FINAL REPORT

Making the Most of Limited Data in the Evaluation of Advanced Traveler Information Services (ATIS) Through Experimental Resampling:

Cincinnati Case Study

February 2004

Meenakshy Vasudevan
Dr. Karl Wunderlich
Alan Toppen
James Larkin

Contract Sponsor:	Federal Highway Administration
Contract No.:	DTFH61-00-C-00001
Project No.:	900610-D1
Department:	J190

MTS™ Mitretek Systems
Falls Church, Virginia

ABSTRACT

Researchers and practitioners are commonly faced with the problem of limited data in the evaluation of ITS systems. Due to high data collection costs and limited resources, they are often forced to make decisions about the efficacy of a system or technique based on small data sets. How reliable are conclusions based on small data samples? If only limited data are available, how do we maximize their value? This report addresses these questions to evaluate the potential benefits of a prospective notification-based traveler information service that delivers pre-trip travel time information (ATIS) to simulated drivers in Cincinnati, Ohio. In Cincinnati, travel time data were initially available for only 30 weekdays. Our analysis of this small data set using the Heuristic On-Line Web Linked Arrival Time Estimator (HOWLATE) methodology developed by Mitretek Systems indicated that an ATIS user would reduce his trip disutility by 32% versus a comparable non-user. However, since trip experiences on 30 weekdays might not characterize the typical experience of a commuter, conclusions drawn from the small sample might not accurately represent a more generalized assessment of the benefits of ATIS. Hence, we applied an analog of statistical resampling ("experimental resampling") to generate a large sample of days over which we could evaluate the effectiveness of ATIS. With experimental resampling, the reduction in disutility for an ATIS user was only 22%. Our conclusion was that with experimental resampling we had a more reliable estimate of the benefit. In order to validate our claim, we conducted a larger study, using 154 weekdays spanning a year. Our validation analysis found that when compared to the small sample of 30 weekdays, the resampled cases were better predictors of the benefits for the large sample of 154 weekdays.

KEYWORDS: Intelligent Transportation Systems, benefits, HOWLATE, Advanced Traveler Information Systems, resampling, travel time, reliability, simulated yoked trials, Cincinnati.

ACKNOWLEDGEMENTS

The authors wish to acknowledge numerous colleagues for their contributions to this study:

In the Federal Highway Administration (FHWA): Joseph Peters, the sponsor of this work at the ITS Joint Program Office, for his comments and support; Jeff Paniati (ITS JPO) and Robert Rupert (ITS JPO) for their interest and support.

Colleagues from SmartRoute Systems/Westwood One: Joan Ravier, Peggy Solomon, and Richard Marks. Mitretek constructed and archived travel time data in Cincinnati required for this analysis from data provided by SmarTraveler, the SmartRoute Systems entity providing service in the Cincinnati area

Colleagues at Mitretek who contributed to the study: Donald Roberts and Michael McGurrin.

EXECUTIVE SUMMARY

Researchers and practitioners often contend with the problem of small data sets in the evaluation of ITS systems, because of high data collection costs and limited resources. How reliable are findings that are derived from small data samples? If only limited data are available, how do we maximize their value? This report addresses these questions as they pertain to the HOWLATE methodology using resampling techniques. We use an analog of statistical resampling ("experimental resampling") and the Heuristic On-Line Web Linked Arrival Time Estimator (HOWLATE) methodology, developed by Mitretek Systems, to evaluate the potential user benefits of a prospective notification-based traveler information service that delivers pre-trip travel time information to simulated drivers in Cincinnati, Ohio.

In a 2002 HOWLATE study (*1*), conducted by Mitretek Systems at the request of the Intelligent Transportation Systems (ITS) Joint Program Office of the United States Department of Transportation (USDOT), the potential benefits of a prospective notification-based traveler information service was evaluated in Washington, DC and Minneapolis/St. Paul, two metropolitan areas of varying congestion levels and network structure. Since the modeled traveler information service was shown to be beneficial in both these metropolitan areas, a logical extension was the evaluation of benefits in smaller cities with relatively lower population and congestion levels. To address this issue Mitretek selected Cincinnati, Ohio as a test-bed for the HOWLATE analysis.

In the 2002 Urban Mobility Report, Cincinnati, Ohio, a relatively smaller sized radial network, was reported to have a population of about 1.2 million and 19.9 million miles of daily vehicles-miles traveled on freeways and principal arterials with a Texas Transportation Institute (TTI) Congestion Index of 1.26 (*2*). Given that the population, vehicle miles and congestion levels were lower than the previously studied urban areas (approximately, one-half of Minneapolis/St. Paul and one-third of Washington), Cincinnati was chosen as the next candidate city to examine whether commuter experience could be improved through traveler information. Moreover, as with the two previous cities, there was potential for easy acquisition of data since travel time information was readily available for the region from the SmarTraveler web site (www.SmarTraveler.com).

APPROACH

The analysis was conducted using the HOWLATE methodology and made use of travel time information archived at five-minute intervals from 6:30 AM to 6:30 PM (145 time intervals) for each of the 32 directed facilities in the Cincinnati network. The HOWLATE process reconstructs trips and records trip decisions and outcomes of travelers who make routine commutes at various times of the day. It consists of four modules – the travel time archiver, the travel habituation module, the yoked study simulator, and the output post processor.

In the first module of the HOWLATE process, travel-time reports are archived from the SmarTraveler traveler information web site using an automated process for archiving Internet postings (*3*). In the second module, the travel habituation module, commuters establish their habitual routes and determine their trip departure times that result in an acceptable frequency of on-time arrivals (specified in our study to be 95%), based on the travel times they experience over a number of days called the habituation or the training period. In the third module, simulated yoked trials are conducted between a pair of habitual commuters wherein one uses prospective pre-trip notification-based traveler information and the other does not. In simulated yoked trials two subjects are yoked or paired together to conduct identical trips, i.e., they have the same origin, destination, and target arrival time, and their trips are simulated or synthesized. Finally, in the fourth module, the differences between the commute decisions and trip outcomes of the paired commuters are aggregated to identify the potential effect of pre-trip traveler information services.

Archiving of travel time information commenced on 7 November 2001. However, due to technical problems, Mitretek was unable to archive data from SmarTraveler from 18 January 2002 to 17 March 2002. As a result, we initially had usable data for only 30 weekdays over the period from 7 November 2001 to 17 January 2002. Due to schedule constraints, we had to conduct our analysis using the small data archive of 30 weekdays. To compensate, in the first phase of the study, we applied experimental resampling to the small archived sample of 30 weekdays to conduct a more robust evaluation of the impact of ATIS in Cincinnati.

The intent for conducting this study was to examine if ATIS would be as beneficial in a smaller city, such as Cincinnati, with relatively lower congestion levels, as it was found to be in larger cities with higher congestion levels. As mentioned earlier, initially only 30 days were available for use in this evaluation. Typically, a commuter requires a period of 15 days or more to establish his habitual route and departure time (*4*). In the HOWLATE process this period of habituation is called the training period. Hence, if 15 days are used for the training, then we are left with only 15 days to apply the HOWLATE methodology, which uses travel time data to determine mobility benefits. Since traffic conditions on 30 weekdays might not characterize the typical experience of a commuter over a longer period of time, such as a year, benefits determined using the small sample of available days would be less robust and less representative of those for the year. Conclusions that we draw from the small sample might not accurately represent a more generalized assessment of trip outcomes from use of ATIS in Cincinnati.

To address this issue, we used a concept similar to statistical resampling to generate multiple random samples from the original sample of 30 weekdays to emulate a larger sample of days. We then applied the HOWLATE methodology to this collection of samples to recreate trip outcomes of commuters in a simulated environment. Since the resampled cases were further used to perform simulation experiments to determine mobility benefits, we defined this analog of statistical resampling as "experimental resampling". Experimental resampling differs from statistical resampling in that statistical resampling is used to derive results directly without any additional treatment of data, whereas in experimental resampling, random samples are generated to create an initial state, and additional experiments are performed to obtain the trip outcomes.

A training period was generated by randomly sampling from the original sample of 30 weekdays without replacement, leaving out 15 weekdays, which constituted the corresponding evaluation period, the period during which the benefits of pre-trip ATIS were evaluated. The original sample was treated as a virtual population and random subsamples that constituted the training periods were drawn from it repeatedly. This technique created multiple training periods from one sample through random sampling with replacement. It should be noted that data can appear in more than one subsample, because they are generated by random sampling of the original sample with replacement rather than by dividing them into disjoint subsamples. A thousand such random subsamples or training periods were created, and a thousand evaluation periods were created from those days in the original sample that were not in the corresponding training period.

Phase I: Exploratory Analysis of Use of Experimental Resampling

The objective of the first phase was to examine if experimental resampling can be applied when data are limited in the evaluation of ATIS. We applied the HOWLATE methodology to data sample cases: (i) a small sample case, consisting of only 30 weekdays extending from 7 November 2001 to 17 January 2002, and (ii) a thousand resampled cases, created from the small sample case of 30 weekdays, to examine the potential for applying experimental resampling when data are limited.

Yoked pairs of trips were simulated using the HOWLATE methodology for each of the 506 origin-destination pairs in the Cincinnati network for the small sample case as well as the 1000 resampled cases. Statistics were computed for both data samples, and for the resampled cases, the outcomes were averaged over all resampled cases. Tables ES-1 and ES-2 present the trip outcomes for the small sample case and the resampled cases, respectively. Figure ES-1 graphically depicts the commuter benefits experienced with ATIS plotted against Texas Transportation Institute's (TTI) Congestion Index factors for the three cities.

The key results from the analysis are:

1. Without applying experimental resampling techniques in this study, the small data set would have distorted judgment and resulted in biased conclusions about the efficacy of traveler information service in Cincinnati. User benefits with ATIS were overestimated. Commuter expectations of high travel times and travel variability were skewed as a result of limited data. For the small sample case, as can be seen from Table ES-1, use of ATIS in the PM peak period resulted in a 46% increase in just-in-time arrivals and a 93% decrease in early arrivals. Travel disutility cost decreased by 31%. With experimental resampling, as days with high travel times and travel variability were not concentrated in the training period, we were potentially able to reduce the effect of non-recurrent congestion and consequently, the bias due to limited data. Experimental resampling (Table ES-2) showed that use of pre-trip ATIS

in the PM peak period resulted in an increase of just-in-time arrivals by 26% and a decrease in early arrivals by 88% and travel disutility costs by 23%.

2. Pre-trip ATIS proved to be highly beneficial in Cincinnati (Table ES-2). On an average, commuters using ATIS experienced reductions in travel disutility cost by 13%, late arrivals by 30%, early arrivals by 87%, and early and late schedule delays by 23% and 68%, respectively. Just-in-time arrivals increased by 10%. Travel disutility benefits (Figure ES-1) were comparable to that experienced in Washington and Minneapolis/St. Paul due to high travel time variability and trip times (Table ES-3).

Table ES-1. Average Statistics for Cincinnati for the Small Sample Case (7 November 2001 – 17 January 2002: 30 days)

Aggregate Trip Metrics	All Day		AM Peak		PM Peak	
	Non-ATIS User	ATIS User	Non-ATIS User	ATIS User	Non-ATIS User	ATIS User
% Trips Early	13.4%	1.3%	14.5%	1.4%	33.2%	2.5%
% Trips Just in Time	85.0%	96.9%	83.1%	96.6%	65.6%	95.5%
% Trips Late	1.6%	1.8%	2.4%	2.0%	1.2%	2.0%
When Early, Avg Minutes Early	5.9	4.1	6.2	4.3	8.5	4.5
When Late, Avg Minutes Late	3.2	1.1	5.3	1.1	4.2	1.2
Small's Disutility Value	$1.60	$1.40	$1.70	$1.40	$2.10	$1.40
Travel Expenditure (minutes)	26.7	24.9	27.4	25.5	29.9	25.8
Trip Time (minutes)	20.8	20.8	21.3	21.3	21.4	21.4

Table ES-2 Average Statistics for Cincinnati for the Resampled Cases (7 November 2001 – 17 January 2002: 30 days)

Aggregate Trip Metrics	All Day				AM Peak				PM Peak			
	Non-ATIS User		ATIS User		Non-ATIS User		ATIS User		Non-ATIS User		ATIS User	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
% Trips Early	8.7%	2.6%	1.2%	0.2%	12.3%	4.4%	1.4%	0.3%	20.3%	5.9%	2.5%	0.3%
% Trips Just in Time	88.5%	1.8%	96.8%	0.1%	84.7%	3.3%	96.6%	0.2%	75.7%	4.0%	95.4%	0.2%
% Trips Late	2.7%	0.8%	1.9%	0.1%	3.0%	1.1%	1.9%	0.2%	4.0%	2.0%	2.1%	0.2%
When Early, Avg Minutes Early	5.3	0.4	4.1	0.1	5.9	0.6	4.3	0.1	7.0	0.7	4.6	0.1
When Late, Avg Minutes Late	3.8	0.7	1.2	0.1	4.5	1.6	1.2	0.2	4.0	0.7	1.4	0.1
Small's Disutility Value	$1.60	$0.01	$1.40	$0.00	$1.71	$0.04	$1.40	$0.00	$1.92	$0.05	$1.49	$0.03
Travel Expenditure (minutes)	26.3	0.3	25.1	0.1	27.3	0.5	25.7	0.1	28.8	0.6	26.5	0.2
Trip Time (minutes)	21.1	0.1	21.0	0.1	21.5	0.2	21.5	0.1	22.1	0.3	22.0	0.2

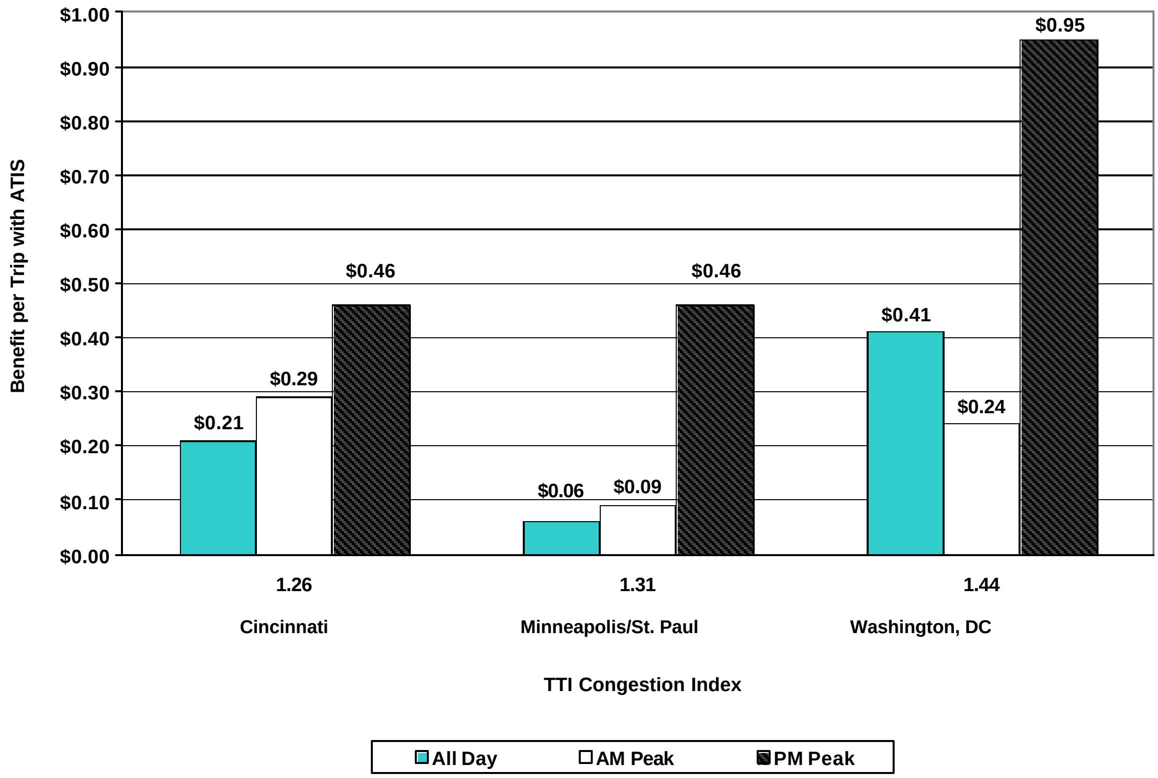


Figure ES-1 Comparison of Travel Disutility Benefits in Cincinnati, Washington, DC and Minneapolis/St. Paul

Table ES-3 Comparison of Average Trip Times and Travel Time Variability in Cincinnati, Washington, DC and Minneapolis/St. Paul

City	Trip Durations (minutes)		Travel Time Variability (minutes)	
	AM Peak	PM Peak	AM Peak	PM Peak
Washington, DC	33.6	34.5	2.1	2.9
Minneapolis/St. Paul	18.6	22.5	1.6	2.4
Cincinnati	22.2	23.1	2.2	2.3

Our premise in the initial phase was that since with experimental resampling we were emulating a larger sample, we would be able to reduce the potential bias in the small sample and make more reliable inferences about traveler information in Cincinnati. However, we were unable to support our claim due to lack of sufficient data at the time the study was initiated. Problems with the data archival process were later resolved and we were able to resume data collection from March 2002. Subsequently, a second phase of the study was implemented to corroborate our assumption that experimental resampling can be beneficial in performing a robust analysis when only limited data are available.

