WISCONSIN TRAFFIC OPERATIONS & SAFETY LABORATORY
UNIVERSITY OF WISCONSIN-MADISON
Department of Civil and Environmental Engineering

Date:	11/30/2008
To:	Marie Treazise, WisDOT Bureau of Highway Operations
From:	Peter Rafferty, Wisconsin TOPS Lab
Subject:	Ramp Metering Evaluation – Technical Memo #3 Analysis Methodology

This memo documents the methodology used to evaluate the operations at metering installations in Wisconsin. It begins by outlining the general approach then goes into further detail on the analysis and economic valuation.

Although transit or road capacity increases are historically very common subjects of economic analysis, the implementations of intelligent transportation systems (ITS) strategies have been receiving increasing attention as resources for capacity expansion become increasingly scarce. There exists a strong business case for lower cost strategies to more efficiently deliver transportation in ways that do not entail additional costly right-of-way or lane-miles. An example of an increasingly common ITS strategy for managing traffic flow in urban areas is ramp metering, wherein traffic signals placed on freeway on-ramps disperse platoons of vehicles or restrict flow to maintain higher throughput and more stable flow on the freeway.

In economic analyses of transportation projects, the value of travel time savings is often the most common benefit assessed and the most substantial component of benefits. A second common component is the value of reducing crashes, either severity or frequency or both. A third common component is change in operating costs, which as a benefit often turns out to be negative as travel speeds increase. A common reference for anybody undertaking a benefit cost analysis of a transportation facility is the American Association of State and Highway Transportation Official's Red Book, which focuses on just these three areas of benefits. Environmental measures are often the next set of benefits, though the current Federal Highway Administration (FHWA) guidance on economic analysis cites the uncertainty of valuing pollution and suggests it may best be left to qualitative mention or to other studies. Then perhaps other more esoteric measures may be included, depending on the study.

The ramp metering operations evaluation focuses on travel time, safety, and travel time reliability. Reliability, however, is a relative newcomer to economic analysis. For example, a book chapter on project appraisal by Ken Small in 1999 discusses travel time, safety, and environmental benefits, but there is no mention of reliability, although he himself had done considerable research on the value of reliability by that time.

3.1 METHODOLOGY

There are two general instances providing with and without metering information. The first is upon initial ramp meter start-up. The second arises from the many brief time intervals on the shoulders of metering periods upon a change in operating times.

The source for the operations data is the WisTransPortal transportation data hub. The volume-speed-occupancy (VSPOC) export tool provides the necessary data in a comma separated file format (CSV). For this project VSPOC is called on to generate tables of lane-specific five-minute volume, speed, and occupancy data for all weekdays 30 days prior to a change and another 30 days following. The entries are then coded by whether the meter was active or not, whether the loop is one of the mainline loops to be used in the analysis, and whether the day is a holiday or not, for filtering.

The data from the inductive loop detectors is checked for zero and null values, and the percentage of these observations is tallied for each location and day. If too many observations are missing, or if there is a pattern of missing values that would introduce bias, these are flagged and not included in the core analysis.

The initial statistical check of performance changes is a two-tailed unweighted t-test of significance for speed change. From this point, the data are processed to generate a weighted travel time mean, median, and 95th percentile travel time.

$$T = \frac{\sum_{i=1}^{P} \sum_{j=1}^{D} \sum_{k=1}^{L} \frac{x}{v_{ijk}} \cdot q_{ijk}}{\sum_{n=1}^{Q} q_n} \cdot 60$$

where
- T = segment travel time (minutes)
- P = five-minute time period
- D = day or date
- L = lane number or loop pairs, e.g., 1, 2, 3
- x = distance represented by the detectors, e.g., one mile
- q = five-minute lane volume
- v = five-minute aggregate speed measured at loop pair (miles per hour)
- Q = total observed volumes
- i, j, k, n = summation index variables

The weighted median and 95th percentile travel time are calculated in the same manner. The difference between 95th and 50th percentile (median) travel time is the measure of reliability applied here. Some additional background on reliability as a user benefit is included later in this memo.

Because there are thousands of observations at each location on each change date, the processing is achieved by a separate routine using the Stata analysis software. The results are summarized in a standard spreadsheet, and these are discussed in subsequent tech memos for this project.

The benefits focused on from an operational perspective include the difference in average travel time with and without the meter active and the difference in travel time reliability with and without the meter.

$$\Delta T = T_{with} - T_{without}$$

$$\Delta R = (T_{95} - T_{50})_{with} - (T_{95} - T_{50})_{without}$$

where *T* = segment travel time (minutes), subscript indicates percentile
R = reliability (minutes)

And the annual equivalent benefit is calculated by multiplying by the volume of traffic affected and scaling to one year.

$$Benefit = \frac{\Delta T + \Delta R}{60} \times \frac{V_{without} + V_{with}}{2} \times L \times P \times D \times VOT$$

where *T* = segment travel time (minutes)
R = reliability (minutes)
V = five-minute lane volume
L = number of lanes
P = number of five-minute periods, e.g., 30 minutes = 6 periods
D = number of days, i.e., 254 work days in one year
VOT = aggregate value of time ($/hr)

Further discussion is provided later in this memo of the consumer surplus framework implicit in the formula and the value of time to be applied.

3.1.1 Ramp Meter Startups and Changes in Operating Times

There are two with and without metering times evaluated. The first is upon initial metering turn-on. In that case, the full metering period is evaluated, for example, 6:30-8:30 AM. This is the far less frequent of the two situations. In most cases of an initial turn-on, there is either no data available prior to start up, or the period prior to turn on was during construction with other capacity impacts affecting travel. There are just a handful of these scenarios available.

The second source of with and without metering period is when a timing change is made. In this case, it may be a change in metering start time to 15 minutes earlier or later. Changes of less than 15 minutes are not evaluated. For example, a metering location may have been metered from 6:30-9:00 AM each weekday. On some date, this timeframe was extended to 6:00-9:00 AM. This evaluation looks at the period 6:00-6:30 AM for differences in operations before and after that date (without and with the meter active).

The initial site selection was made by sorting the list of ramp meters by those which were installed after VSPOC data became available, generally 1996 or later, depending on the location. For instance, if a site had VSPOC data on the mainline from 2000 on and then a ramp meter was installed in 2001, then this site was desirable because one could directly compare the effects of the installation of the ramp meter. In comparison, a site whose ramp meter was installed prior to VSPOC data availability would only be valuable to analyze its meter on/off timing changes.

There is also a question about how reliable the record of the change is. The files have written comments throughout that show the retiming process. Typically a “1070 Configuration” page is printed out and then the documentation is completed by the person making the timing changes. The recommended thresholds and on/off timings are not always followed; the person making the changes pencils in the adjustments on an analysis page. The files then show a printout to confirm the ramp meter timing changes. A subjective estimate of how accurate the timing change is included in the analysis and is referred to as a confidence level for that change. Only a relatively small subset of potential locations and dates were considered acceptable for evaluation based on this confidence. See tech memo #5.

3.2 ECONOMIC VALUATIONS

By estimating the benefits and costs of the system in monetary terms we can determine the net present value, benefit cost ratio, and the internal rate of return for any change or improvement. Benefits that are included in this assessment are time savings, operating costs, and crash reduction savings. Other potential benefits that are not considered in this assessment include travel time reliability and environmental costs such as noise and vehicle emissions. As noted earlier, WisDOT has no standard values for these types of economic measures, so this section of the paper discusses values to be applied.

Unless noted, all monetary values presented in this paper are in 2006 dollars. Where values from references varied from 2006, they were adjusted using the gross domestic product implicit price deflator or the consumer price index, whichever was appropriate. This deflator is sometimes applied in lieu of the consumer price index because its broader inclusion of non-consumer goods and services better reflects the implications of a government-provisioned congested corridor management strategy such as ramp metering.

