were based on RP data used for the estimation of the previous parameter set
- Demand elasticities based on aggregate data, which had been developed in Vrtic, Meyer-Rühle, Rommerskirchen, Cerwenka und Stobbe 2000

The parameters of the underlying logit-based public transport assignment were estimated using SP data and were used in all three cases.



Source: Vrtic, Axhausen, Maggi and Rossera (2003)

Figure 1 Northerly ICN route within the Swiss East-West corridor

The paper will report the main results from the study, focussing on the comparison of the forecasts with the results from the after – situation. The preceding steps of establishing a before model, performing the needed SP survey, estimating the various models and applying them for a scenario including all relevant improvements to generate the three forecasts will also be discussed.

2. Before - model

The before model for 1999, which was the basis for the forecasts, had to be established with an acceptable accuracy, to track the expected, potentially small, changes in traffic demand. Due to budget constraints we had to limit ourselves to the calibration of suitable network models and their associated origin-destination matrices.

The rail passenger model includes 2587 nodes, 5087 links, 589 lines with 9700 runs and 1108 zones, of which are 137 zones outside Switzerland. The timetable-based logit-type assignment of VISUM 8.0 (PTV AG, 2002; Friedrich, Hofsäss and Wekeck, 2001) was used. For the generalised costs of each connection are a function of in-vehicle-time, start waiting time, transfer time and number of transfer. The start waiting time was calculated as a function of the headway. The relevant parameters were estimated from the stated preference data with a multinominal logit model (see below and also Vrtic, Axhausen, Rossera and Maggi (2001) or Vrtic and Axhausen (2003).

The road model includes 7426 nodes, 10240 links and 3066 zones of which 2975 zones are within Switzerland. The internal zones correspond to the municipalities. Large cities were divided into several zones. An average daily working-day origin-destination matrix previously developed for the ARE (Fusseis and Sigmaplan, 1998) was assigned to obtain a deterministic user equilibrium solution and calibrated to the available traffic counts. See for a comparison of differences after the automatic calibration.

Figure 2 Before-model: Comparison of assignment results with counts after calibration [Railway network]

3. The stated-preference survey

The SBB (Swiss railways) carries out a continuous survey of travel behaviour of the Swiss population (KEP). The KEP is an annual telephone survey of about 16800 persons between 15 and 84 years addressing their mobility behaviour during the preceding seven days, but limited to trips longer than 3 km and crossing a municipal boundary by any means of transport. The following items are recorded: origin and destination municipality, number of transfers, waiting times, in-vehicle time, access time to the station, travel purpose, car availability, car ownership, household size, occupation, employment, possession of season ticket, etc. From January to September 2001 the interviewees were asked about their readiness to participate in a further written SP survey, which was then based on the reported trips during the KEP.

The SP survey was formulated as a stated choice survey. The familiarity of this task allows to describe the alternatives with a relatively large number of attributes without making excessive requirement of the interviewed person (FGSV, 1995; Axhausen and Sammer, 2001).

Public transport trips and car trips were selected to form a representative sample of the KEP reported trips and therefore of the population. Each interviewed person received a complete description of his reported trip as an introduction, and eight situations where he or she had to choose between two alternatives:

- Public transport (PT; train) and car for the mode choice experiments
- Two railway connections/route for the route choice experiments

For the mode choice survey seven attributes described the public transport and three the car alternativ. For the route choice six attributes were considered. Variables had three or four levels. In the experimental design the attribute level changes were formulated as percent or absolute changes from the values, which were reported for the selected RP - trip in the KEP survey. All together 2'234 questionnaires were mailed with a response rate of 68%.

4. Route and mode choice model

The multivariate logit (MNL) (Ben-Akiva and Lerman, 1985) were estimated using the standard software LIMDEP (Econometric Software, 1998). The parameters were generally formulated at mode-specific. In addition, trip purpose specific models were estimated. For the detailed results see Vrtic, Axhausen, Rosera and Maggi, 2003. The parameters show the correct sign and are significant and comparable with surveys in other countries (Abay and Axhausen, 2001; Vrtic et al., 2003). To reflect the inertia of the respondent, they were characterised by their car availability and the season ticket ownership (annual season ticket, half-price-discount-card, local public transport season ticket), which had the expected significant effects.