Phase II: Validation Analysis

The objective of the second phase of our analysis was to examine if experimental resampling can be beneficial in providing reliable estimates of user benefits, and to substantiate these results by comparing them to those estimated from a simple time series application of HOWLATE on a large data sample covering nearly a year.

We applied the HOWLATE methodology to three cases: (i) a large sample case, consisting of 154 weekdays extending from 2 September 2002 to 30 May 2003, (ii) a small sample case, consisting of only 30 weekdays extending from 3 September 2002 to 1 November 2002, to replicate the problem of lack of data encountered in phase 1 of the study, and (iii) a thousand resampled cases, created from the small sample case of 30 weekdays. Tables ES-4, ES-5, and ES-6 present the trip outcomes for the large sample case, the small sample case and the resampled cases, respectively. Figures ES-2 and ES-3 compare the relative benefits with use of ATIS for the three cases in the AM and PM peak periods, respectively. The figures show the percentage increase in benefits with use of ATIS for six performance metrics (late and early arrivals, late and early schedule delays, dollar-valued disutility, and trip time).

The key results from the validation analysis are:

1. When data were limited, experimental resampling proved to be an effective technique in providing reliable estimates. Figures ES-2 and ES-3 showed that when compared to the small sample case of 30 weekdays, resampled cases generated using the experimental resampling technique were better predictors of ATIS benefits for the large sample of 154 weekdays. If we had used the small sample to generalize ATIS benefits for commuters in Cincinnati over a year, estimation errors would have ranged from -15% to 32%. With experimental resampling of the small sample, we were able to reduce the estimation errors.

2. Pre-trip ATIS proved to be highly beneficial in Cincinnati (Table ES-6). Our analysis of 154 weekdays of travel time data showed that on average, commuters using ATIS experienced reductions in travel disutility cost by 19%, late arrivals by 37%, early arrivals by 91%, and early and late schedule delays by 30% and 69%, respectively. Just-in-time arrivals increased by 17%

Table ES-4 Average Statistics for Cincinnati for the Large Sample Case (3 Sep 2002 – 30 May 2003)

Aggregate Trip Metrics	All Day				AM Peak				PM Peak			
	Non-ATIS User		ATIS User		Non-ATIS User		ATIS User		Non-ATIS User		ATIS User	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
% Trips Early	14.0%	2.5%	1.3%	0.6%	16.8%	5.4%	1.6%	1.6%	31.4%	9.5%	2.4%	0.8%
% Trips Just in Time	83.3%	3.3%	97.0%	1.3%	80.1%	6.6%	96.6%	2.8%	65.6%	6.9%	95.5%	1.7%
% Trips Late	2.7%	3.8%	1.7%	0.7%	3.1%	8.3%	1.8%	1.6%	3.1%	5.6%	2.2%	1.2%
When Early, Avg Minutes Early	6.2	0.3	4.3	0.1	6.6	0.7	4.5	0.3	8.4	1.2	4.7	0.2
When Late, Avg Minutes Late	6.5	2.9	2.0	1.1	6.2	2.0	2.1	0.9	5.6	2.8	1.7	1.1
Dollar-Valued Disutility Cost	$1.73	$0.26	$1.40	$0.09	$1.82	$0.46	$1.44	$0.18	$2.18	$0.27	$1.52	$0.13
Travel Expenditure (minutes)	27.5	0.4	25.7	0.8	28.3	0.7	26.2	1.6	31.0	0.4	27.4	1.5
Trip Time (minutes)	21.47	0.8	21.43	0.8	21.83	1.6	21.80	1.5	22.87	1.7	22.76	1.5

Table ES-5 Average Statistics for Cincinnati for the Small Sample Case (3 Sep 2002 – 1 Nov 2002)

Aggregate Trip Metrics	All Day				AM Peak				PM Peak			
	Non-ATIS User		ATIS User		Non-ATIS User		ATIS User		Non-ATIS User		ATIS User	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
% Trips Early	9.4%	2.6%	1.7%	0.6%	8.5%	4.3%	1.6%	0.4%	25.8%	9.2%	3.3%	1.1%
% Trips Just in Time	85.4%	4.0%	96.3%	1.1%	85.1%	4.1%	95.8%	1.3%	68.8%	4.1%	94.6%	1.9%
% Trips Late	5.2%	5.8%	2.0%	0.6%	6.4%	6.1%	2.6%	1.2%	5.4%	8.6%	2.0%	1.4%
When Early, Avg Minutes Early	5.3	0.4	4.2	0.1	5.3	0.7	4.3	0.1	7.7	1.3	4.8	0.2
When Late, Avg Minutes Late	4.9	1.9	1.4	0.6	3.8	1.6	1.7	0.8	4.5	1.9	1.7	0.9
Dollar-Valued Disutility Cost	$1.73	$0.28	$1.42	$0.08	$1.78	$0.26	$1.48	$0.10	$2.18	$0.30	$1.55	$0.16
Travel Expenditure (minutes)	26.8	0.4	25.8	0.9	27.3	0.3	26.4	1.0	30.5	0.4	27.7	1.9
Trip Time (minutes)	21.72	1.0	21.63	0.9	22.34	1.1	22.23	1.0	23.12	2.1	22.99	1.9

Table ES-6 Average Statistics for Cincinnati for the Resampled Cases (3 Sep 2002 – 1 Nov 2002)

Aggregate Trip Metrics	All Day				AM Peak				PM Peak			
	Non-ATIS User		ATIS User		Non-ATIS User		ATIS User		Non-ATIS User		ATIS User	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
% Trips Early	13.8%	3.5%	1.6%	0.2%	15.2%	3.7%	2.0%	0.3%	26.3%	7.0%	3.1%	0.5%
% Trips Just in Time	82.8%	3.8%	96.2%	3.6%	81.6%	4.0%	95.6%	3.5%	69.2%	5.5%	94.5%	3.5%
% Trips Late	3.2%	1.2%	1.9%	0.2%	3.1%	1.2%	2.2%	0.3%	4.3%	2.2%	2.2%	0.3%
When Early, Avg Minutes Early	6.0	0.5	4.2	0.0	6.3	0.5	4.4	0.1	7.9	0.9	4.7	0.1
When Late, Avg Minutes Late	4.5	0.6	1.3	0.1	3.3	0.5	1.4	0.2	4.9	1.4	1.5	0.2
Dollar-Valued Disutility Cost	$1.74	$0.07	$1.41	$0.05	$1.78	$0.08	$1.46	$0.06	$2.18	$0.12	$1.54	$0.06
Travel Expenditure (minutes)	27.4	1.1	25.6	1.0	28.1	1.1	26.3	1.0	30.7	1.4	27.6	1.1
Trip Time (minutes)	21.52	0.8	21.46	0.8	21.99	0.8	21.97	0.8	23.09	0.9	22.93	0.9

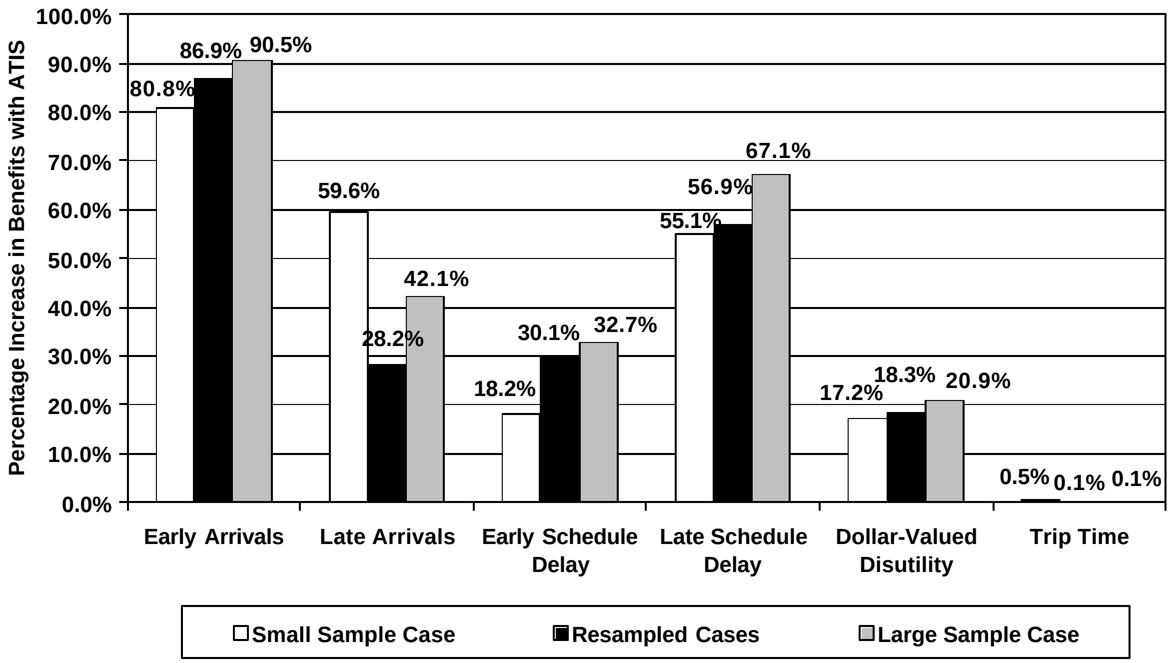


Figure ES-2 Comparison of ATIS Benefits during the AM Peak Period for the Three Data Sample Cases

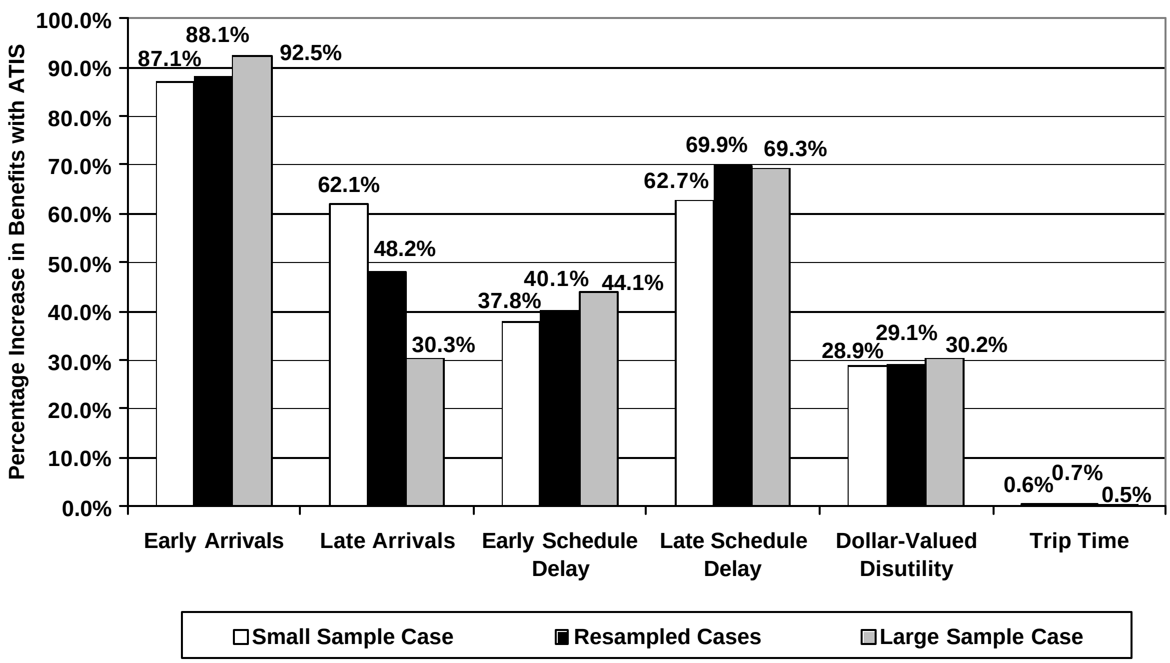


Figure ES-3 Comparison of ATIS Benefits during the PM Peak Period for the Three Data Sample Cases

HYPOTHESES AND KEY FINDINGS

Hypothesis: When compared to a limited data sample, experimental resampling should yield more robust and valid outcomes, and would be more representative of user benefits estimated from a larger sample.

Findings: Applying the HOWLATE process to a sample of 154 weekdays showed that on average, simulated ATIS users in Cincinnati would have increased their trip reliability by 37% and reduced trip disutility by 19% (Table ES-4). If we had been constrained by availability of sufficient data, and had resorted to a direct application of the HOWLATE process to a small sample we would have overestimated trip reliability benefits to users of ATIS by 51% (Table ES-5). Our analysis using a small sample of 30 weekdays showed that on average for the whole day, ATIS users would have reduced their late trips by 61%. Similarly, the small sample showed a reduction in trip disutility by 18.2%. In comparison, applying experimental resampling techniques in this study resulted in a more robust evaluation of benefits to users of pre-trip ATIS in Cincinnati (Table ES-6). Resampling from the small sample of 30 weekdays showed that ATIS resulted in a reduction of late trips by only 39.6% and trip disutility by 19.3%, which were nearly identical to the benefits computed for the large sample case.

Hypothesis: A commuter making use of ATIS will experience higher trip reliability and other mobility benefits than a commuter who does not use ATIS.

Findings: Pre-trip ATIS proved to be highly beneficial in Cincinnati (Table ES-4). Our analysis of 154 weekdays of travel time data showed that on average, commuters using ATIS experienced reductions in travel disutility cost by 19%, late arrivals by 37%, early arrivals by 91%, and early and late schedule delays by 30% and 69%, respectively. Just-in-time arrivals increased by 17%.

Hypothesis: The potential for ATIS benefit is greatest in cities with greatest network variability.

Findings: Our study (Figure ES-1) showed that the potential for ATIS benefits during the peak periods was higher in the city of Washington ($1.19/trip) than in Cincinnati (37 cents/trip). This is in agreement with the TTI congestion index values for these two cities (1.44 for Washington and 1.26 for Cincinnati), and the observed network variability (2.5 minutes in DC compared to 2.1 minutes in Cincinnati). Conversely, the potential for ATIS benefit is higher in Cincinnati than in Minneapolis/St. Paul (6 cents/trip), while the TTI congestion index value for Minneapolis/St. Paul (1.31) is higher than that for Cincinnati. Although the TTI congestion index is strongly correlated with variability, ATIS benefits are affected more by the day-to-day variability than by the magnitude of peak period congestion (*1*). When measuring travel time variability (Table ES-3) from data sets in these cities we found that for the study period, the overall network variability in Minneapolis/St. Paul (1.6 minutes) was lower than that in Cincinnati (2.1 minutes). Thus, we conclude after all that the potential for ATIS benefit is greater in cities with greater day-to-day network variability.

CONCLUSIONS

In this paper we presented an exploratory analysis of maximizing the value of small data sets in the evaluation of traveler information services using an analog of statistical resampling, defined in this paper as experimental resampling. We examined the potential benefits of a prospective notification-based pre-trip traveler information service in Cincinnati using experimental resampling to make reliable inferences. This technique proved to be effective in reducing estimation errors when generalizing traveler information service impacts for commuters in Cincinnati. When compared to the small data sample, resampled cases proved to be better predictors of user benefits estimated from a larger sample. Our analysis showed that a user of ATIS experienced lower disutility, late and early schedule delays and higher trip reliability than his non-ATIS counterpart.

In the evaluation of ITS systems, when data are limited, one can use statistical resampling techniques such as cross-validation, bootstrapping, jackknife, etc. (*5*, *6*) Moreover, small data sets need not be worthless. Their value can be maximized through resampling. The studies mentioned earlier could potentially benefit from the application of bootstrapping. In this study we used experimental resampling as the intent was to create experimental samples to validate commuter experiences.

TABLE OF CONTENTS

1 INTRODUCTION

A common problem in the evaluation of ITS systems is the limited availability of data upon which to base policy or operational decisions. With high data collection costs and limited resources, researchers and practitioners are often faced with the difficult task of making decisions about the usefulness of a system or technique based on insufficient data. Srinivisan and Mahmassani (*7*) investigated the effect of traveler information on commuter behavior in a laboratory environment, and calibrated indifference thresholds of commuters for switching route and departure time. Although the study provides an insight into commuter behavior, it is constrained by the small sample size of less than 15 days. A study conducted to examine the benefits of Houston TransStar showed that because of the freeway monitoring provided by the system, HOV lane restrictions could be promptly lifted when incidents occurred, thereby resulting in annual vehicle delay savings of $42,500 to $85,100 (*8*). However, these results were based only on seven major incidents that occurred in 1996 on monitored freeways with HOV lanes. Another study (*9*) that was performed to examine the impacts of Automatic Vehicle Location (AVL) and Computer Aided Dispatching technologies at Tri-Met, the transit provider for Portland, Oregon, reported that on-time performance of buses increased by 9.4% and wait time reduced by 6.7%, which translated into an annual savings of $1.61 million. These results were based on a before and after study with each period consisting of only ten weekdays.