An internal cost estimate for new metering installations shows approximately $82,000 for installation, including communications, administration, and tuning. Note that ramp metering operations are primarily controlled via a statewide traffic operations center, but this assessment does not estimate any joint or overhead costs for that facility and other related freeway management activity. This should be done, and it will affect the analysis, but it is well beyond the scope of this paper. For ongoing costs, an individual ramp meter incurs direct annual costs of around $1,500 per year for maintenance and $2,800 for operations staff. With six new meters on the study segment, the initial costs were $492,000 and annual costs are estimated to be $25,800 in 2006 dollars.

3.2.1 *Ramp Metering Benefits*

The benefits of ramp metering include savings from reduced crashes and reductions in travel time or operating costs. These benefits can also be negative, or disbenefits, which is common especially for operating costs due to increased travel and higher speeds. An additional benefit introduced here is travel time reliability, which is not trivial. The 2000 Twin Cities ramp meter shutdown evaluation, for example, found that reliability and emissions savings together accounted for nearly three-quarters of total benefits. Reliability measures are becoming more accepted, but this does warrant a separate discussion, below.

For purposes of economic evaluation, the change in consumer surplus is applied. This framework does not commonly appear in engineering analysis, but it is indispensible for estimating economic benefit. Furthermore, a chief advantage is that network effects off the affected segment need not be explicitly determined but are captured implicitly by the price elasticity of demand revealed through the change itself. For example, lowering the price of a commodity increases quantity consumed, but we need not estimate the reduction in consumption of all other goods to assess the net social benefit.

The benefits of crash reduction, travel time reduction, and operating cost savings are outlined below then applied to a previous study of US 45 metering benefits for illustration and validation. Following that is a discussion of travel time reliability, which is applied to a section of freeway near Madison for illustration discussion purposes.

3.2.1.1 Crash Reduction Benefit

Crashes may be categorized by severity, which is key to estimating more accurate costs. There are three commonly used approaches to divide crash types for purposes of applying economic values. The first has three categories: non-injury / property damage only, injury crashes, and fatal crashes. This is the oldest and simplest division and remains valuable where data are limited, for example in developing countries. The second division broadens injury crashes into three severity levels – A, B, and C – for a total of five categories, sometimes referred to as KABCO (K is fatal, O is other). This is common among state departments of transportation, and is used in most project- or program-level economic analysis.

The third division includes seven categories and is used by some insurance companies and national programs, including the National Highway Traffic Safety Administration. The division is called the abbreviated injury scale (AIS) or maximum injury severity (MAIS), and the categories are as follows:

AIS 0 = non-injury
AIS 1 = minor injury
AIS 2 = moderate injury
AIS 3 = serious injury
AIS 4 = severe injury
AIS 5 = critical injury
AIS 6 = fatal

Many studies and references provide costs under the three different categorizations, but none could be found that provided costs on more than one scale or provided a translation between scales. Different agencies of the United States Department of Transportation (US DOT) use each of the three categorizations. The data available here are in the five-category scale, so we limit discussion to those studies and references that provide costs along that scale. A summary from three representative sources is provided in Table 1.

Table 1. Crash Costs (in 2006 dollars)

Crash Type	FHWA	NSC	Mn/DOT
N – Property Damage Only	$ 2,570	$ 2,270	$ 4,400
C – Possible Injury	$ 24,400	$ 24,300	$ 30,000
B – Non-incapacitating Injury	$ 46,300	$ 51,000	$ 61,000
A – Incapacitating Injury	$ 231,400	$ 199,500	$ 280,000
F – Fatal Crash	$ 3,343,000	$ 3,953,000	$ 3,600,000

There exists in the literature a great deal of work on the cost of a fatal crash or the value of a statistical life. The figures commonly accepted in transportation and used in this assessment are based not on an older method of estimating such things as lost future earning, but on the economic willingness to pay principle. In addition to the cost of a fatal crash given above, other values estimated for a fatal crash range from $2.1 million to $4.5 million, in 2006 US dollars. For this assessment the average values from Table 1 and the other resources are applied, as follows:

N – Property Damage Only	$ 3,080
C – Possible Injury	$ 26,200
B – Non-incapacitating Injury	$ 52,700
A – Incapacitating Injury	$ 237,000
F – Fatal Crash	$ 3,500,000

3.2.1.2 Value of Time

The US DOT and others provide typical values of time, but the US DOT also recommends valuing local personal travel time using 50% of the local wage rate, and 100% of the wage for commercial traffic. The Milwaukee County average wage was $19.58 per hour. Other recommendations and estimates are shown in Table 2. The average values are used, and the limited scope of this assessment precludes developing more refined values of time or applying values of time on a more complex basis such as by time of day or by socioeconomic strata.

Table 2. Values of Time

Source	Auto	Commercial
USDOT (16)	$ 13.00	$ 21.00
Mn/DOT (11)	$ 10.46	$ 19.39
Transport Canada (14)	$ 11.09	$ 22.63

Milwaukee Wage Basis (17)	$ 9.79	$ 19.58
Average	$ 11.10	$ 20.70

The values of time and the operating costs discussed next are divided into an auto classification and commercial or heavy vehicle classification. Autos would include motorcycles, cars, light trucks; the commercial classification includes all vehicles with three or more axles. Classification counts on freeways in this area of Milwaukee indicated that traffic is composed of roughly 5.8% commercial or heavy vehicles. While other evidence suggests this figure may be lower during peak periods while meters are operating, that level of detail was not available to verify this.

Values of time per vehicle must also be adjusted for number of adult occupants in a vehicle. This analysis multiplies the average value given in the table above by the average auto occupancy in the region of 1.14 adults per vehicle. The 2001 National Household Travel Survey (NHTS) gives values ranging from 1.14 for work related trips to 2.05 for social/recreational trips. These values vary by day of week and time of day, but the NHTS indicates that the morning commuting period is among the lowest occupancy times. So for this analysis the conservative 1.14 persons per auto for value of time calculations is used. Commercial vehicle occupancy is assumed to be one person.

3.2.1.3 Operating Costs

Operating costs include fuel consumption, tire wear, maintenance, etc., and on a per mile basis are known to vary with travel speed and other operating conditions. Operating costs sometimes decrease with transport improvements, but in congested conditions this is generally not the case, and the operating costs therefore deduct from net benefits. Although benefit cost guidance generally includes operating costs, most do not include operating costs that are variable with speed. Transport Canada excludes operating costs from any economic calculation. Caltrans includes operating costs consisting largely of a fixed portion and a fuel consumption component that does vary with operating speed, and this analysis proceeds using that methodology.

The nationwide average gasoline price – excluding taxes – in 2000 was $1.11 per gallon. This value excludes federal and state taxes because taxes are regarded as transfers and are not to be included in economic evaluations. The federal gas tax has been 18.4 cents per gallon since 1997; the Wisconsin state gas tax was 26.4 cents per gallon in 2000. An illustration of the costs is shown in Figure 1. Costs for heavy vehicles – not shown on the graph - are over twice as great as for autos and are included in the analysis that follows.