Table 1 summarises the results by presenting value-of-travel-time-savings (VTTS) and other relevant ratios of the parameters estimated. The pattern of the VTTS matches expectations with higher values of car-based VTTS and for commuting and business trips. Note also, the limited, but noticeable willingness-to-pay for the additional comfort of better rolling stock. The value of reliability is substantial, if one re-expresses the value in mean number of minutes saved.

The implied time-weight of the transfer includes the transfer time to some extent, as the mode choice SP experiment did not include transfer time, as a separate variable.

The parameter of the cost variable could not be estimated in the RP models, due to the high correlation between costs and distances/travel time. In this study, only a time model (without price variable) was estimated for the RP-data in line with former RP and SP-results (Vrtic and Axhausen, 2000; Vrtic, Axhausen, Rossera and Maggi, 2003), where this formulation had the higher explanatory power in comparison with models, which included only costs.

The valuations derived from the public transport route choice model (in-vehicle-time, price, headway, transfer, transfer time, comfort) were very similar to those derived from the mode choice data. The decision maker evaluated the variable independently from and consistent between the choice contexts. (see Vrtic, Axhausen, Magi and Rossera, 2003 or Vrtic and Axhausen, 2003).

Table 1 Mode choice: Ratios of the model parameters (SP)

		All purposes	Commuters	Business	Shopping	Leisure / vacation
VOT Cars-in-vehicle-time	[CHF/h]	27.7	15.2	74.6	25.3	25.3
VOT PT-in-vehicle-time	[CHF/h]	18.5	8.7	55.8	16.0	16.7
VOT headway	[CHF/h]	8.3	4.5	23.7	3.1	6.6
Transfer	[CHF /transfer]	7.6	2.6	21.0	3.9	7.2
Reliability Cars*	[CHF /Prob.%]	0.1		0.7		0.1
Reliability PT*	[CHF /Prob.%]	0.1		0.6		0.1
IR-doubledecker**	[CHF]	2.9	1.0	1.3	3.1	2.6
IC/EC**	[CHF]	6.3	1.7	0.9	3.2	6.9
ICN**	[CHF]	4.9	2.3	14.2	1.6	4.3
in-vehicle-time cars / PT		1.5	1.8	1.3	1.6	1.5
Reliability cars / PT		1.2		1.1		1.2
transfer/in-vehicle-time PT	[min./transfer]	24.6	18.1	22.6	14.6	25.8
headway/in-vehicle-time PT		0.4	0.5	0.4	0.2	0.4
access time PT/in-vehicle-time PT		2.7	2.0	1.5	2.2	2.4

(*) Probability for 10 min. delay
(**) Compared to the regional train

The observed variable means and choice probabilities of the SP data are unrepresentative of the population. The elasticities presented with estimated mode choice parameters (See Table 2) were therefore calculated for the sample means of the KEP survey. These elasticities are valid for the trips over 10 km. Due to matching RP data the elasticities for reliability and comfort were calculated on the basis of the SP survey data.

5. Traffic forecasts

As mentioned above, the traffic forecasts were focused on the mode and route choice changes The supply-induced changes have to overlaid with changes due to demographic and economic change. For this reason, in a first step, the supply-independent demand growth between the 1999 and 2001 was estimated. On the basis of existing, aggregated data for the economic and population trends appropriate models for the road and rail traffic were estimated by Rossera and Maggi, 2003 for the country as a whole, as regionally differentiated data were not available.

The changes in demand due to the supply changes in year 2001 were quantified in three steps:

- Assignment to estimate route choice changes
- Change in mode choice due to the new generalised costs using a pivot-point formulation (Ortuzar and Willumsen, 1995)
- Final assignment and estimation of changes in link traffic volumes.

It was found that this single iteration was sufficient in this case. This process was repeated for each of the three parameter sets: SP, RP and aggregate elasticities.