The point here is not that these studies are somehow flawed; rather, it raises a critical question: If for many ITS evaluations only limited data are available, how do we maximize their value or improve the accuracy of such studies? This report addresses these questions to evaluate the potential benefits of prospective notification-based traveler information service delivering pre-trip travel time information (referred as ATIS for purposes of this report) to simulated drivers in Cincinnati, Ohio, by applying an analog of statistical resampling to create multiple random samples of days. This paper uses the term “experimental resampling” for this technique since resampling is done to create experimental samples of days. The study applied HOWLATE, a technique developed by Mitretek Systems, to these random samples of days to perform simulation experiments to emulate trip outcomes of drivers in Cincinnati.

Background on the motivation for this study is presented in Section 1.1, followed by the main hypotheses of the study in Section 1.2. An overview of the HOWLATE methodology is presented in Section 2. Section 3 provides a description of the data used in this study, including the geographic area that was covered, and the period during which travel time data were downloaded from the SmarTraveler web site. Section 4 introduces the experimental resampling technique that was used to create multiple random sample cases. Section 5 presents the first phase of this study, and shows how experimental resampling was used in evaluating mobility benefits for a small sample of 30 weekdays. Section 6 presents the second and final phase of the study, which was performed to validate our claim that experimental resampling provided a more reliable estimate. Finally, the key findings are summarized in Section 7.

1.1 Background

In a 2002 HOWLATE study (*1*), conducted by Mitretek Systems at the request of the Intelligent Transportation Systems (ITS) Joint Program Office of the United States Department of Transportation (USDOT), the potential benefits of a prospective notification-based traveler information service was evaluated in Washington, DC and Minneapolis/St. Paul, two metropolitan areas of varying congestion levels and network structure. A logical extension was the evaluation of benefits in smaller cities with relatively lower population and congestion levels. To address this issue Mitretek selected Cincinnati, Ohio as a test-bed for the analysis. In the 2002 Urban Mobility Report (*2*), Cincinnati is reported to have a lower population (1.2 million) and lesser congestion (ranked by TTI as the 29th most congested region in the United States), compared to Washington, DC (population of 3.5 million and a TTI congestion index rank of three) and Minneapolis/St. Paul (population of 2.4 million and a TTI congestion index rank of 14). Moreover, as with the two previous cities, there was potential for easy acquisition of data since travel time information was readily available for the region from the SmarTraveler web site (www.SmarTraveler.com).

In DC and Minneapolis/St. Paul, data were available for 200 weekdays. In Cincinnati, however, when we initiated this project data were available for only 30 weekdays from 7 November 2001 to 17 January 2002. Mitretek was unable to archive data from SmarTraveler from 18 January 2002 to 17 March 2002 due to technical problems. To compensate, in the first phase of the study, we applied experimental resampling to the small archived sample of 30 weekdays to conduct a more robust evaluation of the impact of ATIS in Cincinnati. We hypothesized that by experimental resampling we had a more reliable estimate of benefit, but we could not verify our claim in the initial phase because of insufficient data. Technical issues were later resolved, and archiving of travel time data resumed from March 2002. Subsequently, a second phase of the study was implemented to corroborate our assumption that experimental resampling can be beneficial in performing a robust analysis when only limited data are available. We applied experimental resampling to a larger sample of 154 weekdays extending from 2 September 2002 to 30 May 2003.

1.2 Study Hypotheses

The main hypothesis of this paper is that a commuter making use of ATIS will experience higher trip reliability and other mobility benefits than a commuter who does not use ATIS. Moreover, when compared to a limited data sample, experimental resampling should yield more robust and valid outcomes, and would be more representative of user benefits estimated from a larger sample.

Furthermore we also hypothesize that commuters using ATIS in Cincinnati will reap some benefits as experienced in Washington, DC and Minneapolis. However, ATIS benefits will be greatest in cities with the greatest network variability.

2 OVERVIEW OF THE HOWLATE METHODOLOGY

The HOWLATE methodology (*10, 11*) applies dynamic programming to archived observed roadway travel time data to quantify the mobility benefits of ATIS on simulated commuters. HOWLATE constructs synthetic, retrospective yoked driving trials between a pair of habitual commuters, one who does not use ATIS, and the other who uses ATIS, who have the same origin, destination, desired arrival time and habitual route. A yoked study is a field trial in which, two subjects are coupled together so that they start their trip simultaneously, and report their trip experiences at the end of the trip. One subject uses ATIS, while the other does not (*12*). In this study, we simulated the yoked trials based on a study done by Lappin *et* *al*. (*13*), so that each simulated pair has the same origin, destination, desired arrival time and habitual route, but unlike in the field trial, is not restricted to leave at the same time. Key assumptions of HOWLATE are:

- Both commuters (one who uses ATIS and the other who does not) aspire to arrive at their destination on-time 95% of the time rather than reduce travel times.
- Both commuters are habitual commuters, and are familiar with their route and the travel times on their customary routes at various times of the day.
- The commuter who uses ATIS is familiar with the accuracy of the service.

It consists of four modules - the travel time archiver, the travel habituation module, the yoked study simulator, and the output post processor.

2.1 Travel Time Archiver

The travel time archiver is a software application that monitors travel time reports from SmarTraveler, and stores these reports at five-minute intervals (*3*). The archiver compiles and saves a daily profile of travel times by roadway, by time of day, and date.

2.2 Travel Habituation Module

As one of the assumptions of this study is that the commuters are habitual commuters, habitual routes and trip start times that result in an acceptable frequency of on-time arrivals (95%) need to be determined. This is done in the travel habituation module. The period of habituation, when commuters establish their regular routes and determine their trip start times based on the travel times they experience, is called “the training period.”

In order to establish a habitual route and trip start time, HOWLATE tries to re-create an “actual day.” As SmarTraveler does not update travel time information in real-time, statistical distribution of error between the SmarTraveler travel time reports and the observed travel times are estimated. The distribution of error is based on findings from a study (*14*) conducted on one freeway and one arterial in the Washington, DC region. We applied the same error distribution to travel times reported in Cincinnati based on our assumption that the service provided by SmarTraveler in Washington, DC and Cincinnati were comparable. The error distribution and travel time reports are used to construct multiple “actual day” profiles through independent Monte Carlo trials. Average travel times at five-minute intervals are obtained for each link across all days in the training period using the actual day profiles. Next, for each trip fastest time-variant path and the associated travel time are identified using the technique of Kaufman (*15*). The fastest path is then selected as the habitual route for non-ATIS users. The travel time variability for each route is estimated by computing the variability of its travel time over the days in the training period. The habitual trip start time is determined by subtracting from the target arrival time, the average travel time on the habitual path plus a time buffer (arrival buffer) proportional to the travel variability and level of on-time arrival confidence (95% in this study). The buffer size is computed under the assumption that day-to-day variation in travel times in the training period is normally distributed.

In addition, ATIS users discount or inflate the travel time estimates provided by the ATIS service based on the observed accuracy of those reports in the training period.

2.3 Yoked Study Simulator

After habitual routes and trip start timings are determined, yoked driving trials are simulated between ATIS and non-ATIS commuters. The trials are simulated in the sense that the travel experiences are retrospectively reenacted from historical archives. Traffic simulation or other similar tools are not utilized.

The period during which commuter experiences (i.e., trip variability, preferred route and departure time, and accuracy of ATIS service), learned during the training period, are validated is known as the "evaluation period." Note that validation is performed by comparing the trip outcomes of the ATIS and non-ATIS commuters. One realization of travel congestion is generated for each day of the evaluation period.

Simulated yoked trials are conducted using a single Monte Carlo realization for each day in the evaluation period. The non-ATIS user departs from the origin at his habitual time and takes his customary route. In contrast, the ATIS user is notified by the ATIS service of his trip departure time and route based on existing travel times. The ATIS service recommends departure at the first five minute departure time where the arrival time is within the arrival buffer. Note that, if by taking the current fastest path the commuter is projected to arrive at his destination prior to the arrival buffer, the service postpones notifying the user about a trip start by five minutes. When the trip can no longer be postponed, the ATIS service determines if the travel time on the fastest route is less than the travel time on the commuter's habitual route. If travel-time-savings are more than a pre-determined indifference threshold (three minutes; *16*), the ATIS service alerts the user of the projected trip start time and the alternate fastest route; otherwise the commuter is advised to take the habitual route and leave at the projected trip start time. Once the route and trip start time are determined, the ATIS user traverses the network on the suggested path and does not alter the route. Note that the service may also contact the traveler to suggest trip start timing later than the habitual start time if traffic conditions are better than normal during that particular day.

2.4 Output Post Processor

Each simulated yoked trial is analyzed in the output post-processor module. The performance measures computed for both ATIS and non-ATIS users in this analysis were on-time reliability, schedule delay, and the dollar-valued travel disutility. All metrics were computed for unitary trips as there were no data on network flows.

On-time reliability is defined as the proportion of trips wherein a traveler arrives at the destination at or prior to the target arrival time. If a traveler arrives at the destination more than ten minutes prior to the target arrival time, then it is called an *early arrival*. If a traveler arrives prior to, but within ten minutes of the target arrival time, it is called a *just-in-time arrival*, and if he arrives after the target arrival time, it is called a *late arrival*.

Schedule delay is defined as the difference between the actual arrival at the destination and the target time of arrival. If schedule delay is negative, it is *early schedule delay*, and if positive it is *late schedule delay*.

The *dollar-valued travel disutility* is a measure of disutility associated with a trip by assigning a cost to the duration of travel time and how early or late one reaches one's destination (*17*).

3 CHARACTERISTICS OF PRIMARY DATA

This section describes the data used in the analysis. A description of the geographic area for which the study was conducted is first described in Section 3.1. As the HOWLATE methodology uses knowledge of trip durations to determine the impacts of traveler information services, travel time data are essential to the HOWLATE process. Section 3.2 describes the period and extent of coverage of the travel time data. Section 3.3 discusses how the peak periods are determined.

3.1 Geographic Coverage

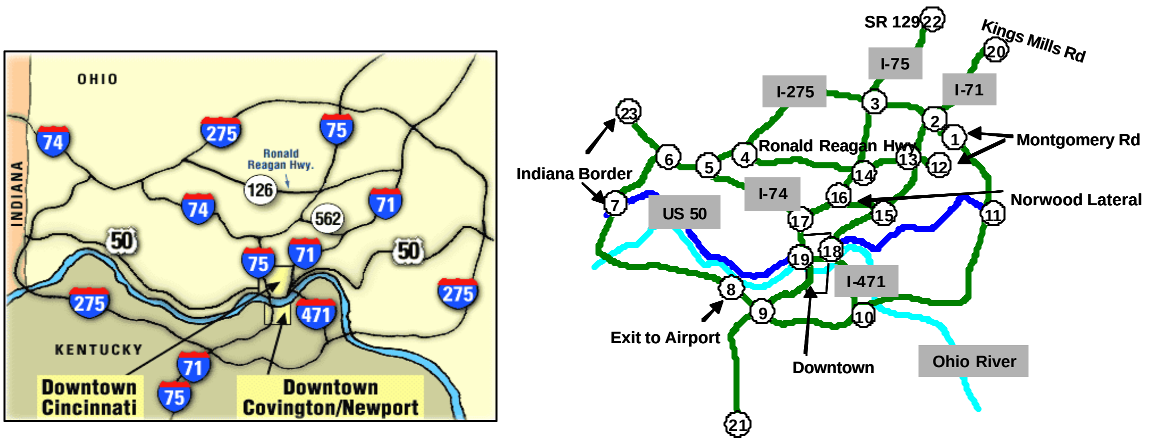
The geographic area used in conducting this study was dictated by the coverage of travel time data SmarTraveler provided for this region. For Cincinnati, this region is bounded by the Indiana border on the West and Clermont County on the East, Kentucky on the South, and Kings Mills on the North, and includes 16 roadway segments or facilities. The travel time information posted on SmarTraveler is provided by the ARTIMIS system, which is a joint venture of the Ohio-Kentucky-Indiana Council of Governments, Kentucky Transportation Cabinet, ODOT, FHWA and City of Cincinnati (*18*). The ARTIMIS system gathers traffic information data using cameras, detectors, DMS, HAR, freeway service patrol vans, etc.

SmarTraveler reports travel times in both directions for 16 roadway segments (termed facility), resulting in travel time reports for a total of 32 directional facilities (28 freeways and four arterials). These facilities constitute a total of 473.2 directional miles, of which 391.8 miles are freeway miles and the remaining 81.4 miles are arterial miles. The average length of a facility is 15 miles, with a standard deviation of 7.3 miles. The longest facility is I-275 between the airport and the Ronald Reagan Highway and is 26.6 miles. The shortest facility is the Norwood Lateral between I-75 and I-71 and is 3.2 miles. Table 1 gives a description of the 16 facilities.

Table 1. SmarTraveler's Descriptions of Facilities for Cincinnati, Ohio

Facility Number	Facility Description	Type	Length (miles)
1	I-275 between I-75 and Montgomery Rd.	Freeway	6.3
2	I-275 between Montgomery Rd. and I-471	Freeway	25.2
3	I-275 between the Ronald Reagan Highway/SR 126 and I-75	Freeway	12.4
4	I-71 between Kings Mills Rd. and the Norwood Lateral/SR562	Freeway	19.4
5	I-74 between the Indiana Border and I-75	Freeway	19.5
6	I-75 between the Norwood Lateral/SR 562 and the Michael A. Fox Hwy./SR 129	Freeway	16.2
7	Norwood Lateral/SR 562 between I-75 and I-71	Freeway	3.2
8	Ronald Reagan Highway/SR 126 between Montgomery Rd. and I-275	Freeway	17.2
9	US 50 between Downtown and the Indiana Border	Arterial	22.7
10	I-275 between the Airport and Ronald Reagan Highway	Freeway	26.6
11	I-275 between I-471 and the Airport	Freeway	14.0
12	I-471 between I-275 and Downtown	Freeway	4.9
13	I-71 between the Norwood Lateral/SR 562 and the Ohio River	Freeway	7.5
14	I-71/75 between the Ohio River and the I-71/75 Split	Freeway	19.7
15	I-75 between the Norwood Lateral/SR 562 and the Ohio River	Freeway	7.3
16	US 50 between I-275/Clermont County and Downtown	Arterial	18.0

To apply the HOWLATE methodology to the Cincinnati region, the SmarTraveler network was converted to a link-node representation. The 32 bi-directed facilities in the SmarTraveler network were further divided into 64 single direction links (60 freeway links and four arterial links) and 23 nodes (potential trip origins or trip destinations). Thus, a total of 506 (23 x 22) origin-destination pairs are simulated across the network of 23 nodes and 64 links. Figure 1 shows the geographic coverage by SmarTraveler for Cincinnati, Ohio, and the corresponding HOWLATE link-node representation. The average directed length of a HOWLATE link was 7.4 miles, with a standard deviation of 5.1 miles. The longest link was US 50 between the Indiana border and downtown Cincinnati and was 22.7 miles. The shortest link was the Ronald Reagan Highway between I-71 and Montgomery Road and was 0.6 miles.



SMARTRAVELER REPRESENTATION

HOWLATE REPRESENTATION

Figure 1. Map of the Cincinnati Network

3.2 Archived Travel Time Data

This section provides a description of the travel time data used in this study. Mitretek started downloading travel time data for each facility in Cincinnati from the SmarTraveler web site on 6 November 2001. Data were archived at five-minute intervals from 6:30 AM to 6:30 PM (145 time intervals) for each of the 32 directed facilities. However, due to technical problems, Mitretek was unable to archive data from SmarTraveler from 18 January 2002 to 17 March 2002. Even though the technical issues were resolved, because of schedule constraints we were unable to use archives from later in 2002 for this study. Altogether, we could archive travel time data for a total of only 44 weekdays, of which only 30 were usable for analysis. The remaining 14 days had to be excluded from the analysis due to missing data, which were caused either due to a temporary failure of the SmarTraveler web site or due to problems in our automated data archiving process.

Table 2 lists the dates when data could be successfully acquired. In November, of the 19 potential weekdays, travel time was archived for 9 days, resulting in 47% acquisition rate. In December and January the acquisition rate was 57% and 64%, respectively. There were a total of 4640 (145 x 32) archived travel-time reports for each of the 30 days. The travel time for each facility was then divided among its corresponding HOWLATE links based on the assumption of uniform speed on the facilities.

Table 2. Travel Time Data Acquisition Dates for Cincinnati

Calendar Month	Complete Days	Potential Days	Percent Coverage	Dates of Coverage*																				
				M	T	W	Th	F	M	T	W	Th	F	M	T	W	Th	F	M	T	W	Th	F	M
Nov '01	9	19	47%		-	7	8	-	-	13	14	15	16	19	-	-	-	-	-	-	-	29	30	
Dec '01	12	21	57%	3	4	5	6	7	10	-	-	-	-	17	18	19	-	-	-	-	-	27	28	31
Jan '02	9	14	64%		-	-	3	-	7	8	-	10	11	14	15	16	17	-						

* Note: Gray dates are potential days (weekends excluded). "-" indicates days dropped due to missing data. The numbers are the actual dates with complete data.