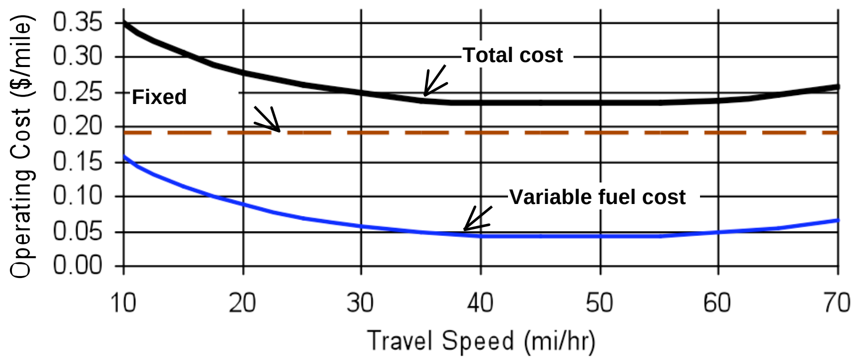

Figure 1. Automobile Operating Costs (2006 dollars)

3.2.2 US 45 Analysis

In this US 45 example, the analysis corridor is divided into five segments. Traffic volumes and speeds were collected on each segment and averaged based on the representative length. The data are presented as an average for the AM and PM peak periods across the weekdays on which the data were collected. The consumer surplus framework is a departure from the study, so the results will vary, although relatively slightly in this case. In the study, the total change in operating measures such as vehicle miles traveled (VMT) is reported rather than the change in consumer surplus. Figure 2 graphically illustrates this difference in a generic sense for a change in conditions from the dashed lines shown to the solid lines.

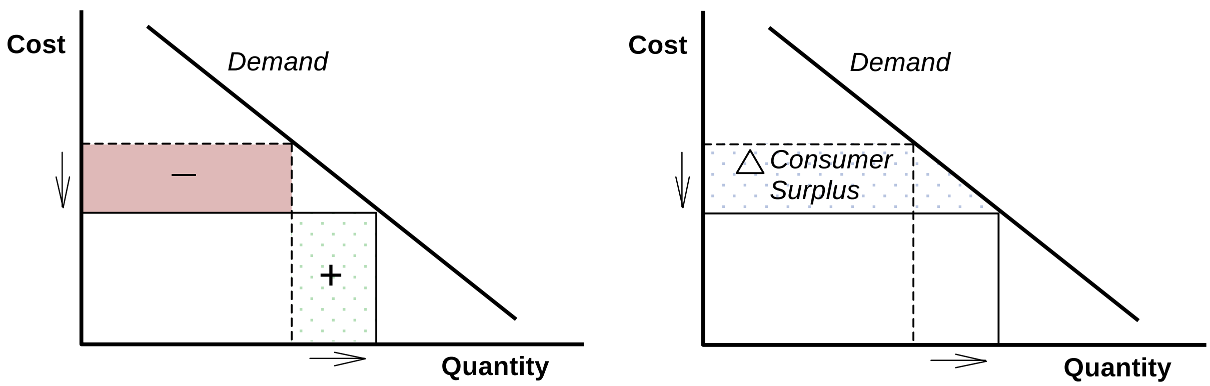

Figure 2. Change in Consumer Surplus

There was also no differentiation of normal versus generated traffic in the study. Although this is not uncommon in this type of analysis, it is usually for the stated reason that demand elasticity is low and the generated traffic effect is relatively minimal. For this analysis, this difference is considered and the rule of one half is applied to arrive at the shaded trapezoidal area shown in the image on the right in Figure 2.

3.2.2.1 Crash Reduction Benefits

The study presented a drop in crash rate from 2.98 crashes per million vehicle miles (MVM) to 2.60 crashes/MVM with metering in effect. There is no distinction made between crash severities, however, so as part of this analysis a distribution was estimated. All crashes were collected along the southbound analysis segment from 1999 to 2001, which includes a full year prior to and after the ramp metering installations. In total there were 920 crashes. However, the effects of ramp metering are only present while operating, so this crash data was filtered to include only the morning and afternoon 1.5-hour peak periods evaluated in the study. The distribution during this time was:

Severity	Total	Distribution
K	0	0.00%
A	2	1.19%
B	6	3.57%
C	54	32.14%
N	106	63.10%

The absence of fatal crashes relieves this analysis of accommodating very infrequent but very high value events in a statistically meaningful manner. Additional years of data may reveal a more significant fatal crash risk. Table 3. US 45 Crash Costs summarizes the estimated crash cost savings benefit from ramp metering. The values in the table represent weekday metering periods. The annual estimate is arrived at by multiplying the weekday average by 260 weekdays.

Table 3. US 45 Crash Costs

		Without Meters	With Meters
VMT (Daily, Metering Hours)		216,546	219,398
VMT (Annual MVM, Metering Hours)		56.30	57.04
Crash Rate (per MVM)		2.98	2.60
Expected Crashes per 100 MVM	K	0.00	0.00
	A	3.55	3.10
	B	10.64	9.29
	C	95.79	83.57
	N	188.02	164.05
Expected Cost per 100 MVM		$4,490,363	$3,917,767
Crash Cost Change (Annual)		-$324,506	

The total crash cost change is calculated by multiplying the difference in crash risk by the average of the without and with annual VMT, i.e., the area of the consumer surplus change – the trapezoid – illustrated in Figure 2. As the data indicate, there were no fatal crashes during metering hours at this location from 1999 through 2001, obviating the need for a sensitivity analysis on those values.

3.2.2.2 Travel Time Reduction Benefits

The change in travel speeds at the five analysis locations are available in the report and not repeated here. From each a segment travel time is calculated by dividing the segment distance by the speed. Then change in total vehicle-hours is calculated using the trapezoid calculation illustrated in Figure 2. The results are summarized in Table 4.

Table 4. US 45 Travel Time Costs

Location	Change in Total Vehicle-Hours (Daily)		
	AM		PM
0 - Waukesha Co Line	2.63		0.69
1 - Congress St	-43.11		0.06
2 - Center St	-9.25		-37.83
6 - Wisconsin Ave	-2.89		-57.18
8 - Belton RR	-2.61		-26.76
Subtotal	-55.23		-121.01
Ramps*	106.17		54.14
Total	50.94		-66.87
Total Both Periods (hrs)		-15.93	
Annual Estimate (hrs)		-4,143	
	Auto		Truck
Travel Time Cost Change	-$49,385		-$4,974
Total (annual)		-$54,359	

* The "Ramps" line shown in Table 4 comes directly from the engineering report and therefore does not explicitly reflect the change in consumer surplus. No other information on ramp operations or crash experience was provided. It is evident that the ramp delay almost entirely offsets the reduced mainline travel time.

The cost changes for Auto and Truck are calculated using the values for travel time from the previous section, including the average auto occupancy and truck percentage.

3.2.2.3 Operating Cost Benefit

Operating cost changes are calculated in a similar fashion. Operating costs per trip on each segment are a function of travel speed, segment length, and vehicle class. The costs per trip per segment are summarized in Table 5.

Table 5. US 45 Operating Costs, By Segment

Location	Without Meters AM	Without Meters PM	With Meters AM	With Meters PM
Automobiles and Light Trucks				
0 - Waukesha Co Line	0.83	0.82	0.82	0.81
1 - Congress St	1.08	1.14	1.12	1.14
2 - Center St	0.47	0.47	0.47	0.47
6 - Wisconsin Ave	0.57	0.57	0.57	0.56
8 - Belton RR	0.45	0.45	0.45	0.45
Heavy Trucks				
0 - Waukesha Co Line	3.16	3.09	3.11	3.08
1 - Congress St	3.75	4.25	4.13	4.25
2 - Center St	1.53	1.32	1.58	1.46
6 - Wisconsin Ave	1.83	1.45	1.84	1.58
8 - Belton RR	1.52	1.21	1.54	1.28

Change in operating cost is also calculated similarly to change in travel time. The change in cost per trip per segment is multiplied by the average of the without and with traffic volume, i.e., the trapezoid in Figure 2. The results are summarized in Table 6.