The procedure allows us to compare the effect of route choice only with the joint effect of route and mode choice. Table 3 and Table 4 provide these results for a selection of rail and road links. As expected the bulk of the changes is due to route choice. The largest demand increases from 10% to 20% can be found on the "Jurafuss" section (Grenchen - Lausanne), where the ICN has the largest impact on travel times and service levels. An additional 2% to 8% increase is due to mode choice in this part of the corridor. It can be seen, that the three parameter sets lead to different mode choice changes: in the "Jurafuss" section the RP set implies smaller and the elasticities EL larger link volumes than the SP approach.

Table 2 Mode choice: Elasticities (travel distance > 10 km; Railway- and Road Passenger traffic)

Variable	Mode	All purposes	Commuters	Business	Shopping	Leisure / vacation
		Calculated with SP parameters and RP mean values				
In-vehicle-time Cars	Cars	-0.425	-0.665	-0.680	-0.545	-0.530
	PT	0.671	0.776	1.531	1.008	0.937
Price Cars	Cars	-0.121	-0.312	-0.076	-0.156	-0.174
	PT	0.191	0.365	0.171	0.288	0.308
In-vehicle-time PT	Cars	0.365	0.480	0.615	0.460	0.456
	PT	-0.575	-0.560	-1.386	-0.850	-0.805
Price PT	Cars	0.157	0.435	0.092	0.223	0.217
	PT	-0.247	-0.508	-0.206	-0.412	-0.383
Access time PT	Cars	0.172	0.272	0.111	0.279	0.127
	PT	-0.272	-0.318	-0.249	-0.515	-0.224
Headway	Cars	0.144	0.320	0.154	0.121	0.116
	PT	-0.227	-0.374	-0.346	-0.224	-0.205
Transfer	Cars	0.115	0.133	0.151	0.101	0.134
	PT	-0.181	-0.156	-0.339	-0.186	-0.237
		Calculated with SP parameters and SP mean values				
Reliability (Cars*)	Cars	-0.024		-0.064		-0.024
	PT	0.049		0.146		0.044
Reliability (PT*)	Cars	0.016		0.050		0.015
	PT	-0.035		-0.114		-0.030
IR-doubledecker**	Cars	-0.012	-0.014	-0.002	-0.024	-0.011
	PT	0.025	0.030	0.006	0.076	0.021
IC/EC**	Cars	-0.027	-0.030	-0.002	-0.024	-0.032
	PT	0.051	0.056	0.004	0.073	0.053
ICN**	Cars	-0.020	-0.042	-0.027	-0.010	-0.018
	PT	0.040	0.071	0.053	0.039	0.034

(*) Probability for 10 min. delay
(**)Compared to the regional train

Table 3 Change of link volume by link on the railway network considering all railway- and road improvements

	Index 1999=100 (*)			
Link	2001 - With RC	2001 – With RC and MC (SP)	2001 – With RC and MC (RP)	2001 – With RC and MC (EL)
St.Gallen – Will	103	104	101	103
Winterthur – Effretikon	106	107	104	106
ZH Flughafen – Zürich	109	110	107	109
Killwangen – Lenzburg	110	112	109	111
Rupperswil – Aarau	106	108	105	107
Aarau – Olten	107	108	106	107
Grenchen – Biel	122	126	125	129
Biel – Neuchatel	119	122	120	125
Neuchatel – Yverdon	120	123	121	127
Yverdon – Lausanne	118	121	118	123
Lausanne – Geneve	105	104	100	102
Herzogenb. – Burgdorf	100	99	97	97
Fribourg – Lausanne	96	94	91	93
Bern – Fribourg	98	98	97	98

RC- route choice change, MC – mode choice change
SP – Stated Preference, RP – Revealed Preference, EL – Elasticity's
(*) Without supply change, 2001= 105 for all link

Table 4 Change of link volume by link on the road network considering all railway- and road improvements

	Index 1999=100 (*)			
Link	2001 - With RC	2001 – With RC and MC (SP)	2001 – With RC and MC (RP)	2001 – With RC and MC (EL)
Payerne - Yverdon	250	255	258	263
Fribourg - Vevey	92	92	93	92
Yverdon - Lausanne	124	124	125	125
Neuchatel - Yverdon	95	94	95	95
Payerne – Moudon - Lausanne	73	73	73	73

RC- route choice change, MC – mode choice change
SP – Stated Preference, RP – Revealed Preference, EL – Elasticity's
(*) Without supply change, 2001= 105 for all link

On the road network the opening of the motor-way section Payerne-Yverdon leads to an increase of the demand in comparison with the former link volume (1999) of approximately 150%. This traffic is diverted from the highway sections Fribourg - Vevey and Neuchatel - Yverdon, as well as from the regional road Payerne - Moudon - Lausanne. With mode choice change, the link volume along the new motorway section will increase by an additional 5% (SP) to 13% (elasticities).