3.3 Defining Peak Periods

A previous study (*1*) using the HOWLATE methodology has shown that ATIS benefits are the greatest during PM peak periods when network-wide travel times are high. Hence, to examine the effect of ATIS during the congested periods in Cincinnati, it is important to first determine the peak periods. Moreover, the start and end times of peak periods are required inputs in the HOWLATE process.

A cluster analysis was performed on the average facility speeds. Cluster analysis is a technique whereby data can be organized into homogeneous groups or clusters. To perform the cluster analysis, first, the average speed for each facility was determined for each five-minute time interval from 6:30 AM to 6:30 PM across the 30 days using the archived travel times. Next, network-wide speeds were calculated for each five-minute time interval by averaging the previously calculated average facility speeds weighted by the lengths.

A two-cluster analysis was performed to separate the average speeds into distinct clusters of peak and off peak periods. Figure 2 shows the average speeds by time of day. The AM peak speeds for southbound and westbound facilities were higher than the AM peak speeds for northbound and eastbound facilities or the PM peak speeds. The degree of association between the AM peak and off peak proved higher than that between the AM peak and the PM peak. Hence, the cluster analysis was repeated by defining three clusters to distinguish the AM peak from the off peak. The AM peak period was determined to extend from 7 AM to 9 AM and the PM peak period from 4 PM to 6:30 PM. The off peak periods were from

6:30 AM to 7:00 AM and 9 AM to 4 PM. The average travel time during the AM peak periods of the thirty days was 17.7 minutes with a standard deviation of 2.4. During the PM peak period, the average travel time was 18.0 minutes and standard deviation was 2.2, while during the off peak period these were 16.7 minutes and 1.7, respectively.

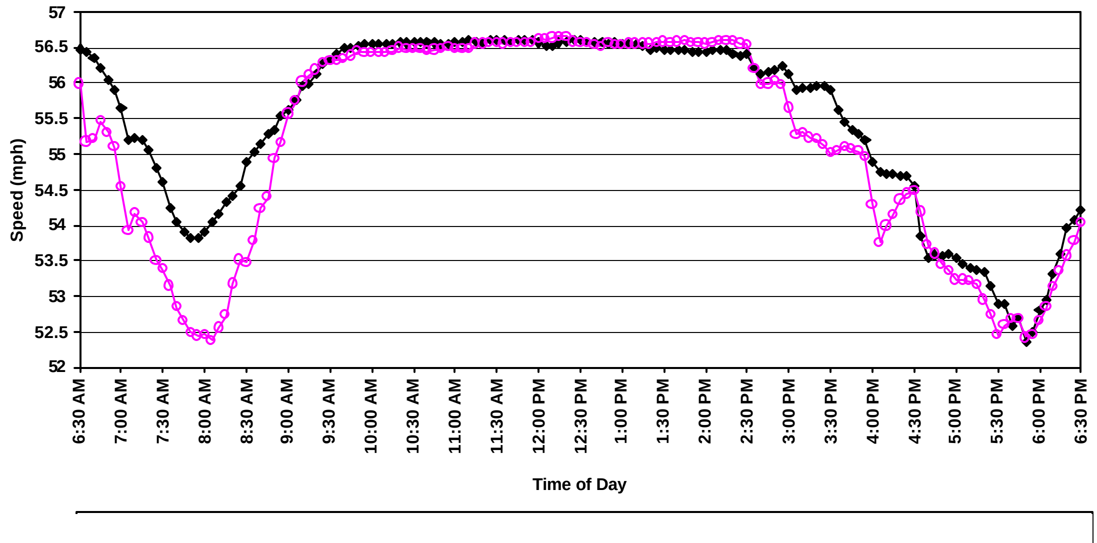


Figure 2. Average Speed by Time of Day

4 EXPERIMENTAL RESAMPLING

The intent for conducting this study was to examine if ATIS would be as beneficial in a smaller city, such as Cincinnati, with relatively lower congestion levels, as it was found to be in larger cities with higher congestion levels. As mentioned earlier, only 30 days were available for use in this evaluation. Typically, a commuter requires a period of 15 days or more to establish his habitual route and departure time (*4*). In the HOWLATE process this period of habituation is called the training period. Hence, if 15 days are used for the training, then we are left with only 15 days to apply the HOWLATE methodology, which uses travel time data to determine mobility benefits. Since traffic conditions on 30 weekdays might not characterize the typical experience of a commuter over a longer period of time, such as a year, benefits determined using the small sample of available days would be less robust and less representative of those for the year. Conclusions that we draw from the small sample might not accurately represent a more generalized assessment of trip outcomes from use of ATIS in Cincinnati.

To address this issue, we used a concept similar to statistical resampling to generate multiple random samples from the original sample of 30 weekdays to emulate a larger sample of days. We then applied the HOWLATE methodology to this collection of samples to recreate trip outcomes of commuters in a simulated environment. Since the resampled cases were further used to perform simulation experiments to determine mobility benefits, we defined this analog of statistical resampling as “experimental resampling”. Experimental resampling differs from statistical resampling in that statistical resampling is used to derive results directly without any additional treatment of data, whereas in experimental resampling, random samples are generated to create an initial state, and additional experiments are performed to obtain the trip outcomes.

A training period was generated by randomly sampling from the original sample of 30 weekdays without replacement, leaving out 15 weekdays, which constituted the corresponding evaluation period, the period during which the benefits of pre-trip ATIS were evaluated. The original sample was treated as a virtual population and random subsamples that constituted the training periods were drawn from it repeatedly. This technique created multiple training periods from one sample through random sampling with replacement. It should be noted that data can appear in more than one subsample, because they are generated by random sampling of the original sample with replacement rather than by dividing them into disjoint subsamples. Multiple such random subsamples or training periods were created, and multiple evaluation periods were created from those days in the original sample that were not in the corresponding training period.

The recommended number of resampled cases range from 50 to 2000 depending on the application (*19*). We chose 1000 as the number of random samples to replicate commuter experiences over a larger sample. Results are dependent on the number of samples, but one of the objectives of this study was to show that when data are limited, experimental resampling from the limited data will provide better predictors of a commuter’s experience over a longer period of time. Hence, our basis for selecting the number to be 1000 was that we needed to create a large sample and the larger the sample, the better predictor it would be of the overall benefits. The next section describes the first phase of the study.

5 PHASE I: EXPLORATORY ANALYSIS OF USE OF EXPERIMENTAL RESAMPLING

The objective of the first phase was to examine if experimental resampling can be applied when data are limited in the evaluation of ATIS. This section describes how experimental resampling was applied to the small sample of archived travel time data for 30 weekdays to determine user benefits from ATIS.

5.1 Creating Data Samples

We applied the HOWLATE methodology to two data sample cases: (i) a small sample case, consisting of only 30 weekdays, and (ii) a thousand resampled cases, created from the small sample case of 30 weekdays. This section presents the method used to create the two data sample cases.

5.1.1 Small Sample Case

The small sample case was defined by using the customary approach in the HOWLATE process of using the first few days for training and the remaining for evaluation. Thus, the training period extended from 7 November 2001 to 10 December 2001, and the evaluation period from 17 December 2001 to 17 January 2002. The training and evaluation periods each consisted of 15 weekdays.

5.1.2 Resampled Cases

In the HOWLATE process since we were constrained by the fact that we needed a minimum of 15 distinct days for the training period, a random sample of size 15 was created from the small sample of 30 days without replacement. This random sample made up the training period and the remaining 15 days made up the corresponding evaluation period. A thousand such random samples or training periods were created by repeatedly sampling from the original sample of 30 days, and evaluation periods were created from those days in the original sample that were not in the corresponding training period. The period of days that were randomly sampled extended from 7 November 2001 to 17 January 2002.

Figure 3 illustrates how the small sample case and resampled cases were generated from the original sample of 30 days.

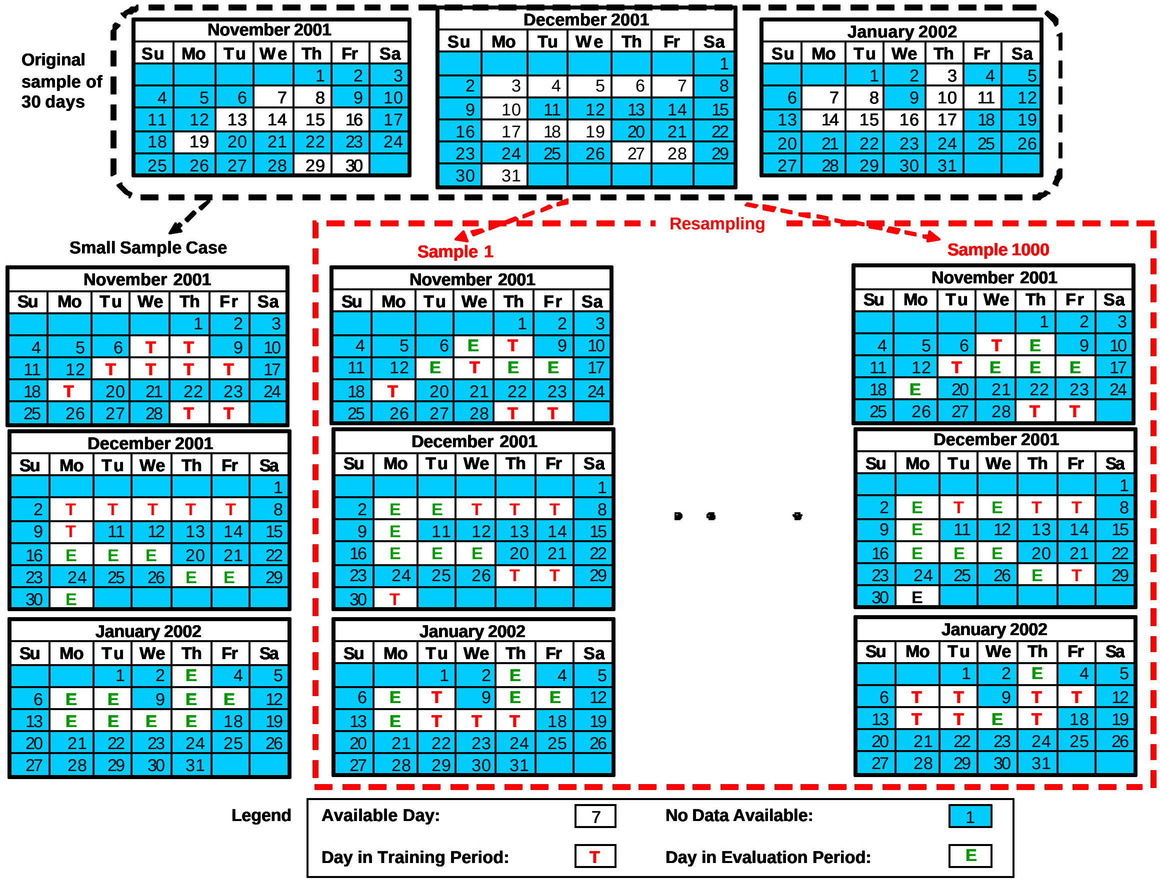


Figure 3. Process for Generating Small Sample Case and Resampled Cases

5.2 Travel Time Characteristics

This section compares the network travel times and variability for the two cases, and presents the hypothesis for each case. Table 3 shows the average network travel times and variability for the training and evaluation periods for the two cases.

5.2.1 Small Sample Case

The SmarTraveler website reported incidents and heavy demand during the AM and PM peak periods on eight of the 15 days in the training period, compared to four days in the evaluation period. Consequently, the training period had relatively higher travel times and travel variability than the evaluation period, possibly also influenced by the fact that the evaluation period is a holiday period. Given that, for the small sample case we hypothesized that the non-ATIS user would be more conservative and experience higher early schedule delays than his ATIS counterpart.

Figure 4 illustrates the variation of the average facility travel times reported by SmarTraveler during the AM and PM peak periods. Over the study period, there is a 12% reduction in travel time during the PM peak period and a 3% reduction during the AM peak period.

5.2.2 Resampled Cases

It was expected that by experimental resampling, the confounding effects of non-recurrent congestion and holiday traffic patterns would be reduced since days were evenly distributed among the training and evaluation periods. For the resampled cases, our hypothesis was that the non-ATIS user would have relatively smaller frequency of early arrivals in comparison to the small sample case.

For both cases, however, the ATIS user was expected to have lower travel disutilities and delays.

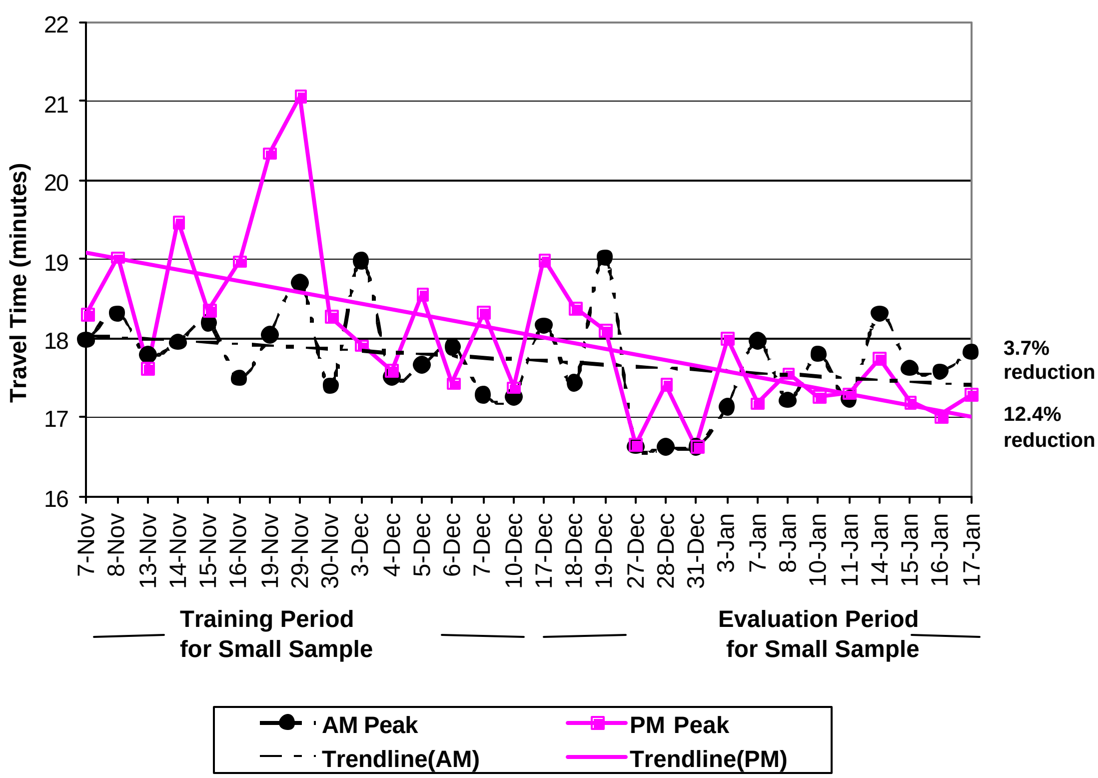


Figure 4. Variation of Facility Travel Time Over the Study Period

Table 3. Facility Travel Times and Travel Variability

Time of Day	Baseline Case				Resampled Cases			
	Training Period		Evaluation Period		Training Period		Evaluation Period	
	Average (minutes)	Standard Deviation (minutes)	Average (minutes)	Standard Deviation (minutes)	Average (minutes)	Standard Deviation (minutes)	Average (minutes)	Standard Deviation (minutes)
AM Peak	17.9	2.3	17.5	2.2	17.7	2.3	17.7	2.2
PM Peak	18.6	2.6	17.5	2.1	18.1	2.4	18.0	2.3

5.3 Simulated Yoked Study

Simulated yoked trials (defined in §2) were conducted between the non-ATIS user (F95) and the ATIS user (A95) for each origin-destination pair. Five Monte Carlo realizations were conducted for each day in the evaluation period for target arrival times at 15-minute intervals between 6:30 AM and 6:30 PM. The training of commuters and evaluation of ATIS were performed for the small sample case and the thousand resampled cases. Statistics were computed for each of the cases, and for the resampled cases, the experimental resampling standard errors for each performance measure were also determined. Please note that the users or commuters mentioned here are simulated users or commuters, and the trip outcomes are the simulated trip outcomes.

The trip decisions made by the A95 commuter for the small sample case and the resampled cases are discussed in Section 5.3.1. Section 5.3.2 presents the trip outcomes for the two data sample cases, and compares the results.

5.3.1 Trip Decisions

This section presents the choices made by the A95 commuter, with respect to changes in his habitual trip start time and/or route, as a result of the reported traffic conditions and his prior experience and knowledge of the accuracy of the traveler information service. The analysis is performed for the small sample case and the resampled cases.