Table 6. US 45 Change in Operating Cost

Location	Daily Op. Cost Change AM	Daily Op. Cost Change PM
0 - Waukesha Co Line	-$58.99	-$14.28
1 - Congress St	$494.52	-$1.17
2 - Center St	$44.45	$70.77
6 - Wisconsin Ave	$12.21	$45.64
8 - Belton RR	$16.73	$30.54
Total	$508.91	$131.50
Total Both Periods	$640.41	
Annual Estimate	$166,508	

3.2.2.4 US 45 Benefit Summary

The total estimate of annual benefits from the three components evaluated here are:

Crash Costs	-$324,506
Travel Time Costs	-$54,359
Operating Costs	$166,508
Total	-$212,357

The most significant component is the $324 thousand reduction in crash costs. Next most significant is the $166 thousand increase in operating costs, which as expected is a disbenefit in this case. Travel time savings are often the most significant component, but in this analysis it amounts to just a $54 thousand reduction. Nonetheless, the estimated total benefit of the ramp meter installations in the first year is $212 thousand.

To estimate a net present value (NPV) or a benefit cost ratio, additional analysis years would need to be included, which they are not in the original study nor in this assessment. If that information were available, all values should be adjusted to constant dollars with a discount rate applied to all future values. For evaluations of this sort, the United States Office of Management and Budget used a 10% discount rate until 1992 when it dropped to 7%. Transport Canada continues to use 10%, while WisDOT in their economic analyses use a rate of 5%. While this may be low if considering opportunity costs and constrained budgets, and may favor more capital intensive projects, a sensitivity analysis could check a range of values.

In this situation it is highly plausible that as congestion grows on the corridor, the benefits from ramp metering will grow. For illustration, holding annual benefit constant at $212 thousand will return a conservative estimate for those economic measures. Assuming a simple 10-year analysis horizon and setting aside life cycle related accounting, given the values derived in this paper, and including a sensitivity check of 5% and 10% discount rates, the results are:

Discount Rate	5%	10%
NPV	$903,377	$594,830
B/C	2.37	2.01
IRR	36%	36%

The NPV and benefit cost ratio (B/C) would both be significantly lower with a higher discount rate, although the breakeven point in either case occurs just a few years after deployment. The third measure shown is an internal rate of return (IRR), a less common measure but useful as a comparison to the discount rate. The IRR in this case is substantially higher than the discount rate. All measures indicate that ramp metering here is a sound investment of tax payer dollars.

3.2.3 Travel time Reliability

This section outlines a review of measures of travel time reliability, dollar values estimated for those measures, and the application to benefit-cost assessment. Following that is a prototypical assessment and discussion of travel time reliability on a section of moderately congested freeway in Madison, Wisconsin. The objective is to illustrate the reasonableness and practicality of folding reliability benefit (or disbenefit) into the ramp metering assessment, or to any broader assessment of an ITS strategy.

Two key aspects of the travel time reliability discussion are how to measure it from whatever operations data are available, and how to value it with a monetary cost for use in economic analysis. The general state of the practice of evaluating or monitoring travel time reliability is summarized in the Economic Analysis Primer by the FHWA. They cite the common understanding that motorists value unexpected delay just as much – or more – than average delay or average travel time, which is validated in numerous economic studies, some of which are mentioned below.

3.2.3.1 Measuring Reliability

Measuring reliability hinges on numerous factors, including data availability, time increments, trip lengths, and common bases for comparison. Standard deviation is an older measure used for travel time reliability, as was the coefficient of variation, but these are often not used for a few reasons. For example, travel time distributions are not symmetric. There always exists a lower bound, whether legal or practical, and as congestion increases, the upper tail of the distribution grows. It is specifically this upper tail of increased uncertainty that travelers value when considering cost of travel. Then for purposes of presentation to non-technical audiences, some statistical measures are more difficult to conceptually internalize. More common practice now is to identify the 90^{th} or 95^{th} percentile travel time. Conceptually for public consumption, that would represent the worst day or two of a trip among typical work days in one month. This percentile value could then be used directly or compared to median or mean travel time and presented as a buffer index, travel time index, planning time index, and so forth. Unlike the straightforward value of time for average travel, the value of reliability must also specify the units, e.g., dollars per minute of standard deviation or per minute of difference between 90^{th} percentile and median or 50^{th} percentile time.

In terms of measuring travel times on a large scale, there exists little alternative to automated detection equipment, which in turn largely limits studies to freeways. Much of the work estimating the values people place on different aspects of travel is based on surveys. The priced SR 91 corridor in southern California is by far the most researched test bed for this, and there have been more recent efforts to tie the detector data into the discussion on that location. In Wisconsin there are no priced roads, so it is not so straightforward to estimate values via the disaggregate choice (logit) models common in the literature, but there may be opportunities to use detector data in more sophisticated ways to estimate the relative values placed on travel time, reliability, and departure time choice.

The research using detector data used the difference between 80^{th} and 50^{th} percentile travel times. It appears some jurisdictions and researchers also use the median value for comparison, while others use the mean. The FHWA guidance mentioned above suggests the mean, though it does not align with their own framework of being on time a given number of days out of 20 work days in a month, which is based on the median. For example, if planning for 90^{th} percentile travel time a traveler will be on time 18 days, planning on median time they will be on time 10 days, but planning on mean time they will be on time some number (or fraction) of days greater than 10, depending on variability. Also, as congestion and reliability worsens, the gap between 90^{th} percentile and mean grows more slowly than the gap between 90^{th} percentile and median.

An answer to this may be to test several measures of reliability in a logit specification for time and reliability values for the same set of data to see what fits best, or perhaps to see that different measures fit equally well. Essentially, rather than evaluate the same set of engineering calculations with multiple model specifications, there may be value in evaluating multiple measures of reliability using the same model specifications. In the end, it remains a matter of standardizing on one of them for the sake of consistent project appraisal among practitioners.

3.2.3.2 Valuing Reliability

An extensive amount of research on basic values of travel time has existed for quite some time. Much less had been analyzed regarding the value of time differentiated by travel conditions, level of service (LOS), or reliability. The last two or three decades have seen a rapid increase in more prevalent congested travel conditions in the United States, bringing greater attention to this issue. Earlier examples of research on the subject of travel time reliability date to 30-40 years ago and include conceptualizations of early departures and schedule delay.

Substantial work in estimating the value of travel time reliability has been done by Small and others in California. A benchmark effort was through the National Cooperative Highway Research Program (NCHRP) Project 2-18, which from 1991 to 1997 pushed development of highway user cost estimates and methodologies. A report for that project focused on key aspects of user costs related to travel time. One aspect studied is the potentially greater value of travel time that exists in highly congested conditions versus less congested or uncongested conditions, i.e., a minute saved in stop and go traffic is valued higher than a minute saved in free flow conditions. A second and more significant aspect is the value placed on unpredictable or variable travel time, i.e., travel time reliability. This is a component of user cost that had to this point in the mid-90s received relatively little attention.

The NCHRP evaluation was based on mail surveys and interviews of residents in the Los Angeles, California area along the priced SR 91 corridor. Standard linear utility models are estimated via a maximum likelihood logit analysis. The measure of travel time variability used in this study is the standard deviation, calculated from the mean travel time and five early/late possibilities presented under each option on the survey. The effect of travel-time uncertainty is modeled both with and without schedule costs – early or late arrivals – included, but including that along with reliability reduced significance. The conclusion was that including unreliability accounts for any effect of scheduling costs, and the respective variables should not all be in the same model.