6. After study: Comparison of the traffic forecasts with the 2001 observed volumes

The 2001 link volumes can be compared directly with the three forecasts. While counts for the railway network cover the whole network, road traffic counts were only available for about 70 automatic counting stations. We treat these counts are error free knowing that both types have problems. The railway counts are annualised estimates based on large, but not 100% count. The road counts include intrazonal traffic, which by definition is not modelled in the assignment model.

The SBB produces a station-to-station demand matrix based on on-board surveys of origins and destinations. The small number of surveys any one train services produces instabilities in the estimates from year to year, which we could not correct on a pair-by-pair basis. The SBB produced matrix was automatically calibrated to correct for any error in the grossing up procedure. The estimate for the total number of road trips was derived by repeating the automatic calibration of road matrix with the 2001 counts. In both cases, the SP-parameter based forecasts were closer to the changes estimated from the comparison of the calibrated OD-matrices (Table 5):

Table 5 Predicted and observed change in demand

Change 2001 against 1999 calibrated matrices for the average working day	Rail matrix [additional trips]	Rail link counts [average difference	Rail link counts [average difference, weighted by link volume]	Road matrix [additional trips]
2001 calibrated matrix	+6.37%	+10.7%		+4.26%
SP parameter based – forecast	+5.12%	+10.5%	+7.2%	+4.66%
RP parameter based – forecast	+3.79%	+8.9%	+7.7%	+4.80%
Elasticities based – forecast	+4.93%	+10.0%	+7.5%	+4.74%

The remaining errors could be due to errors in the forecast background growth or in forecast changes due to the network improvements. We do know, that the lack of regional differentiation is problematic. A good example is the effect of the Swiss national exhibition in the summer of 2001, which increased the demand to and from the four cities hosting the show considerably, even before and after the exhibition.

Table 5 shows also the mean differences between the 2001 rail link counts and the forecast values, unweighted and weighted by the relevant link volumes. Again, the SP-based forecasts match the data best. A comparison of road traffic counts and forecasts confirmed this conclusion. The difference between the mean errors for the link counts and the matrix values are due to the fact, the each trip will observed multiple times in the link counts.

Considering these differences, one has to remember that the remaining error after the automatic calibration of the origin-destination matrix has a similar order of magnitude, e.g. in our case an error of 6.9% for the rail network. Clearly, it would be desirable to match the observed counts better, but we are reaching the limit of the relatively simple approached used here. The

shows the differences between the forecasted (SP approach) and the surveyed link volume.

Figure 3 Absolute differences in the link volume: SP-based forecast and traffic counts [Railway network]

7. Conclusion

The three parameter sets and the resulting forecasts be evaluated on the basis the following criteria:

- Completeness of the description of travel behaviour
- Plausibility of the model parameters
- Plausibility of the predicted changes in transport demand
- Differences between observed counts and forecast values

In this study in all four respects, the SP data and the SP parameter set were the best alternative. The SP experiments provided detailed estimates for the most relevant aspects of the two choices modelled. The usual problem of SP experiments, i.e. the lack of relevant constants, did not create problems here, as route choice models do not have constants and the mode choice model was formulated as a pivot-point model. The SP experiments allowed in particular to evaluate the valuations of the comfort of the different types of rolling stock and to separate the valuations of the fact of a transfer, of the transfer time and of the associated unreliability.

The RP parameter set was incomplete and lacked an estimated cost/fare coefficient due to the inherent correlations in the data. This biased the remained parameters, especially in our case the frequency coefficient, which in turn contributed to some strongly biased results. One could argue, that we handicapped the RP model by not including a fixed cost coefficient using a priori knowledge. While true, in many cases, such as Switzerland, this is not possible due to a lack of local experience and prior local results.