5.3.1.1 Small Sample Case

In Cincinnati, on an average during the day, traveler information services recommended a change in departure time more often than a change in route. The A95 commuter changed his departure times on 31% of the trips and routes on 1% of the trips. Table 4 illustrates the trip decisions made by the A95 commuter when compared to the F95 commuter for the small sample case for the peak periods and the whole day. When changing the departure time, the A95 commuter left earlier than normal by 5.8 minutes or later by 7.1 minutes.

During the PM peak period A95 commuters deviated from their habitual behavior on 57% of the trips. Overall, they altered their departure times on 55% of the trips and routes on 3% of the trips. They changed only their departure times on 54% of the trips, only routes on 2%, and both route and departure times on 1% of the trips. The A95 commuter typically took a longer route when taking a different route from normal, although on an average these routes were longer only by 0.4 miles. When departing earlier than the regular departure time, the A95 commuter left early by 6.5 minutes and when departing later than normal, he left late by 8.1 minutes.

Table 4. Trip Decisions of A95 Compared to F95 for the Small Sample Case (7 November 2001 – 17 January 2002: 30 days)

Travel Choice Category	All Day	AM Peak	PM Peak
Trips with Both Route and Departure Time Changes	0.6%	0.7%	1.4%
Trips with Only Route Changes	0.8%	1.1%	1.6%
Trips with Only Departure Time Changes	29.9%	34.1%	53.6%
Trips with No Change	68.8%	64.1%	43.4%
Trips with Route Changes:			
% Resulting in Shorter Routes (with respect to length)	14.4%	11.4%	22.8%
% Resulting in Longer Routes (with respect to length)	85.6%	88.6%	77.2%
Avg. Miles Route is Shorter (when taking shorter route)	1.8	1.3	2.3
Avg. Miles Route is Longer (when taking longer route)	0.3	0.4	0.4
Trips with Departure Time Changes:			
% With Early Departure	7.7%	11.3%	4.3%
% With Late Departure	92.3%	88.7%	95.7%
Avg. Minutes Early Departure (when departing early)	5.8	5.5	6.5
Avg. Minutes Late Departure (when departing late)	7.1	6.8	8.1

5.3.1.2 Resampled Cases

The trip decisions made by the A95 commuter for each of the thousand resampled cases were aggregated for the AM peak, off peak and PM peak periods, and the averages over all thousand samples were determined, which are shown in Table 5. On an average, the A95 commuter changed his departure time on 26% of the trips and route on 2% of the trips. The A95 commuter, when changing his departure time, left earlier than normal by 6 minutes or later by 6.3 minutes.

During the PM peak period A95 commuters deviated from their habitual behavior on 48% of the trips. They changed only their departure time on 44% of the trips, only route on 2%, and both route and departure time on 2% of the trips. The A95 commuter left later than his regular departure time five times more often than he left early. When departing earlier than his regular departure time, he left early by 6.3 minutes, and when departing later than normal he left late by 6.8 minutes. When deviating from the habitual route, the A95 commuter typically took a longer route, which was on an average longer than the habitual route by 0.6 miles.

Table 5. Trip Decisions of A95 Compared to F95 for Resampled Cases (7 November 2001 – 17 January 2002: 30 days)

Travel Choice Category	All Day	AM Peak	PM Peak
Trips with Both Route and Departure Time Changes	0.7%	0.7%	1.6%
Trips with Only Route Changes	1.1%	1.2%	2.5%
Trips with Only Departure Time Changes	25.2%	32.7%	43.7%
Trips with No Change	73.0%	65.4%	52.2%
Trips with Route Changes:			
% Resulting in Shorter Routes (with respect to length)	8.5%	9.4%	10.1%
% Resulting in Longer Routes (with respect to length)	91.5%	90.6%	89.9%
Avg. Miles Route is Shorter (when taking shorter route)	0.5	0.5	0.8
Avg. Miles Route is Longer (when taking longer route)	0.4	0.4	0.6
Trips with Departure Time Changes:			
% With Early Departure	16.5%	16.1%	15.2%
% With Late Departure	83.5%	83.9%	84.8%
Avg. Minutes Early Departure (when departing early)	6.0	5.5	6.3
Avg. Minutes Late Departure (when departing late)	6.3	6.4	6.8

It should be noted that in Cincinnati a commuter did not have too many viable alternate routes to choose from, as the network modeled was not dense and well connected owing to SmarTraveler covering only the major freeways and arterials. Although alternate routes may have existed, these were not modeled, as travel time information was not available on those routes from SmarTraveler. Hence, in this study the A95 commuter was more likely to improve his trip performance by changing his departure time rather than his route.

5.3.1.3 Comparative Analysis between the Small Sample Case and Resampled Cases

The most significant differences in the behavior of A95 commuters between the small sample and resampled cases occurred in the PM peak period. The A95 commuter modified his trip 9% more often in the small sample case than in the resampled cases. Departure behavior for the PM peak is compared between the two cases in Figure 5. When A95 commuters changed departure times, they left later than the regular time with a frequency of 96% for the small sample case compared to 85% for the resampled cases.

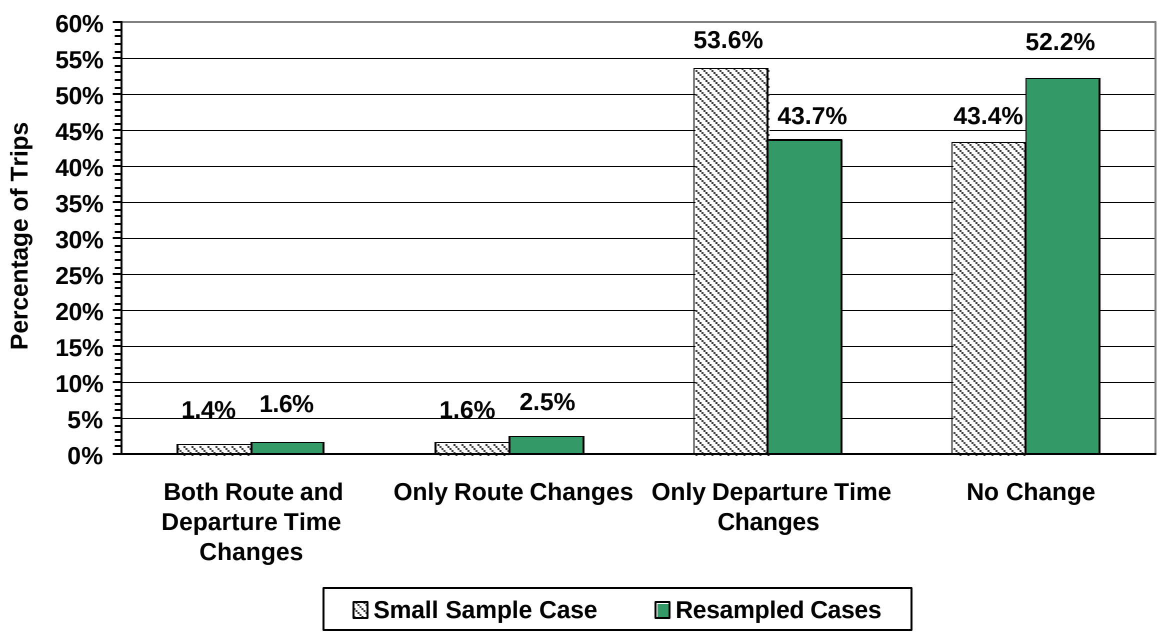


Figure 5. Comparison of A95 Commuter Behavior During the PM Peak Period

The training period is used in this study for the F95 and A95 commuters to get habituated and determine their regular routes and trip start times. The two commuters establish departure times for future trips using the travel times experienced during the training period as a benchmark for the levels of congestion on their routes. Hence, during the evaluation period if the expected travel times are lower than what was experienced during the training period, the A95 commuter will delay his departure time to minimize early arrivals at the destination. As A95 commuters experienced higher travel times and travel variability during the training period of the small sample case, they had a higher incidence of delayed departures in the small sample case than in the resampled cases in which the effect of incidents was reduced as days with high travel times and travel variation were more evenly distributed among the training and evaluation periods.

Thus, the key findings from this analysis can be summarized as follows:

1. Based on fifteen days of analysis, without resampling, we would have expected the notification-based traveler information service to recommend a change in 36% of the AM peak trips and 57% of the PM peak trips. The A95 commuter changed his departure time more often than his route due to limited route choices in the modeled network. During the PM peak periods the A95 commuter changed his departure time on 55% of the trips and routes on 3% of the trips. Due to high travel times and travel variability in the training period of the small sample case, caused by incidents, commuters allocated more travel time than was necessary. Hence, to minimize early arrivals, the A95 commuter left later than his regular time on 96% of the trips.

2. With experimental resampling, the notification-based traveler information service would have recommended changes in 35% of the AM peak trips and 48% of the PM peak trips. As observed for the case without resampling, the A95 commuter changed his departure time more often than his route. During the PM peak periods, the A95 commuter changed his departure time on 46% of the trips and routes on 4% of the trips. The effect of non-recurrent congestion was reduced with experimental resampling since days with incidents were more evenly distributed. Hence, the A95 commuter delayed his trips 85% of the time.

5.3.2 Trip Outcomes

This section discusses the experiences of the F95 and A95 commuters and assesses if pre-trip ATIS based on the SmarTraveler data stream would have improved commute reliability in Cincinnati.

5.3.2.1 Small Sample Case

The statistics for the small sample case is shown in Table 6. Pre-trip ATIS proved to be beneficial on an average for the whole day. A commuter using ATIS was able to reduce his early trips by 90% and increase his just-in-time arrivals by 14%. However, there was an increase in late trips by 13%. The F95 commuter was late on 1.6% of the trips while the A95 commuter was late on 1.8% of the trips. On average, the A95 commuter was late only by twelve seconds on days when late. For the small sample case, high trip times and variability during the training period caused F95 commuters to be highly conservative. They allocated more travel time than was necessary and determined a trip start time that would give them sufficient buffer for on-time arrivals. Hence, during the evaluation period, while the F95 commuters, who were unable to change their departure times, left early, the A95 commuters tended to delay their departure times to minimize early schedule delays thereby resulting in occasional late arrivals. Please note that the desired frequency of on-time arrivals for both commuters is 95%, which the A95 commuter was able to satisfy (98%) despite being late on 1.8% of the trips. When early, the F95 commuters were early by 5.9 minutes and when late, they were late by thirty seconds. In contrast, the A95 commuters experienced reductions in their early and late schedule delays by 31% and 60%, respectively. On an average, they were early by 4.1 minutes and late only by twelve seconds. The A95 commuter also experienced reductions in travel disutility cost by 18% and travel expenditure by 7%. However, the trip times experienced by the F95 and A95 commuters were the same.

Table 6. Average Statistics for Cincinnati for the Small Sample Case (7 November 2001 – 17 January 2002: 30 days)

Aggregate Trip Metrics	All Day		AM Peak		PM Peak	
	F95	A95	F95	A95	F95	A95
% Trips Early	13.4%	1.3%	14.5%	1.4%	33.2%	2.5%
% Trips Just in Time	85.0%	96.9%	83.1%	96.6%	65.6%	95.5%
% Trips Late	1.6%	1.8%	2.4%	2.0%	1.2%	2.0%
When Early, Avg Minutes Early	5.9	4.1	6.2	4.3	8.5	4.5
When Late, Avg Minutes Late	3.2	1.1	5.3	1.1	4.2	1.2
Small's Disutility Value	$1.60	$1.40	$1.70	$1.40	$2.10	$1.40
Travel Expenditure (minutes)	26.7	24.9	27.4	25.5	29.9	25.8
Trip Time (minutes)	20.8	20.8	21.3	21.3	21.4	21.4

During the AM and PM peak periods, use of pre-trip ATIS resulted in significant benefits. There were reductions of 90% and 93% in early arrivals during the AM and PM peak periods. Travel disutility cost reduced by 18% and 33% and just-in-time arrivals increased by 16% and 46%. Although, late arrivals reduced by 17% during the AM peak period, there was an increase in late trips during the PM peak period from 1.2% to 2.0%. It should be noted that the A95 commuter was late on an average by only twelve seconds in the PM peak period. As mentioned earlier, incidents during the training period made the F95 commuters more conservative than their A95 counterparts who on an average postponed their trips by 8.1 minutes (Table 4). During the AM peak period, the reductions in early and late schedule delays were 31% (reduced from 6.2 minutes to 4.3 minutes) and 80% (reduced from 5.3 minutes to 1.1 minutes), respectively. During the PM peak period, these reductions were 47% (reduced from 8.5 minutes to 4.5 minutes) and 71% (reduced from 4.2 minutes to 1.2 minutes), respectively. Travel expenditure decreased by 7% and 14%.

5.3.2.2 Resampled Cases

The statistics for each of the thousand resampled cases were aggregated for the AM peak, off peak and PM peak periods, and the averages over all thousand samples were determined, which are shown in Table 7. Benefits of pre-trip ATIS were high for the whole day as well as during the peak periods. During the day on an average, the F95 commuters arrived early on 9% and late on 3% of the trips. They were just-in-time 88% of the time. In comparison, the A95 commuters were early 1%, just-in-time 97% and late 2% of the time. Use of pre-trip ATIS resulted in a reduction of early and late trips by 87% and 30%, respectively, and an increase in just-in-time arrivals by 10%. When early, the F95 commuters were early by 5.3

minutes and when late, they were late by 3.8 minutes. The A95 commuters experienced reductions in their early and late schedule delays by 23% and 68%, respectively. On an average, they were early by 4.1 minutes and late by 1.2 minutes. The travel disutility cost reduced by 13% and travel expenditure by 5%. A nominal reduction of less than 1% was observed in trip time.

During the AM and PM peak periods, use of pre-trip ATIS resulted in reductions of 88% each in early arrivals, 35% and 47% in late arrivals, and 18% and 23% in travel disutility cost. Just-in-time arrivals increased by 14% and 26%. During the AM peak period, the reductions in early and late schedule delays were 27% (reduced from 5.9 minutes to 4.3 minutes) and 73% (reduced from 4.5 minutes to 1.2 minutes), respectively. During the PM peak period, these reductions were 34% (reduced from 7 minutes to 4.6 minutes) and 66% (reduced from 4 minutes to 1.4 minutes), respectively. Trip time reduced by less than 1% for both peak periods.

Please note that the benefits shown in Tables 6 and 7 are the averages computed for unitary trips between each origin and destination for a target arrival time. In reality, the number of trips varies greatly between origin-destination pairs, and from peak to off-peak period.

Table 7. Average Statistics for Cincinnati for the Resampled Cases (7 November 2001 – 17 January 2002: 30 days)

Aggregate Trip Metrics	All Day				AM Peak				PM Peak			
	F95		A95		F95		A95		F95		A95	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
% Trips Early	8.7%	2.6%	1.2%	0.2%	12.3%	4.4%	1.4%	0.3%	20.3%	5.9%	2.5%	0.3%
% Trips Just in Time	88.5%	1.8%	96.8%	0.1%	84.7%	3.3%	96.6%	0.2%	75.7%	4.0%	95.4%	0.2%
% Trips Late	2.7%	0.8%	1.9%	0.1%	3.0%	1.1%	1.9%	0.2%	4.0%	2.0%	2.1%	0.2%
When Early, Avg Minutes Early	5.3	0.4	4.1	0.1	5.9	0.6	4.3	0.1	7.0	0.7	4.6	0.1
When Late, Avg Minutes Late	3.8	0.7	1.2	0.1	4.5	1.6	1.2	0.2	4.0	0.7	1.4	0.1
Small's Disutility Value	$1.60	$0.01	$1.40	$0.00	$1.71	$0.04	$1.40	$0.00	$1.92	$0.05	$1.49	$0.03
Travel Expenditure (minutes)	26.3	0.3	25.1	0.1	27.3	0.5	25.7	0.1	28.8	0.6	26.5	0.2
Trip Time (minutes)	21.1	0.1	21.0	0.1	21.5	0.2	21.5	0.1	22.1	0.3	22.0	0.2

5.3.3 Comparative Analysis between the Small Sample Case and Resampled Cases

This section examines the differences in trip outcomes and mobility benefits of ATIS estimated for the two cases, to ascertain if there was indeed a need for experimental resampling. Figure 6 illustrates the arrival performances of F95 and A95 commuters during the PM peak period for the small sample case and resampled cases. If we had not applied experimental resampling and used only the small sample case for the analysis, we would have reported a 46% increase in just-in-time arrivals with ATIS (increased from 66% to 96%), and a 93% decrease in early arrivals (decreased from 33% to 3%). Note that these conclusions would have been based on results for only 15 days. Expecting higher travel times and variability, F95 commuters departed early to leave a large buffer for on-time arrivals, consequently arriving more than ten minutes prior to their desired arrival times. This also resulted in lesser number of late arrivals (1%) in comparison to the resampled cases. On applying the experimental resampling technique we saw slightly lower user benefits with ATIS. There was an increase in just-in-time arrivals by 26% and a reduction in early arrivals by 88%. Late arrivals reduced by 48%. Hence, if we had used the existing data set without applying resampling techniques overestimation of user benefits with ATIS.