The resulting values found for value of travel time are consistent with the large body of evidence present. Experimenting with many model specifications, the report concludes that the value placed on travel time reliability, measured as the standard deviation of travel time, is 131% of the value of time. If the value of time is taken as half the mean wage rate, a common basis of comparison is then the value of reliability as roughly 66% of the mean wage rate. Considering data limitations and practical implementation, the recommendation in the report is to apply a mark-up factor of 2.5 to the value of time in congestion where detailed breakdowns of travel conditions and traveler characteristics are generally not available. However, better access to automated detector data now may alleviate that first limitation.

Subsequent revealed preference work, most often also based on SR 91 or I-15 in southern California, have shown roughly similar values of reliability. A revealed preference analysis along the SR 91 corridor used the difference between the 90th and 50th percentile travel times as the measure of reliability, and the value of reliability found was around 111% of the value of time, depending on the model. A later measure of reliability as the difference between 80th and 50th percentiles found the value to be about 91% of the value of time. A newly revised textbook on the subject concludes that values of reliability tend to be the same as the value of time, when measured as the difference between 90th and 50th percentile travel times. Several survey based estimates using standard deviation as the measure generally show the value to be 0.8 to 1.3 the value of time.

An additional consideration is how values vary by user characteristics. For example, Lam and Small included variables for gender and transponder use, and estimated values of time and values of reliability using nested logit models. In that case, an installed transponder had little effect on values, but females were showed to value reliability about twice as highly as males. More recently, Brownstone and Small, again using data from priced expressways in southern California, evaluated heterogeneity with more sophisticated specifications and methods, e.g., various interactions, random parameters, Bayesian inference. Specifying characteristics such as income, travel flexibility, and gender, they find that value of time is clearly nonlinear, but inherent statistical complications in interpreting the results pose a barrier to practical implementation. The recommendation on that point is to pursue better measurement of individual characteristics rather than keep tinkering with the models and statistics.

3.2.3.3 Example Application to Ramp Metering

Without question, the most famous study of ramp metering occurred in late 2000 in the Twin Cities when all 430 meters in the region were shut off for several weeks by legislative mandate. The evaluation of this unique with and without metering data showed a 15 to 1 benefit cost ratio. What had not been commonly seen prior to this report, however, was the inclusion of travel time reliability benefit. The inclusion of this new benefit is remarkable in that it amounted to 64% of the total benefit. The remaining 36% included the more common travel time savings, crash reductions, and fuel consumption, as well as less common measures of emissions.

The benefit of greater reliability was estimated by applying the same value of travel time to the change in standard deviation of travel time, which is consistent with the literature findings. However, the study points out that the reliability benefit is a conservative estimate because its typically valued at three times average travel times. That statement does not appear to be consistent with other findings in the literature.

Had the larger factor been applied, reliability gains would have increased to 84% of total benefits, dwarfing travel time savings by a factor of 309 to 1. The factor of three is sometimes applied in other project appraisal work, and it does also appear as the default value in the Intelligent Transportation Systems Deployment Analysis System (IDAS) software. The cited source for this is a paper that in turn is largely based on the NCHRP Report 431 by Small et al. discussed above. However, the "three times" figure is in reference to the shift from congested travel to uncongested travel, not to the value of reliability, *per se*, as is becoming increasingly common.

If the use of standard deviation persists as a measure of reliability, an appropriate aggregate value is likely close to the value of time used, which may be half the wage rate or whatever is assumed for the project, but perhaps not three times the value of time, the issues with standard deviation notwithstanding. A recommended measure in lieu of a central dispersion is an upper tail-only dispersion such as the difference between 90th and 50th percentile.

3.2.3.4 Travel time Reliability Methodology

The prototype study corridor is US 12/18 in Madison, Wisconsin. This is referred to locally as the Beltline freeway, which runs south and west of central Madison. The specific location evaluated here is a 5.7 mile segment of westbound US 12/18 in the southeast portion of the metro area. This segment runs from the system interchange with I-39 / I-90 to the interchange with US 151 / South Park Street. Two-way daily volumes along this six-lane section of freeway currently average around 110,000 to 130,000 vehicles per day. Jobs in the area are relatively more concentrated in central Madison, at major hospitals, the University of Wisconsin, and immediately surrounding area. The westbound Beltline is a common route – with little alternative – for weekday morning inbound travelers coming from points east or southeast of Madison.

The Wisconsin Department of Transportation (WisDOT) maintains a network of in-pavement inductive loop detectors that collect volume, speed, and occupancy data at 20-second intervals. The data are processed internally and archived in 5-minute intervals for each lane at each station. The archived data management system (ADMS) for this is called the WisTransPortal, which is housed at the University of Wisconsin-Madison Traffic Operations and Safety (TOPS) Laboratory. The WisTransPortal is primarily an Oracle 10g database archiving a variety of data related to traffic operations, safety/crashes, road weather, and construction/lane closures. The information is available via web-based applications, direct SQL access, and real-time XML feeds, e.g., used for furnishing the 511 traveler information system.

On this westbound segment there are five system detection stations (SDS), each with pairs of loop detectors across the three lanes of travel. Loop pairs are sometimes referred to as speed traps, and they are installed in such a way to enable direct measurement of speeds, with calibration, as individual vehicles travel across the pair. Flow (volume) data are also collected, as are occupancy data, which can in turn be converted to density. Figure 3 schematically illustrates the study segment.

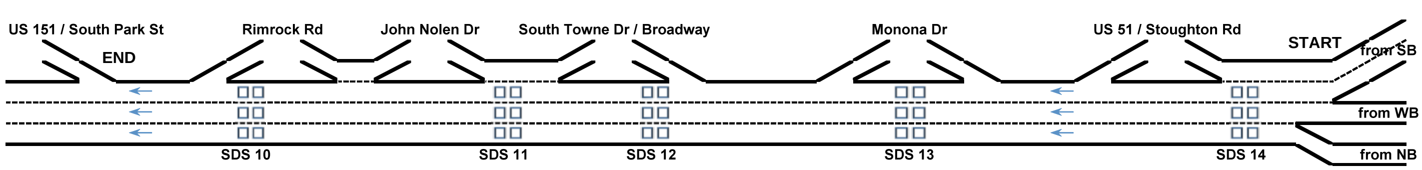

Figure 3. Schematic Segment Depiction (not to scale)

Speeds at each of the 15 detector pairs were collected for aggregated 15-minute intervals from 6:30 AM to 9:00 AM for every Tuesday, Wednesday, and Thursday from January, 2005 to April,

2008. The 15-minute interval is used due to its magnitude similar to the segment travel time as well as to follow related studies on the value of travel time reliability.

Tuesdays to Thursdays are selected to represent the typical commuting days while conveniently avoiding confounding effects from several holidays or adjoining weekend travel. This results in 77,550 potential observations. However, loop detectors are famously unreliable, and those on this section of freeway are no exception. Of those 77,550 potential observations, 41,995 (54%) contained no information. In addition, the following specific holidays were excluded:

- Thursday, Nov 24, 2005 (Thanksgiving)
- Tuesday, Jul 4, 2006 (Independence)
- Thursday, Nov 23, 2006 (Thanksgiving)
- Wednesday, Jul 4, 2007 (Independence)
- Thursday, Nov 22, 2007 (Thanksgiving)
- Tuesday, Dec 25, 2007 (Christmas)
- Tuesday, Jan 1, 2008 (New Years)

The remaining observations were compiled by date and time. For simplicity and to avoid some possible bias from non-random missing data, the observations were further filtered by those 15-minute intervals wherein all 15 detectors reported data. Of the 2,721 remaining 15-minute intervals with data, around 39% had one or more detectors not reporting. Nonetheless, there remain over 1,600 time intervals with complete speed data from which to estimate travel time, as follows:

$$T = \sum_{l=1}^{5} \left[\frac{d_l}{\sum_{i=1}^{3} \frac{v_i}{60} \Big/ 3} \right]$$

where T = segment travel time
l = location of system detector station (1 to 5)
d = distance or length of travel segment represented by l
i = loop pair, for lanes 1, 2, and 3
v = 15-minute aggregate speed measured at loop pair i, in miles per hour

The sum of the five travel segment lengths ($d_1 + d_2 + \ldots + d_5$) constitute the total 5.7 mile section of the Beltline under evaluation. The method here of summing travel times rather than averaging speeds and converting to time is intentional and is akin to the important difference in calculating space mean speed instead of time mean speed.