The set of elasticity estimates covered only a subset of all variables and more importantly did not include cross-elasticities. Given that we only used elasticities for travel time and headway the results are surprisingly good. Even with the now available estimates (see Table 2) the general limitations in their application remain.

The forecasts with the RP parameters and the elasticities benefited from the use of the SP route choice parameters. Still, the SP parameter forecasts were consistently better then the other two, even if by a smallish margin. While it might go to far to claim that the SP results are always to be preferred, it also clear, that this study has demonstrated that SP based forecasts due not necessarily perform worse than the other two data sources.

Overall, when looked at in detail the SP – based results were more consistent and reasonable. The effects of variables not considered in the RP parameters or in the elasticities produced local or regional biases, which limit the usefulness of the results in many cases.

In summary, this study has contributed substantially to the local knowledge base by providing detailed route and mode choice parameters sets and the direct and cross elasticities derived from them. In addition, the project updated and calibrated the national road and rail networks and origin – destination matrices.

On a methodological level, the study has confirmed the ability of SP-based parameter sets to perform as well, maybe even outperform RP-based parameter sets in large scale applications of travel demand models.

References

Abay, G. and Axhausen, K.W., 2001. Zeitkostenansätze im Personenverkehr: Vorstudie, SVI Forschungsberichte, 42/00, Vereinigung Schweizerischer Verkehrsingenieure (SVI), Zürich.

Axhausen, K.W. and Sammer, G., 2001. Stated Responses: Überblick, Grenzen, Möglichkeiten, Internationales Verkehrswesen, 53 (6) 274-278.

Ben-Akiva, M.E. and Lerman, S.R., 1985. Discrete Choice Analysis, MIT Press, Cambridge.

Econometric Software, 1998. LIMDEP 7.0, Econometric Software, Sydney.

FGSV, 1995. Hinweise zur Messung von Preferenzstrukturen mit Methoden der Stated Preferences, FGSV, Köln

Friedrich M., Hofsäss, I., and Wekeck, S., 2001. Timetable-based transit assignment, 80th Annual Meeting of the Transportation Research Board, Washington, D.C., Januar 2001.

Fusseis / Sigmaplan, 1998. Grundlagenuntersuchung zu Bahn 2000, Bericht an die SBB und den Dienst GVF, Bern.

PTV AG, 2002. Benutzerhandbuch VISUM (06.03.01) Version 7.5, Karlsruhe

Rossera, F., and Maggi, R., 2003. Varianzanalyse und Nachfragewachstum, Zusatzbericht zu Verifizierung von Prognosemethoden im Personenverkehr, report for SBB and the Bundesamt für Raumentwicklung (ARE), USI, Lugano.

Schweizerischen Bundesbahnen, 2001. Kontinuierliche Erhebung Personenverkehr (KEP), SBB, Bern.

Schweizerischen Bundesbahnen, 2000. Nachfrage Schienenpersonenverkehr 1999/2000; Erzeugung einer optimierten O/D-Matrix basierend auf den FQ-Erhebungen, Bern.

Vrtic, M. and Axhausen, K.W., 2000. Modelle der Verkehrsmittelwahl: Regionale Wege in der Schweiz, Stadt Region Land, 69, 193-204.

Vrtic M. and Axhausen, K.W., 2002. Modelle der Verkehrsmittelwahl aus RP- und SP Datengrundlage, Heureka ´02, 293-309, FGSV, Köln.

Vrtic, M. and 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 83th Annual Meeting of the Transportation Research Board, Washington, D.C., January 2003.

Vrtic, M., Meyer-Rühle, O., Rommerskirchen, S., Cerwenka, P., and Stobbe, W., 2000. Sensitivitäten von Angebots- und Preisänderungen im Personenverkehr, SVI Forschungsberichte, 44/98, Bundesamt für Strassen, Bern..

Vrtic, M., Axhausen, K.W., Maggi, R., and Rossera, F., 2003. Verifizierung von Prognosemethoden im Personenverkehr, report for SBB and the Bundesamt für Raumentwicklung (ARE), IVT, ETH Zürich and USI, Zürich und Lugano.