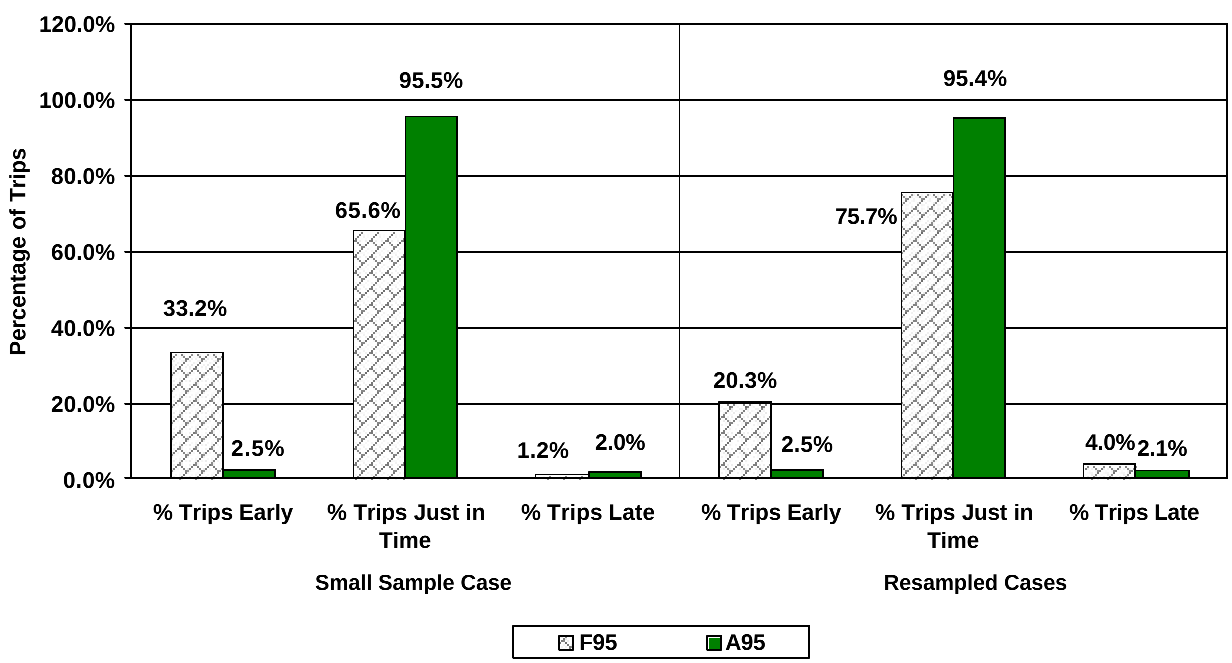


Figure 6. Comparison of Arrival Performances during the PM Peak Period

With ATIS, the percentage of trips with early, just-in-time and late arrivals are nearly the same for both the small sample case as well as the resampled cases. This is because the objective of A95 commuters is to arrive on time 95% of the time either by not deviating from their habitual behavior or by changing their departure time or route or both, according to the reported traffic conditions. Hence, regardless of the travel times experienced during the training period, the A95 commuter can react to the current reported conditions and improve his trip performance. The A95 commuter only uses the training period to determine a habitual route and trip start time, and is not forced to adhere to it unlike the F95 commuter whose performance during the evaluation period is highly dependent on the traffic conditions during the training period since he does not react to the current traffic conditions.

Figures 7 to 9 show the just-in-time arrivals, on-time arrivals and travel disutility costs for F95 commuters during the PM peak period for the small sample case and each of the resampled cases. For the resampled cases, the percentage of just-in-time arrivals range from 62% to 85% (Figure 7) and on-time arrivals vary from 86% to 99.5% (Figure 8). Figure 9 shows similar disparity among the resampled cases in the travel disutility costs experienced by F95 commuters. The disutility costs varied from $1.80 to $2.10. These figures illustrate how results can fluctuate based on the case chosen for analysis due to the small sample size. Thus, had we used a single small sample to generalize commuter experiences in Cincinnati, we would have made inaccurate observations. By resampling we are emulating a larger sample, which potentially reduces the bias due to a small sample.

By applying the experimental resampling technique we saw ATIS users in Cincinnati reduce their travel disutilities by $0.31 per trip during the AM peak period and $0.43 per trip during the PM peak period. In comparison, ATIS users in Washington, DC experienced reductions in travel disutility of $0.24 per trip during the AM peak period and $0.95 per trip during the PM peak period; in Minneapolis, the corresponding benefits were $0.09 per trip and $0.46 per trip, respectively (*1*). Figure 10 graphically depicts the commuter benefits experienced with ATIS plotted against Texas Transportation Institute's (TTI) Congestion Index factors for the three cities. From the figure it can be seen that despite a relatively lower congestion index, during the AM peak ATIS users in Cincinnati garnered similar travel disutility benefits as ATIS users in Washington. Likewise, during the PM peak, ATIS users in

Cincinnati and Minneapolis/St. Paul had nearly identical benefits. Although the TTI congestion index is strongly correlated with variability, ATIS benefits are affected more by the day-to-day variability than by the magnitude of peak period congestion (*1*). On measuring travel time variability in these cities, shown in Table 8, we find that the variability in Cincinnati during the AM peak period is comparable to that in Washington, and higher when compared to Minneapolis. Therefore, in the AM peak period, potential ATIS benefits are comparable in Cincinnati and Washington, and lower in Minneapolis/St. Paul.

During the PM peak period, the trip times and travel variability were lower in Cincinnati in comparison to Washington. Hence, the disutility benefits were lower in Cincinnati. However, Cincinnati and Minneapolis/St. Paul experienced similar travel variability, and consequently, their benefits were also the same. Thus, although Washington and Minneapolis/St. Paul had higher congestion indices, higher trip times and travel time variability in Cincinnati during the study period, resulted in disutility benefits analogous to those experienced in other two cities.

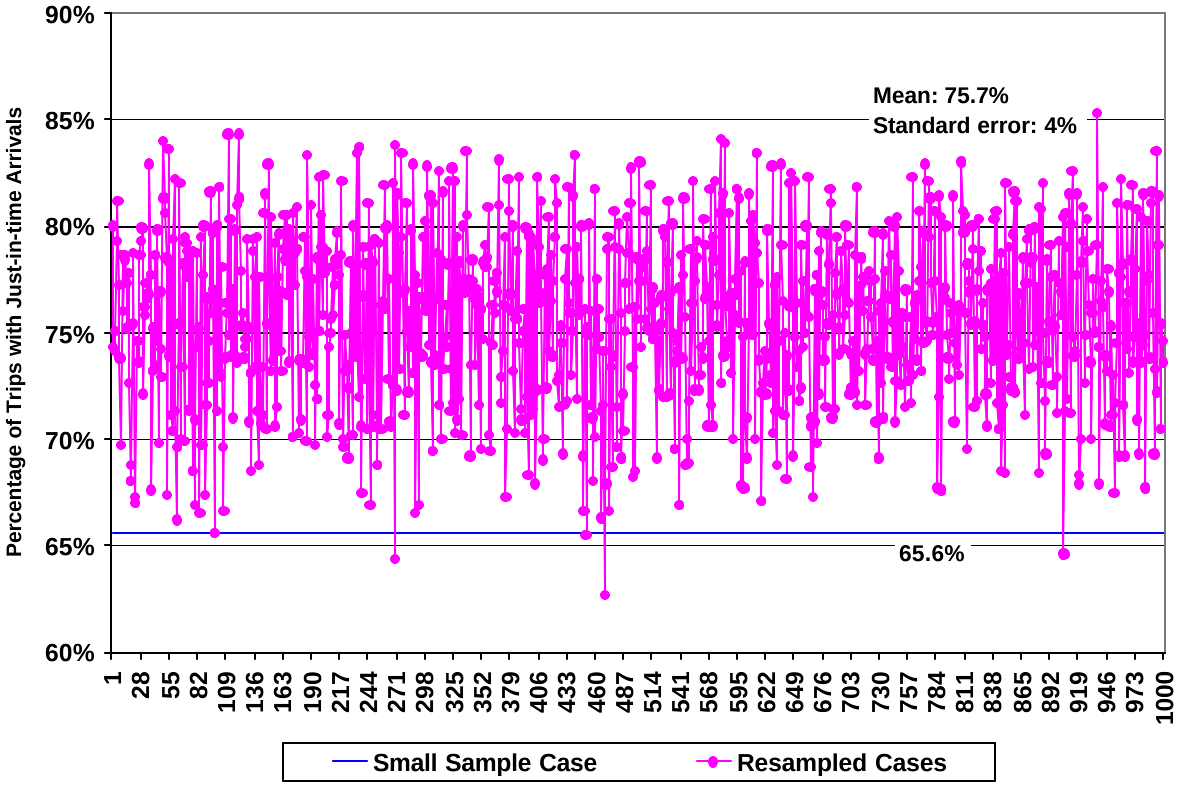


Figure 7. Comparison of Just-In-Time Arrivals of F95 Commuters during the PM Peak Period

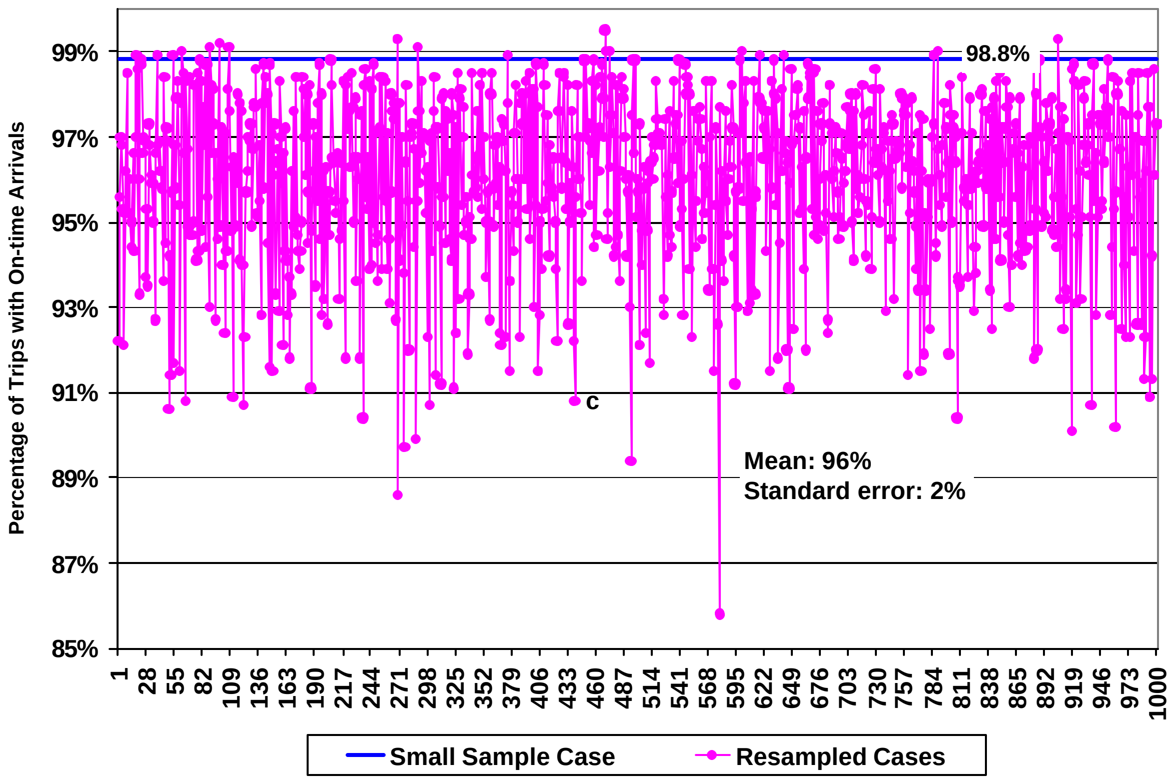


Figure 8. Comparison of On-Time Arrivals of F95 Commuters during the PM Peak Period

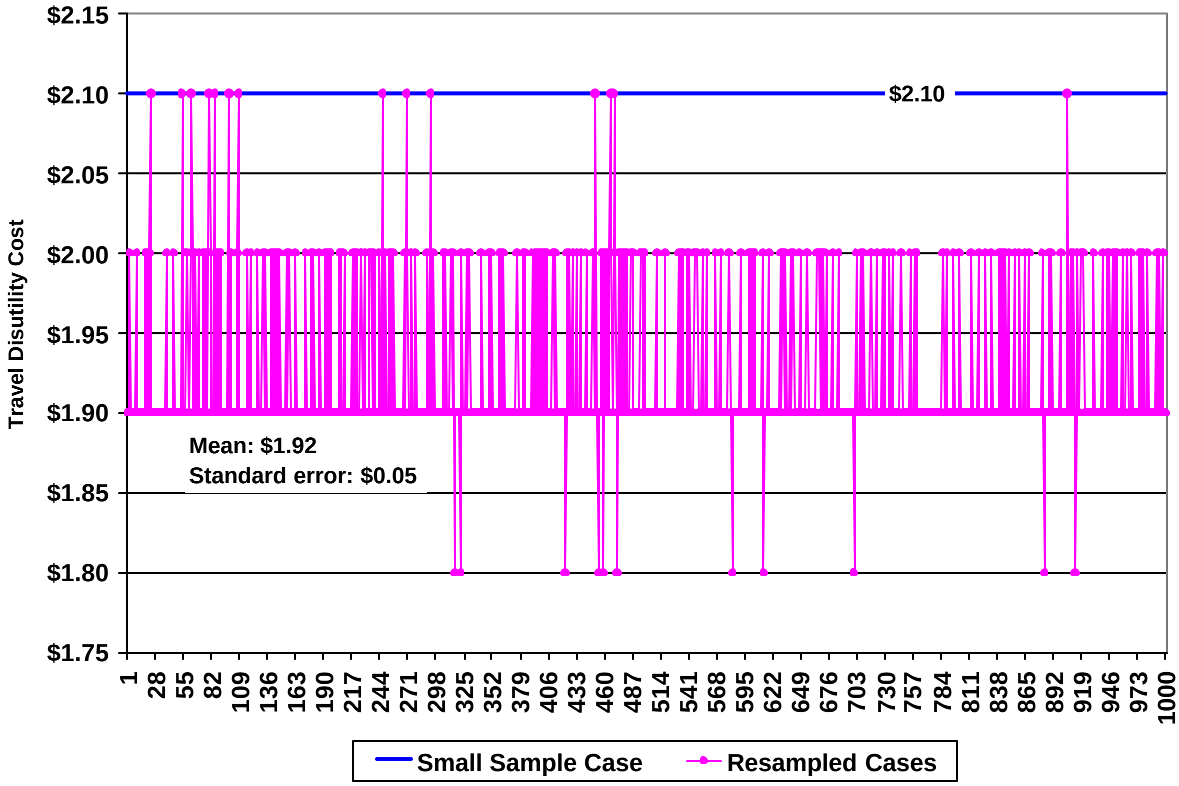


Figure 9. **Comparison of Travel Disutility Costs for F95 Commuters for PM Peak Period**

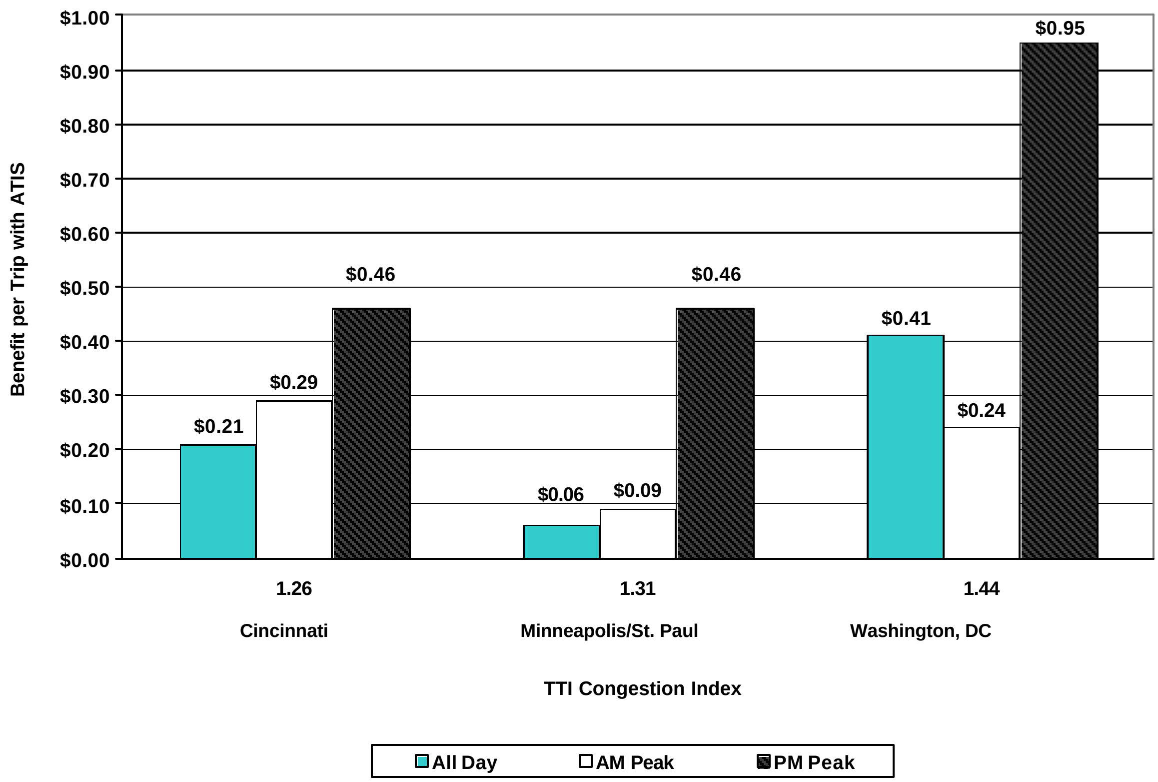


Figure 10. Comparison of Travel Disutility Benefits in Cincinnati, Washington, DC and Minneapolis/St. Paul

Table 8. Comparison of Average Trip Times and Travel Time Variability in Cincinnati, Washington, DC and Minneapolis/St. Paul

City	Trip Durations (minutes)		Travel Time Variability (minutes)	
	AM Peak	PM Peak	AM Peak	PM Peak
Washington, DC	33.6	34.5	2.1	2.9
Minneapolis/St. Paul	18.6	22.5	1.6	2.4
Cincinnati	22.2	23.1	2.2	2.3

The key results from this analysis are:

1. Without applying experimental resampling techniques in this study, the small data set would have distorted judgment and resulted in biased conclusions about the efficacy of traveler information service in Cincinnati. User benefits with ATIS were overestimated. Commuter expectations of high travel times and travel variability were skewed as a result of limited data. For the small sample case, use of ATIS in the PM peak period resulted in a 46% increase in just-in-time arrivals and a 93% decrease in

early arrivals. Travel disutility cost decreased by 33%. With experimental resampling, as days with high travel times and travel variability were not concentrated in the training period, we were potentially able to reduce the effect of non-recurrent congestion and consequently, the bias due to limited data. Experimental resampling showed that use of pre-trip ATIS in the PM peak period resulted in an increase of just-in-time arrivals by 26% and a decrease in early arrivals by 88% and travel disutility costs by 23%.