Incorporating or controlling for other factors that contribute to unreliable travel times is beyond the scope of this analysis. Examples of these include road weather conditions, e.g., wet, snow, ice; construction activity that impacts capacity on the study segment or on a nearby segment, adjoining or parallel; and traffic incidents such as crashes, special events, and so forth. This analysis is evaluating travel time variations, and many factors contribute to unreliability, but there is no cause to exclude any data from incidents or other impacts. Adverse weather, incidents, and construction are normal and frequent occurrences, and travelers grow accustomed to their effects and plan accordingly.

That said, travelers could be expected to make particular adjustments for longer term, predictable effects, so one exception that is checked for is a construction project during the period of collected data. A two and a half mile segment of the Beltline west (downstream) of this location – Fish Hatchery Road to Verona Road – underwent major reconstruction from April 30 to July 20, 2007. Depending on the 15-minute interval, the number of valid samples – i.e., non-holiday, Tuesday-Thursday, all 15 detectors reporting – was between nine and twelve during that timeframe. The significance of the observed difference in mean travel time was evaluated via a t test. This result is shown in the following section.

For the comparisons in the analysis that follows, five values of travel time reliability are applied, taken from the references discussed in the previous section. They are listed here as a percent of mean wage rate:

- 66%, based on standard deviation (NCHRP Report 431)
- 46%, based 80th – 50th percentile
- 55%, based on 90th – 50th percentile
- 50%, based on 90th – 50th percentile
- 50%, based on 90^{th} percentile – mean travel time

3.2.3.5 Reliability Findings

Table 7 summarizes the travel time measurements. N is the number of observations (with complete detector data); the mean and median values are shown; and then measures of unreliability – the standard deviation and the difference between the 90^{th} and 50^{th} percentile travel times. This last value would be the additional travel time to allow, on average, in order to arrive on time 18 out of 20 work days in a month.

Table 7. Travel Times (minutes)

Start Time	6:30	6:45	7:00	7:15	7:30	7:45	8:00	8:15	8:30	8:45
N	162	161	159	164	164	166	167	171	176	172
Mean	5.98	6.19	6.23	7.30	8.98	9.86	8.62	7.59	6.98	6.51
Median (50th %ile)	5.86	5.95	5.95	6.65	7.94	9.01	6.90	6.38	6.11	5.89
Standard Deviation	0.92	1.38	1.65	2.54	3.33	4.10	4.39	3.83	3.07	2.48
90th %ile	6.64	6.92	6.87	9.52	12.55	14.73	12.60	10.96	10.22	8.29
90th - 50th %ile	0.78	0.98	0.92	2.87	4.60	5.72	5.69	4.58	4.11	2.41

The series of plots in Figure 4 on the following page illustrate the travel times for three 15-minute intervals each day. The horizontal axes do not show specific dates, but the scale covers the 2005-2008 collection timeframe. The horizontal dashed line shows the mean travel time, and the dashed box calls out the Beltline reconstruction period in early 2007.

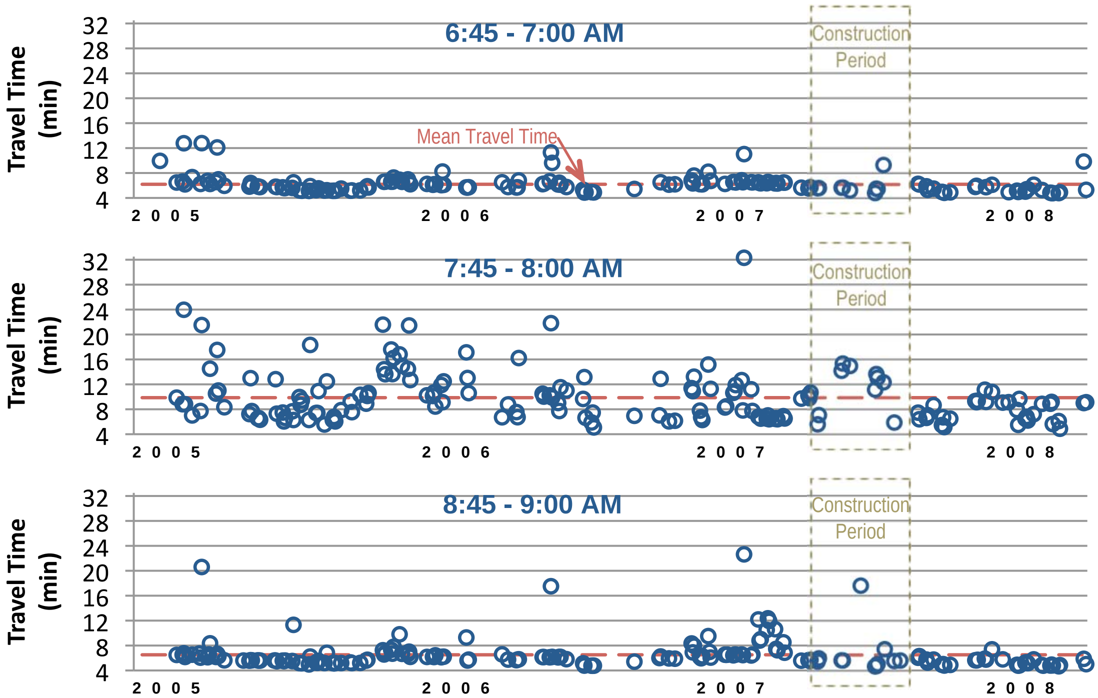

Figure 4. 15-Minute Travel Times

The effects of the Beltline reconstruction are summarized in Table 8. The table shows the 15-minute time period, the mean travel time during construction and otherwise, the difference, and the significance of the t statistic.

Table 8. Construction Effect on Travel Time

	Average Travel Time			
Time	Normal	Construction	Difference	Signif.
6:30	5.15	4.92	-0.22	0.256
6:45	5.28	8.52	3.24	**0.000**
7:00	5.30	5.25	-0.05	0.891
7:15	6.06	7.15	1.09	**0.035**
7:30	7.46	9.77	2.30	**0.004**
7:45	8.37	9.94	1.57	0.142
8:00	7.33	7.99	0.66	0.540
8:15	6.45	6.55	0.10	0.909
8:30	5.94	5.20	-0.74	0.256
8:45	5.62	5.86	0.24	0.694

Travel times in three of the 15-minute periods (6:45, 7:15, and 7:30) were greater during construction, with significance at the 0.05 level. However, as stated earlier, construction impediment is a normal part of travel, and travel times during this period are included in the remainder of this analysis.

A visually illustrative method of depicting the travel time data is shown below in Figure 5. By sorting all travel times from least to most, for each time period, they may be presented as a cumulative probability distribution. The arrows were added to call out the transition from a less congested period, i.e., more reliable travel times, through the morning peak congested period, and back again to a less congested time.

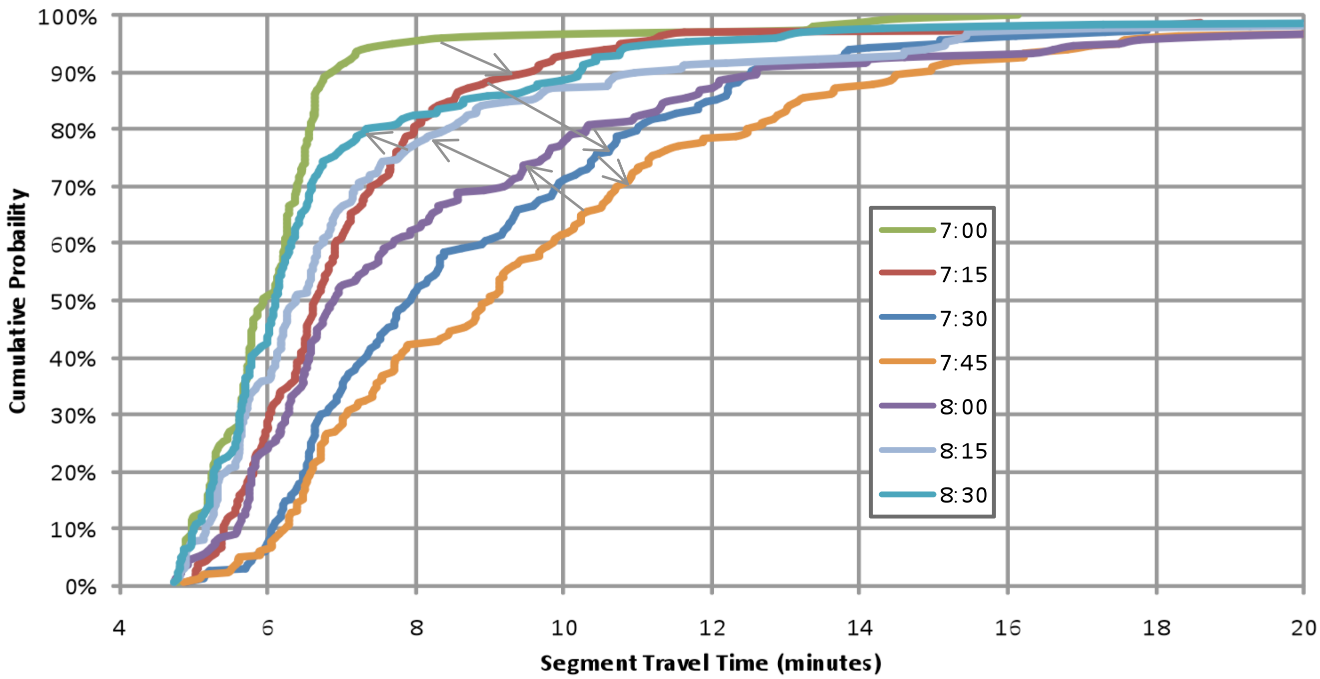

Figure 5. Travel Time Probability Distribution

The graph shows the median travel time varying by about three minutes, between six minutes and nine minutes. The 90th percentile travel time varies by about eight minutes, between about seven minutes and 15 minutes.

Four different measures of reliability discussed are depicted in Figure 6. They are all evidently closely related, but statistically the strongest correlation is between the measures for 90th percentile – median and 90th percentile – mean (0.988 Spearman rank correlation). The weakest correlation is between the measures for standard deviation and 80th percentile – median (0.927 Spearman rank correlation).

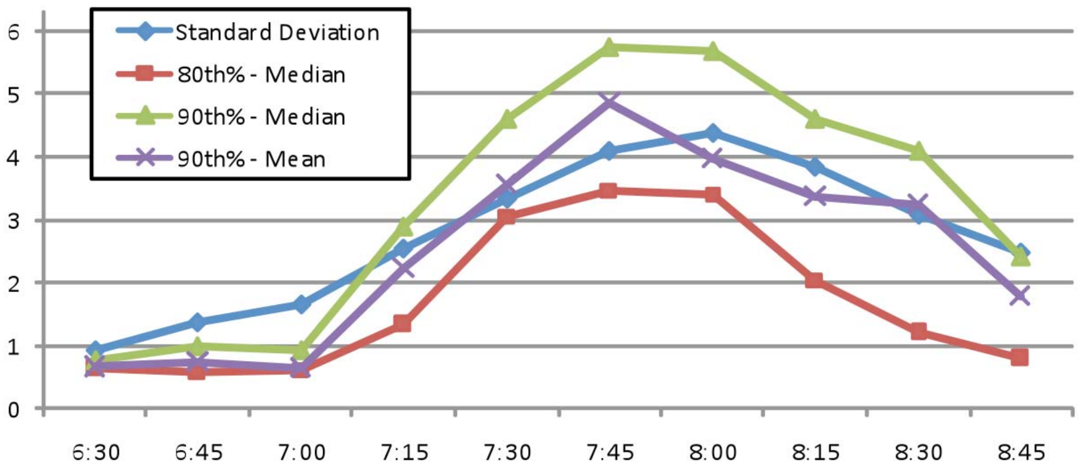

Figure 6. Measures of Travel Time Reliability

It is apparent from the plot that the order from smallest to greatest is consistent across time periods for three of the four measures – the exception is standard deviation. During times of low congestion, the standard deviation is greater than the other variability measures. This may reinforce the shift to using the upper tail only measures as a better gauge of travel time reliability.

The mean wage rate in Madison was $19.25 per hour in 2006. If the value of time is equivalent to 50% of the mean wage rate, the value of time is 16 cents/minute. In turn, the results for the various measures of reliability are summarized in Table 9 and Figure 7.

Table 9. Value of Reliability

Measure of Reliablity	% of Mean Wage	Value of Reliability ($/min)	6:30	6:45	7:00	7:15	7:30	7:45	8:00	8:15	8:30	8:45
Std. Deviation	66%	$0.21	$0.19	$0.29	$0.35	$0.53	$0.70	$0.86	$0.92	$0.81	$0.64	$0.52
80^{th}% - Med.	46%	$0.15	$0.09	$0.09	$0.09	$0.19	$0.45	$0.51	$0.50	$0.30	$0.18	$0.12
90^{th}% - Med.	55%	$0.18	$0.14	$0.17	$0.16	$0.51	$0.82	$1.02	$1.01	$0.81	$0.73	$0.43
90^{th}% - Med.	50%	$0.16	$0.12	$0.16	$0.15	$0.46	$0.74	$0.92	$0.91	$0.74	$0.66	$0.39
90^{th}% - Mean	50%	$0.16	$0.11	$0.12	$0.10	$0.36	$0.57	$0.78	$0.64	$0.54	$0.52	$0.29

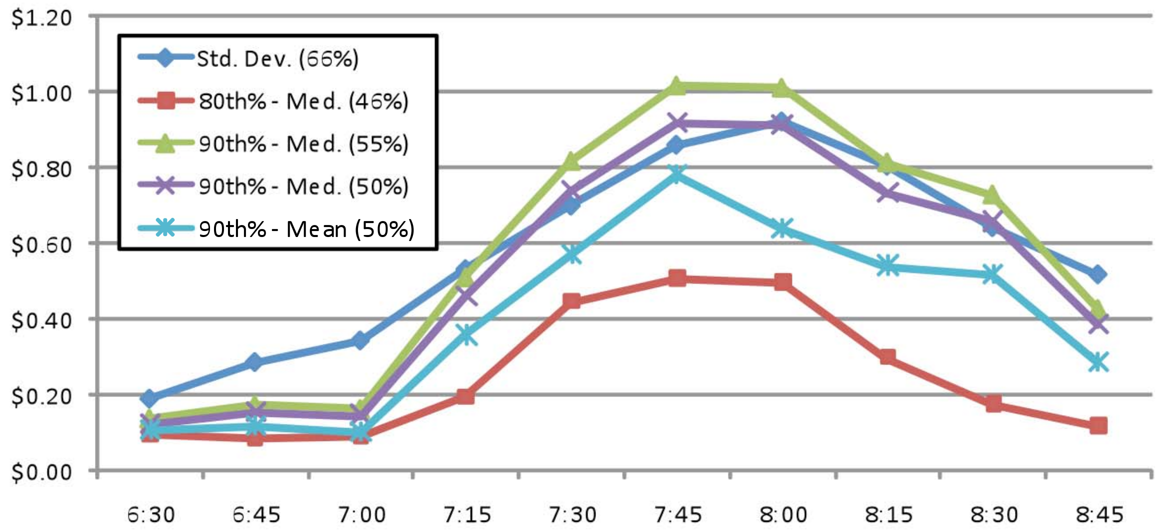

Figure 7. Value of Reliability

The values based on standard deviation carry over the same issue with the values being relatively higher in less congested times. The unit value of reliability placed on the 80th percentile measure appears relatively low compared to other estimates.