2. Pre-trip ATIS proved to be highly beneficial in Cincinnati. On an average, commuters using ATIS experienced reductions in travel disutility cost by 13%, late arrivals by 30%, early arrivals by 87%, and early and late schedule delays by 23% and 68%, respectively. Just-in-time arrivals increased by 10%. Travel disutility benefits were comparable to that experienced in Washington and Minneapolis/St. Paul due to high travel time variability and trip times.

Our premise in the initial phase was that since with experimental resampling we were emulating a larger sample, we would be able to reduce the potential bias in the small sample and make more reliable inferences about traveler information in Cincinnati. However, we were unable to support this premise due to lack of sufficient data at the time the study was initiated. Problems with the data archival process were later resolved and we were able to resume data collection. The next section describes the second phase of the study, wherein we conducted a larger study using travel time data for 154 weekdays to validate our claim that experimental resampling is beneficial when data are limited.

6 PHASE II: VALIDATION ANALYSIS

In the initial phase of the study we hypothesized that by experimental resampling we had a more reliable estimate of benefit. In order to corroborate our assumption that resampling can be beneficial in performing a robust analysis when only limited data are available, we conducted a larger study for the period extending from 2 September 2002 to 30 May 2003 covering 154 weekdays to more precisely determine the value of experimental resampling.

The objective of our analysis was to examine if experimental resampling can be beneficial in providing reliable estimates of user benefits, and to substantiate these results by comparing them to those estimated from a simple time series application of HOWLATE on a large data sample covering nearly a year.

6.1 Creating Data Samples

We applied the HOWLATE methodology to three cases: (i) a large sample case, consisting of 154 weekdays, (ii) a small sample case, consisting of only 30 weekdays, and (iii) a thousand resampled cases, created from the small sample case of 30 weekdays. This section presents the method used to create each of the three cases.

6.1.1 Large Sample Case

The large sample case was defined as having a training period from 3 September 2002 to 22 November 2002, and an evaluation period from 25 November 2002 to 30 May 2003. The training period consisted of 45 weekdays, while the evaluation period consisted of 109 weekdays.

6.1.2 Small Sample Case

Despite having a large stream of data that spanned nearly a year, we constructed a small sample case to replicate the problem of lack of data encountered in phase 1 of the study. The small sample case was defined as having a training period from 3 September 2002 to 23 September 2002, and an evaluation period from 24 September 2002 to 1 November 2002. The training and evaluation periods each consisted of 15 weekdays.

6.1.3 Resampled Cases

As described previously in Section 5.1.2, a random sample of size 15 was created from the small sample of 30 days without replacement. This random sample made up the training period and the remaining 15 days made up the corresponding evaluation period. A thousand training periods were created by repeatedly sampling from the small sample of 30 days, and evaluation periods were created from those days in the small sample that were not in the corresponding training period. The period of days that were randomly sampled extended from 3 September 2002 to 1 November 2002.

6.2 Travel Time Characteristics

This section compares the network travel times and variability for the three data samples, and presents the hypothesis for each case. Table 9 shows the average network travel times and variability for the training and evaluation periods for the three data samples.

6.2.1 Large Sample Case

During the AM peak period, SmarTraveler reported incidents on 39 days in the training period and 64 days in the evaluation period. In the PM peak period, there were incidents on 41 days in the training period and 67 days in the evaluation period. Table 9 shows that the training period had higher travel times in both peak periods. However, the evaluation period had significantly higher variability than the training period in the AM peak period than in the PM peak period. Due to the high travel times in the training period, we hypothesized that the non-ATIS user would be conservative and experience high early delays. However, the frequency of early arrivals would be greater in the PM peak period, due to the higher variability seen in the AM peak period of the evaluation period.

6.2.2 Small Sample Case

During the AM peak period, incidents were reported by SmarTraveler to last longer during the evaluation period than during the training period. The evaluation period experienced higher travel times and travel variability compared to the training period (Table 9). Consequently, the non-ATIS user, who habituated during the training period, was less informed of the variability in travel time. Given that, we hypothesized that the non-ATIS user would be less conservative and arrive at his destination later than his ATIS counterpart. In the PM peak period, the duration of incidents was longer during the training period than during the evaluation period. This is further evidenced by the higher travel times and variability in the training period than in the evaluation period. Hence, we expected the non-ATIS user to be more conservative than was necessary and leave a larger buffer for on-time arrivals, thereby resulting in frequent early arrivals of more than ten minutes.

6.2.3 Resampled Cases

It was expected that by experimental resampling, the confounding effects of non-recurrent congestion that occurred in the 30 days that were used to create the thousand resampled cases would be either reduced or magnified to reproduce the effects of incidents that occurred during the larger period of 154 weekdays. For the resampled cases, our hypothesis was that the non-ATIS user would have similar trip experiences as that observed for the large sample case.

For all three sample cases, the ATIS user was expected to have lower travel disutilities and delays. Additionally, for the resampled cases, our hypothesis was that the ATIS user would have similar mobility benefits as that experienced for the large sample case.

6.3 Validation Results

Simulated yoked trials were conducted between the F95 and the A95 commuters for each origin-destination pair. As in the previous phase, five Monte Carlo realizations were conducted for each day in the evaluation period for target arrival times at 15-minute intervals between 6:30 AM and 6:30 PM. The training of commuters and evaluation of ATIS were performed for the small and large sample cases, and the 1000 cases created by experimental resampling the small sample of 30 weekdays. Statistics were computed for each of the cases, and for the resampled cases, the experimental resampling standard errors for each performance measure were also determined.

This section discusses the experiences of the F95 and A95 commuters for the large sample case and compares the estimated ATIS benefits with those from the small sample case and the resampled cases. The goal here is to validate our hypothesis that a commuter using ATIS will have higher reliability and other mobility benefits than a commuter who does not make use of ATIS. We also corroborate our claim that when compared to a limited data sample, experimental resampling should yield more robust and valid outcomes, and would be more representative of user benefits estimated from a larger sample covering a year.

When the HOWLATE methodology was applied to the large sample, we found commuter benefits from ATIS to be significantly high for the whole day and during the peak periods. Table 10 shows the statistics for the large sample case. During the day on average, use of ATIS reduced disutility cost from $1.73 to $1.40 (19% reduction), late schedule delay from 6.5 minutes to 2 minutes (69% reduction), early schedule delay from 6.2 minutes to 4.3

minutes (30% reduction), late trips from 2.7% to 1.7% (relative reduction of 37%), early trips from 14% to 1.3% (relative reduction of 91%), and trip time from 21.5 minutes to 21.4 minutes (0.2% reduction). Late trips are an indicator of trip reliability. When a commuter is late 2.7% of the time, his trip reliability is 97.3%. During the AM and PM peak periods, use of ATIS resulted in relative reductions of 21% and 30% in travel disutility cost, 67% and 69% in late schedule delay, 33% and 44% in early schedule delay, 42% and 30% in late trips, and 91% and 93% in early trips. These reductions are relative to the trip experiences of the non-ATIS user. Trip time reductions were nominal for both peak periods. The marginal benefits from use of ATIS with respect to trip time reduction are in agreement with numerous field operational studies (*12, 20, 21*).

The statistics for the small sample are shown in Table 11. The results indicate that even for the small sample ATIS was found to be highly beneficial. During the day on average, use of ATIS resulted in a reduction of early trips from 9.4% to 1.7%, late trips from 5.2% to 2%, early schedule delay from 5.3 minutes to 4.2 minutes, late schedule delay from 4.9 minutes to 1.4 minutes, disutility cost from $1.73 to $1.42, and trip time from 21.7 minutes to 21.6 minutes. During the AM and PM peak periods, use of ATIS resulted in relative reductions of 17% and 29% in disutility cost, 55% and 63% in late schedule delay, 18% and 38% in early schedule delay, 60% and 62% in late trips, and 81% and 87% in early trips. Reductions in trip time were negligible.

The statistics for each of the resampled cases were aggregated for the AM peak and PM peak periods, and the averages over all 1000 samples were determined, which are shown in Table 12. Experimental resampling standard errors were computed for all metrics. During the day on average, A95 commuters experienced reductions in their disutility cost by 19%, late schedule delays by 72%, early schedule delays by 30%, late and early trips by 40% and 88%. The ATIS commuter experienced similar high benefits during the peak periods. During the AM and PM peak periods, use of ATIS resulted in reductions of 18% and 29% in travel disutility cost, and 57% and 70% in late schedule delay.

Table 9. Comparison of Average Network Travel Times and Travel Variability for the Three Data Samples

Peak Period	Small Sample Case (30 days)				Resampled Cases (30 days)				Large Sample Case (154 days)			
	Training		Evaluation		Training		Evaluation		Training		Evaluation	
	Average (minutes)	Standard Deviation (minutes)	Average (minutes)	Standard Deviation (minutes)	Average (minutes)	Standard Deviation (minutes)	Average (minutes)	Standard Deviation (minutes)	Average (minutes)	Standard Deviation (minutes)	Average (minutes)	Standard Deviation (minutes)
AM Peak	18.0	1.0	18.4	1.7	18.1	1.7	18.1	2.2	18.0	1.7	17.8	2.5
PM Peak	18.4	1.7	18.1	1.5	18.4	1.9	18.0	1.6	18.3	2.3	18.0	2.4

Table 10. Average Statistics for Cincinnati for the Large Sample Case (3 Sep 2002 – 30 May 2003)

Aggregate Trip Metrics	All Day				AM Peak				PM Peak			
	F95		A95		F95		A95		F95		A95	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
% Trips Early	14.0%	2.5%	1.3%	0.6%	16.8%	5.4%	1.6%	1.6%	31.4%	9.5%	2.4%	0.8%
% Trips Just in Time	83.3%	3.3%	97.0%	1.3%	80.1%	6.6%	96.6%	2.8%	65.6%	6.9%	95.5%	1.7%
% Trips Late	2.7%	3.8%	1.7%	0.7%	3.1%	8.3%	1.8%	1.6%	3.1%	5.6%	2.2%	1.2%
When Early, Avg Minutes Early	6.2	0.3	4.3	0.1	6.6	0.7	4.5	0.3	8.4	1.2	4.7	0.2
When Late, Avg Minutes Late	6.5	2.9	2.0	1.1	6.2	2.0	2.1	0.9	5.6	2.8	1.7	1.1
Dollar-Valued Disutility Cost	$1.73	$0.26	$1.40	$0.09	$1.82	$0.46	$1.44	$0.18	$2.18	$0.27	$1.52	$0.13
Travel Expenditure (minutes)	27.5	0.4	25.7	0.8	28.3	0.7	26.2	1.6	31.0	0.4	27.4	1.5
Trip Time (minutes)	21.47	0.8	21.43	0.8	21.83	1.6	21.80	1.5	22.87	1.7	22.76	1.5

Table 11. Average Statistics for Cincinnati for the Small Sample Case (3 Sep 2002 – 1 Nov 2002)

Aggregate Trip Metrics	All Day				AM Peak				PM Peak			
	F95		A95		F95		A95		F95		A95	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
% Trips Early	9.4%	2.6%	1.7%	0.6%	8.5%	4.3%	1.6%	0.4%	25.8%	9.2%	3.3%	1.1%
% Trips Just in Time	85.4%	4.0%	96.3%	1.1%	85.1%	4.1%	95.8%	1.3%	68.8%	4.1%	94.6%	1.9%
% Trips Late	5.2%	5.8%	2.0%	0.6%	6.4%	6.1%	2.6%	1.2%	5.4%	8.6%	2.0%	1.4%
When Early, Avg Minutes Early	5.3	0.4	4.2	0.1	5.3	0.7	4.3	0.1	7.7	1.3	4.8	0.2
When Late, Avg Minutes Late	4.9	1.9	1.4	0.6	3.8	1.6	1.7	0.8	4.5	1.9	1.7	0.9
Dollar-Valued Disutility Cost	$1.73	$0.28	$1.42	$0.08	$1.78	$0.26	$1.48	$0.10	$2.18	$0.30	$1.55	$0.16
Travel Expenditure (minutes)	26.8	0.4	25.8	0.9	27.3	0.3	26.4	1.0	30.5	0.4	27.7	1.9
Trip Time (minutes)	21.72	1.0	21.63	0.9	22.34	1.1	22.23	1.0	23.12	2.1	22.99	1.9

Table 12. Average Statistics for Cincinnati for the Resampled Cases (3 Sep 2002 – 1 Nov 2002)

Aggregate Trip Metrics	All Day				AM Peak				PM Peak			
	F95		A95		F95		A95		F95		A95	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
% Trips Early	13.8%	3.5%	1.6%	0.2%	15.2%	3.7%	2.0%	0.3%	26.3%	7.0%	3.1%	0.5%
% Trips Just in Time	82.8%	3.8%	96.2%	3.6%	81.6%	4.0%	95.6%	3.5%	69.2%	5.5%	94.5%	3.5%
% Trips Late	3.2%	1.2%	1.9%	0.2%	3.1%	1.2%	2.2%	0.3%	4.3%	2.2%	2.2%	0.3%
When Early, Avg Minutes Early	6.0	0.5	4.2	0.0	6.3	0.5	4.4	0.1	7.9	0.9	4.7	0.1
When Late, Avg Minutes Late	4.5	0.6	1.3	0.1	3.3	0.5	1.4	0.2	4.9	1.4	1.5	0.2
Dollar-Valued Disutility Cost	$1.74	$0.07	$1.41	$0.05	$1.78	$0.08	$1.46	$0.06	$2.18	$0.12	$1.54	$0.06
Travel Expenditure (minutes)	27.4	1.1	25.6	1.0	28.1	1.1	26.3	1.0	30.7	1.4	27.6	1.1
Trip Time (minutes)	21.52	0.8	21.46	0.8	21.99	0.8	21.97	0.8	23.09	0.9	22.93	0.9

All three cases showed significant benefits from use of ATIS for the whole day and for both the peak periods. For A95 commuters, the trip outcomes are nearly the same for all three cases (Tables 10 to 12). This is because the objective of A95 commuters is to arrive on time 95% of the time either by not deviating from their habitual behavior or by changing their departure time or route or both, according to reported traffic conditions. Hence, regardless of the travel times experienced during the training period, the A95 commuter can react to the current reported conditions and improve his trip performance. The A95 commuter uses the training period to determine a habitual route and trip start time, and is not forced to adhere to it unlike the F95 commuter whose performance during the evaluation period is highly dependent on the traffic conditions during the training period since he does not react to the current traffic conditions. The differences in user benefits are due to the experiences of the F95 commuters rather than that of the A95 commuters. Please note that the benefits shown in Tables 10 to 12 are the averages computed for unitary trips between each origin and destination for a target arrival time. In reality, the number of trips varies greatly between origin-destination pairs, and from peak to off-peak period.