3.2.3.6 Reliability Discussion

These travel times are only for a segment of a freeway, thus they are a portion of a longer trip which invariably will have other factors contributing to unreliable trip times. Nonetheless, measures of travel time and reliability are more difficult to acquire on arterial or local streets, while evaluating what data are available on freeways yields valuable insight into the nature of travel decisions. In this case travelers between 7:45-8:00 AM on the westbound Beltline are choosing to travel then and are not changing departure times, in the aggregate, pointing to the offsetting values of schedule delay, i.e., early versus late arrivals. An individual altering departure time in either direction from this peak would reduce both their average travel time and reliability costs, not to mention the congestion externality from the unpriced facility.

One goal for this research from the outset was to develop a better, i.e., more practical and transparent, methodology to evaluate the effect on travel time and reliability from an ITS strategy like metering. A particular challenge has turned out to be finding usable before data, for several reasons other than just the intermittent missing data. First, automated detector data here were not archived prior to 1996 while many of the ramp meters in Wisconsin were installed before this, as early as around 1970. Second, where ramp meters are installed, there are a set of loop detectors on the freeway mainline that provide the best insight into operations. However, those detectors are often installed at the same time as the ramp meter installation, so there are no before data. Third, ramp meter installations often occur as part of a larger freeway reconstruction project. If detectors were present prior to construction, there are before data. However, the reconstruction project itself often increases capacity of the roadway. Examples of this readily appear in the data, as ramp meters alone cannot possibly cause the dramatic improvements sometimes evident in the after data. But maybe they are. A chief difficulty is collecting data related to construction activity and physical modifications affecting capacity.

Recommendations for further inquiry into the general aspects of using detector data include two things already mentioned above. First is to explore what is the most meaningful measure of travel time reliability, and can a recommendation be made on which to standardize for common practice. The effort behind this paper points to the use of an upper tail distribution (not standard deviation) based on comparison to the median (not the mean) such as the 90^{th} – 50^{th} percentile measure, which is consistent with the somewhat established notion of being on time 18 days out of 20. To reiterate, while reporting different percentiles is easy, the need for a consistent measure hinges largely on the economic value applied to that measure. As suggested earlier, research into what is an efficient – or more efficient – estimator of reliability should be pursued by assessing various measures under the same model specification.

The second recommendation is to explore how detector data can be used estimate relative relationships between values of travel time, reliability, and departure time, especially in the absence of priced roads. An important question may lie in whether there is anything to be learned from aggregate departure time choice when no tolls and no disaggregate income information is available, i.e., the common situation with freeway detection data.

3.3 DISCUSSION AND CONCLUDING REMARKS

For this immediate purpose of evaluating ramp metering effects, the 95^{th} – 50^{th} percentile measure will be used. The data are obtained from the WisTransPortal for each metering location in the timeframe prior to and following a change in operation (e.g., turn on, turn off, or change in timing window). The table of data is then brought into the Stata software for intermediate processing. Data outside the time frame are filtered out, as are observations from loops not representing the mainline travel adjacent to the meter location. Observations with the meter active are flagged separately from observations without the meter active. At this point the observations are automatically inspected for zero or null entries and this percent is included in the output for consideration of a data confidence level, described elsewhere. Travel speeds are converted to travel time on a per mile basis. The subsequent results presented in later memos treat the distance of a meter's influence as one mile along the freeway, consider 2,640 feet upstream and downstream.

The mean speed change without and with the meter operational is checked for significance with the t statistic. Finally, the descriptive statistics mean, 50^{th} percentile (median), and 95^{th} percentile travel time are calculated, weighted by volume, for the mainline adjacent to the meter without and with the meter operational. The following measures are returned from Stata:

- Mean flow with meter inactive
- Mean flow with meter active
- Mean speed with meter inactive
- Mean speed with meter active
- t test of significance between mean speed without and with meter

- Mean travel time with meter inactive
- 50th percentile (median) travel time with meter inactive
- 95th percentile travel time with meter inactive
- Mean travel time with meter active
- 50th percentile (median) travel time with meter active
- 95th percentile travel time with meter active

These measures are brought into Excel for summary analysis and graph creation, as displayed in the next memo. Crash experience is treated differently due to limited data availability, and that is discussed in tech memo #6.

There remain some methodological issues that may be improved upon for a more robust and reliable analysis. To begin, a 20-year analysis period could be fleshed out that includes any additional operating costs and salvage values, or estimated life cycle costs. Cost may need to be estimated more accurately, carefully considering joint costs with operations facilities, staff, or other ITS elements.

Broad implications arise from the distinction between a “before and after” analysis versus a “with and without” analysis. While the before and after analysis is easier and can rely on readily available field data, it may not necessarily provide meaningful results as conditions change or if other factors play a significant role. In the case of older analyses and this follow-on evaluation, the analysis timeframe is relatively short and travel conditions are assumed to be relatively stable. However, to evaluate the effects of ramp metering over a typical 20- or 30-year timeframe, the operating conditions both with and without meters in operation is required. Traffic forecasts would need to be relied upon for future analysis years.

Assuming periodic shutdowns of the meters are not practical, at least two approaches still present themselves. One is to develop a microsimulation model for the corridor and evaluate the with and without conditions. The second is to identify control conditions elsewhere on the freeway network. Both are relatively high resource-consuming processes with their own inherent complications.

Crash risk is known to vary with operating conditions. As conditions change, even applying a simple model of crash risk would more accurately estimate benefits from crash reductions. Furthermore, although there were no fatal crashes on this corridor in the analysis period, there still exists a fatal crash hazard. What that hazard is may be evident by looking at additional years of crash data, and further research is then necessary to estimate what the effect of ramp metering might have on very low fatal crash risk, given the operating conditions with and without ramp metering.

Other factors that may be included in an economic evaluation of this type include travel time reliability and environmental benefits related to reduced emissions, for example. Travel time reliability especially is an increasingly important consideration with significant economic repercussions, especially for commercial vehicle traffic.

A key outcome of this assessment was the research and consideration of economic values for travel time, crashes of different severities, and operating costs, and their application to an ITS operational strategy.

The results of this assessment should be reassuring to managers and policy makers who are faced with scrutiny of the ramp metering program. Despite far from optimal conditions for the ramp meters along the US 45 corridor, the net present value is at least $600 thousand, with a benefit cost ratio of two or better. These are a result of relatively low values of travel time savings, exclusion of some potentially large benefits measures, and a grossly over simplified ten-year analysis period. With this basic methodology expanded into evaluating ramp meters and other ITS projects, results will continue to demonstrate that ITS operational strategies are very competitive from an economic benefit framework.