If we only had a small sample to estimate peak period benefits from use of ATIS for a year, as was the case with the initial study, and had not applied experimental resampling, estimation errors for the performance metrics (late and early arrivals, late and early schedule delays, dollar-valued disutility, and trip time) would have ranged from -15% to 32%. Estimation errors were determined as the differences between the benefits of ATIS, with respect to the six performance metrics, computed for the large sample case and the small sample case. Note that the intent here is to examine the extent of error in generalizing the benefits of ATIS for a typical commuter over a year using a small dataset. Analysis using the small sample showed that during the day on average, use of ATIS resulted in a reduction of early trips from 9.4% to 1.7% (Table 11). Thus, if we had used the small sample case to generalize the percent benefit from ATIS in reducing a commuter's early arrivals, we would have estimated it to be 82%. However, the large sample case showed that the percent benefit from ATIS in reducing early arrivals was 91% (reduced from 14% to 1.3%; Table 10). Thus, in this case our estimation error would have been -9%. By experimental resampling, the analysis showed that with ATIS early trips reduced from 13.8% to 1.6% (Table 12), resulting in a percent benefit of 87%. In this case, the estimation error would have been -4%, as the actual percent benefit was 91%. By applying experimental resampling to estimate commuter benefits for a year with respect to six performance metrics, estimation errors would have ranged from -14% to 18%.

Figures 11 and 12 compare the relative benefits with use of ATIS for the three cases in the AM and PM peak periods, respectively. The figures show the percentage increase in benefits with use of ATIS for each of the six metrics. From the figures it is evident that when compared to the small sample case, ATIS benefits with respect to all performance metrics computed for the resampled cases are better predictors of the benefits for the large sample. For example, Figure 12 shows that conducting the HOWLATE analysis for nearly a year showed that a non-ATIS user could reduce his late trips by 30% in the PM peak period by using ATIS. If we had only 30 weekdays to assess the value of ATIS, we would have overestimated the percentage reduction in late trips by twice the observed benefit for the year (62% reduction compared to 30% over the year). By resampling, we were able to reduce the overestimation by nearly a third (48% reduction in late trips through experimental resampling). Figure 11 shows that in the AM peak, the small sample showed that the ATIS user was able to reduce his trip by 0.5% from 22.3 minutes to 22.2 minutes (Table 11), while the large sample and the resampled cases showed that the ATIS user reduced his trip by 0.1%.

Thus, the key results from the validation analysis are:

1. When data were limited, experimental resampling proved to be an effective technique in providing reliable estimates. Figures 11 and 12 showed that when compared to the small sample case of 30 weekdays, resampled cases generated using the experimental resampling technique were better predictors of ATIS benefits for the large sample of 154 weekdays. If we had used the small sample to generalize ATIS benefits for commuters in Cincinnati over a year, estimation errors would have ranged from -15% to 32%. With experimental resampling of the small sample, we were able to reduce the estimation errors.

2. Pre-trip ATIS proved to be highly beneficial in Cincinnati. Our analysis of 154 weekdays of travel time data showed that on average, commuters using ATIS experienced reductions in travel disutility cost by 19%, late arrivals by 37%, early arrivals by 91%, and early and late schedule delays by 30% and 69%, respectively. Just-in-time arrivals increased by 17%.

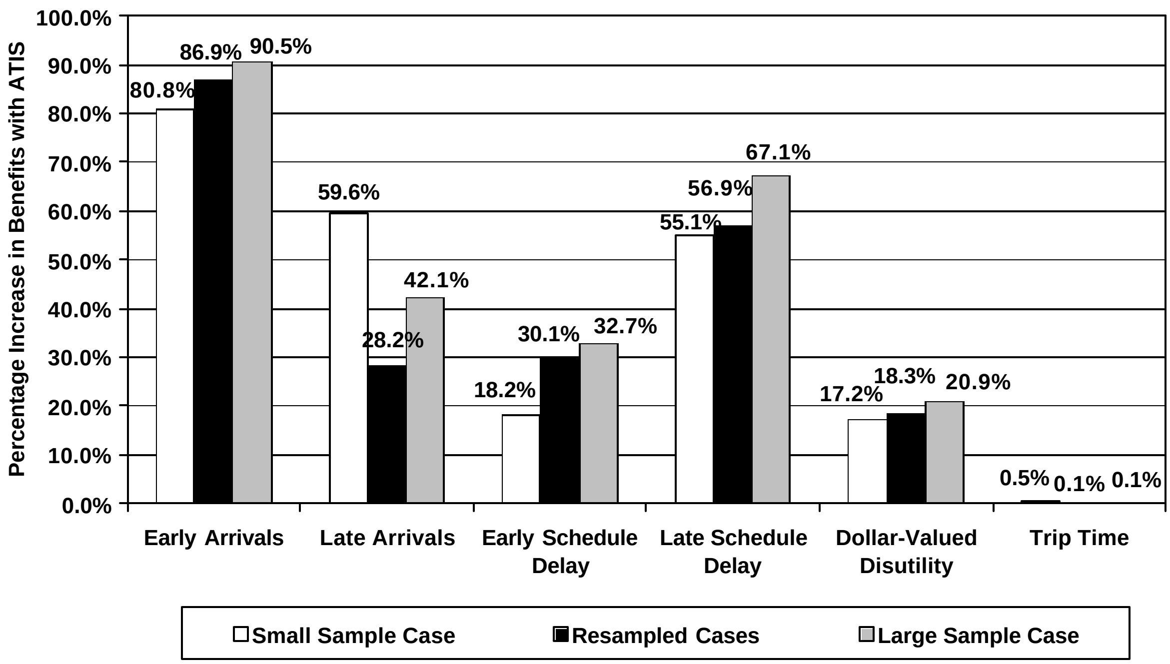


Figure 11. Comparison of ATIS Benefits during the AM Peak Period for the Three Data Sample Cases

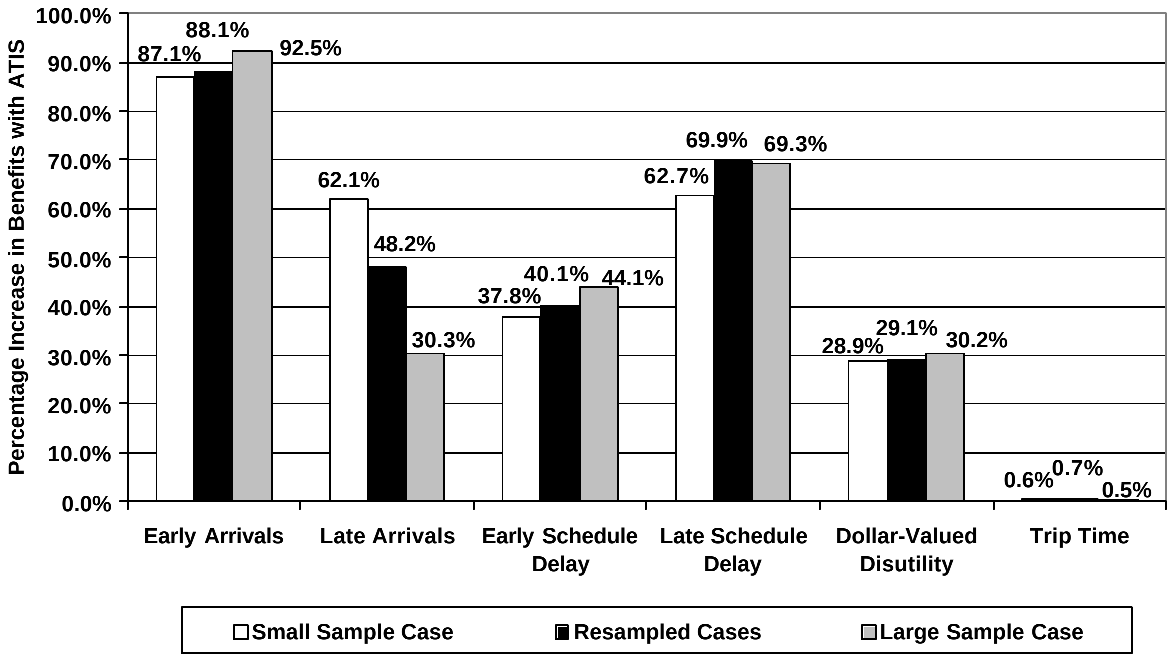


Figure 12. Comparison of ATIS Benefits during the PM Peak Period for the Three Data Sample Cases

7 CONCLUSIONS AND KEY FINDINGS

In this paper we presented an exploratory analysis of maximizing the value of small data sets in the evaluation of traveler information services using an analog of statistical resampling, defined in this paper as experimental resampling. We examined the potential benefits of a prospective notification-based pre-trip traveler information service in Cincinnati using experimental resampling to make reliable inferences. This technique proved to be effective in reducing estimation errors when generalizing traveler information service impacts for commuters in Cincinnati. When compared to the small data sample, resampled cases proved to be better predictors of user benefits estimated from a larger sample. Our analysis showed that a user of ATIS experienced lower disutility, late and early schedule delays and higher trip reliability than his non-ATIS counterpart.

In the evaluation of ITS systems, when data are limited, one can use statistical resampling techniques such as cross-validation, bootstrapping, jackknife, etc (*5*, *6*). Moreover, small data sets need not be worthless. Their value can be maximized through resampling. The studies mentioned earlier could potentially benefit from the application of bootstrapping. In this study we used experimental resampling as the intent was to create experimental samples to validate commuter experiences.

7.1 Hypotheses and Key Findings

This section examines our main study hypotheses presented in Section 1.2 and discusses our findings from the study.

Hypothesis: When compared to a limited data sample, experimental resampling should yield more robust and valid outcomes, and would be more representative of user benefits estimated from a larger sample.

Findings: Applying the HOWLATE process to a sample of 154 weekdays showed that on average, simulated ATIS users in Cincinnati would have increased their trip reliability by 37% and reduced trip disutility by 19%. If we had been constrained by availability of sufficient data, and had resorted to a direct application of the HOWLATE process to a small sample we would have overestimated trip reliability benefits to users of ATIS by 51%. Our analysis using a small sample of 30 weekdays showed that on average for the whole day, ATIS users would have reduced their late trips by 61%. Similarly, the small sample showed a reduction in trip disutility by 18.2%. In comparison, applying experimental resampling techniques in this study resulted in a more robust evaluation of benefits to users of pre-trip ATIS in Cincinnati. Resampling from the small sample of 30 weekdays showed that ATIS resulted in a reduction of late trips by only 39.6% and trip disutility by 19.3%, which were nearly identical to the benefits computed for the large sample case.

Hypothesis: A commuter making use of ATIS will experience higher trip reliability and other mobility benefits than a commuter who does not use ATIS.

Findings: Pre-trip ATIS proved to be highly beneficial in Cincinnati. Our analysis of 154 weekdays of travel time data showed that on average, commuters using ATIS experienced reductions in travel disutility cost by 19%, late arrivals by 37%, early arrivals by 91%, and early and late schedule delays by 30% and 69%, respectively. Just-in-time arrivals increased by 17%.

Hypothesis: The potential for ATIS benefit is greatest in cities with greatest network variability.

Findings: Our study showed that the potential for ATIS benefits during the peak periods was higher in the city of Washington ($1.19/trip) than in Cincinnati (37 cents/trip). This is in agreement with the TTI congestion index values for these two cities (1.44 for Washington and 1.26 for Cincinnati), and the observed network variability (2.5 minutes in DC compared to 2.1 minutes in Cincinnati). Conversely, the potential for ATIS benefit is higher in Cincinnati than in Minneapolis/St. Paul (6 cents/trip), while the TTI congestion index value for Minneapolis/St. Paul (1.31) is higher than that for Cincinnati. Although the TTI congestion index is strongly correlated with variability, ATIS benefits are affected more by the day-to-day variability than by the magnitude of peak period congestion (*1*). When measuring travel time variability from data sets in these cities we found that for the study period, the overall network variability in Minneapolis/St. Paul (1.6 minutes) was lower than that in Cincinnati (2.1 minutes). Thus, we conclude after all that the potential for ATIS benefit is greater in cities with greater day-to-day network variability.

REFERENCES

1. Jung, S., J. Larkin, V. Shah, A. Toppen, M. Vasudevan, and K. Wunderlich. On-time Reliability Impacts of Advanced Traveler Information Services (ATIS), Volume II: Extensions and Applications of the Simulated Yoked Study Concept. Prepared by Mitretek Systems for FHWA, 2002.

2. Schrank, D. and T. Lomax. The 2002 Urban Mobility Report. Prepared by Texas Transportation Institute and The Texas A&M University System, June 2002. http://mobility.tamu.edu/ums/study/final_report.pdf

3. Larkin, J. and K.E. Wunderlich. Archiving Retail ATIS Travel Times. Presented at the 11th Annual Meeting, ITS America, Miami, Florida, 2001.

4. Srinivasan, K. K., and H. S. Mahamassani. Dynamic Decision and Adjustment Processes in Commuter Behavior under Real-Time Information. Report SWUTC/02/167204-1. Office of the Governor of the State of Texas, General Revenue Funds, 1992.

5. Efron, B., R. J. Tibshirani. An Introduction to the Bootstrap. Chapman and Hall, New York, 1993.

6. Shao, J. and D. Tu. The Jackknife and Bootstrap. Springer-Verlag, New York, 1995.

7. Srinivasan, K., and H. S. Mahamassani. Role of Congestion and Information in Tripmakers' Dynamic Decision Processes: An Experimental Investigation. Presented at the 78th Annual Meeting of the Transportation Research Board, January 1999.

8. Estimation of Benefits of Houston TranStar. Prepared by Parsons Transportation Group and Texas Transportation Institute, February 1997.

9. Strathman, J. G., K. J., Dueker, T. Kimpel, R. L. Gerhart, K. Turner, P. Taylor, S. Callas, and D. Griffin. Service Reliability Impacts of Computer-Aided Dispatching and Automatic Vehicle Location Technology: A Tri-Met Case Study. Transportation Quarterly, Vol 54, No. 3, Summer 2000, pp. 85-102.

10. Wunderlich, K., M. Hardy, J. Larkin, and V. Shah. On-time Reliability Impacts of Advanced Traveler Information Services (ATIS): Washington, D.C., Case Study. Prepared by Mitretek Systems for FHWA, 2001. (Document #13335 on the ITS Electronic Document Library. http://www.its.dot.gov/welcome.htm

11. Shah, V., S. Jung, J. Larkin, V. Shah, A. Toppen, M. Vasudevan, and K. Wunderlich. On-time Reliability Impacts of ATIS, Volume III: Implications for ATIS Investment Strategies. Draft report, prepared by Mitretek Systems for FHWA, December 2002.

12. Inman, V., R. Sanchez, C. Porter, and L. Bernstein. TravTek Evaluation: Yoked Driver Study. Report #FHWA-RD-94-139. FHWA, U.S. Department of Transportation, October 1995.

13. Lappin, J. Advanced Traveler Information Services (ATIS): Who Are ATIS Customers? Presented at the ATIS Data Collection Guidelines Workshop, Scottsdale, AZ, February 2000.

14. Hardy, M., J. Larkin, V. Shah, and K. Wunderlich. Accuracy of Travel Time Estimates Obtained From Advanced Traveler Information Services. Proceedings of the 9th Annual Meeting of the Intelligent Transportation Society of America, Miami, Florida, 2001.

15. Kaufman, D., and R. Smith. Fastest Paths in Time-Dependent Networks for Intelligent Highway Systems Application, *IVHS Journal*, Vol. 1, No. 1, pp. 1-11, 1991.

16. Srinivasan, K., and H. Mahamassani. Dynamic Decision and Adjustment Processes in Commuter Behavior under Real-Time Information. Report SWUTC/02/167204-1. Office of the Governor of the State of Texas, General Revenue Funds, 1992.

17. Small, K., R. Noland, and D. Lewis. Valuation and Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation, NCHRP Report #431, National Academy Press, Washington, DC, 1999.

18. Jeannotte, K., P. Sankar, and D. Krechmer. Evaluation of the Advanced Regional Traffic Interactive Management and Information System (ARTIMIS). Presented at the 11th Annual Meeting of ITS America, June 2001.

19. Plutowski, M. E. P. Survey: Cross-validation in Theory and in Practice. Department of Computational Science Research, David Sarnoff Research Center, Princeton, New Jersey, USA.

20. JHK and Associates. Pathfinder Evaluation Report, California State Department of Transportation, 1993.

21. Schofer, J., Koppelman, R., Webster, R., Berka, S., and Peng, T. "Field Test of the Effectiveness of ADVANCE Dynamic Route Guidance on a Suburban Arterial Street Network," from The ADVANCE Project—Formal Evaluation of the Targeted Deployment, Volume II. FHWA, U.S. Department of Transportation, Washington, DC, 1997.