Empirical Evaluation of Drivers' Behavior Response to Accident Information on Freeway Changeable Message Signs

Yiguang (Ethan) Xuan (corresponding author)
Partners for Advanced Transportation Technology (PATH)
University of California, Berkeley
2105 Bancroft Way Suite 300, Berkeley, CA, 94720-3830, USA
xuan.yiguang@path.berkeley.edu

Adib Kanafani
Graduate School
University of California, Berkeley
112 McLaughlin Hall #1712, Berkeley, CA 94720-1712, USA
phone: +1 (510) 642 0367, fax: +1 (510) 643 8919
kanafani@berkeley.edu

Paper submitted to TRB Annual Meeting 2014
July 24, 2013

4980 words + 6 figures + 4 tables ⇒ 7480 'words'

ABSTRACT

In this paper, we focus on the accident messages displayed on freeway changeable message signs (CMS), and study their effect on drivers' route choice behavior. Previous studies on the effect of CMS messages have shown mixed results, due to differences in event types and modeling choices. Therefore, the purpose of this paper is to ascertain the real effect of CMS accident messages, and also to compare two commonly used statistical models. We calculate the proportion of total flow heading to off-ramps or freeway interchanges (called "diversion rate" hereafter) at diverging locations, and use the change in diversion rate as the indicator for behavior change. We first draw insights from two case studies, and find that the effect of CMS accident messages on the diversion rate is minor and the effect of visible congestion is dominant. We then compare two commonly used statistical analyses, accounting for the effect of visible congestion. The correlation analysis compares the diversion rate with and without CMS accident messages, while the causality analysis compares the diversion rate right before and after CMS accident messages are turned on or off. With empirical data from three study sites, we use the causality analysis to show that the real effect of CMS accident messages on diversion rate is insignificant. However, the correlation analysis shows positive correlation between CMS accident messages and diversion rate, indicating that this analysis cannot be used to draw causal inferences and that other factors have played a role in changing the diversion rate.

INTRODUCTION

In this paper, we evaluate the effect of accident messages displayed on freeway changeable message signs (CMS) on drivers' route choice behavior. (Changeable message signs are also known as variable message signs or dynamic message signs.)

CMS is an important component of the Advanced Traveler Information System (ATIS), whose purpose is to generate and disseminate traveler information to enable informed travel decisions. Compared with other traveler information sources like TV, radio, on-board navigation devices, and smartphone apps, CMS has two distinct features. First, CMS only targets en route drivers traveling through certain locations, mainly for them to change their routes (and sometimes travel modes, like park-and-ride) but not their departure time. Second, the CMS message is visible to all drivers within its range. The rest of the information sources either have a limited penetration rate, or may not be active at all times, or both.

CMS messages are mostly about travel time and accident information. Sometimes, CMS will broadcast roadwork information, weather-related cautions, etc. In this paper, we will focus on accident messages instead of travel time or roadwork, for two reasons. First, accidents are not foreseeable, and therefore accident messages are more likely to be valuable to travelers and motivate their behavior change. Second, travel time does not show where the problem is even when its value is higher than normal. Comparatively, accident messages are more accurate, with information on the location and description of the accidents, and therefore are more likely to induce behavior change.

The CMS system has been in use for a long time, dating back to at least the 1960s. The system is expensive: Typical installation cost for each freeway CMS is around $200,000, excluding the cost for operation and maintenance. In California alone, there are about 771 such signs on the freeway, which cost at least $150 million for installation.

However, we have very limited understanding of how effective the CMS system is. This statement is true for ATIS generally. The effectiveness of the system is largely anecdotal. According to a recently published NCHRP report (27), only 30 percent of the agencies reported having evaluation data that demonstrate the benefits of providing information to the traveling public, and only 40 percent have an ongoing program for evaluating the provision of traveler information.

The goal of our research is to evaluate the effectiveness of CMS messages and to understand how travelers react to this information. The main challenge in answering this question is to distinguish the effect of CMS from the effect of other sources of information such as visible congestion and radio.

This paper is organized as follows. In the literature review section, we describe the previous work and the gap in modeling the effect of CMS on travelers. After describing our experimental design, data, and study sites, we start by drawing insights from two case studies. We then perform two types of statistical analyses: The first analysis is simple and focuses on correlation only; the second analysis borrows ideas from regression discontinuity and determines the causal effect of the CMS message. At the end, we conclude the paper and provide directions for future research.

LITERATURE REVIEW

We will start with a brief description of the literature on the network effect of traveler information. The majority of the literature on the effect of CMS uses either the survey method or the empirical method, which are described below in detail.

Network Effect

Earlier efforts to quantify the effect of traveler information (*17*, *2*) tend to make idealized assumptions about driver response, for example that drivers will fully comply with route diversion suggestions, either for system optimal assignment or to avoid incidents. The purpose of these analyses is not to model driver response accurately, but rather to estimate the upper bound of the benefit of traveler information.

Survey Method

The majority of the literature on the effectiveness of CMS is based on surveys (*18*, *24*, *23*, *3*, *12*, *26*, *31*, *25*, *20*, *1*, *7*, *8*). Survey is an effective method of obtaining information on individuals and their thought process. These and many other studies have offered insights into factors that affect drivers' decision for route diversion, including purpose of travel, schedule flexibility, travel distance, cause of congestion on current route, familiarity with alternative routes, information availability on alternative routes, and previous experiences with traveler information. (*20*, *8*) provide good literature reviews on the findings so far.

The main problem with the survey method is that people's stated intention and their actual behavior may not be consistent. Very often, survey overestimates the rate for route diversion (*33*), and therefore may not be appropriate for operational applications. For example, if CMS is used to divert traffic to arterial streets, the amount of traffic diverted needs to be estimated so that traffic signals on local streets can account for it. Survey generally does not provide the level of accuracy needed for such operational purposes.

Empirical Method

The alternative to study the effectiveness of CMS is to rely on field data to see what travelers actually did. This is the approach taken in this paper. The number of studies we are aware of in this category is much smaller compared to those using surveys, and these studies are summarized in TABLE 1. Overall, the effect of CMS reported by these studies varies over a very wide range. While almost all the studies report the effect of CMS to be statistically significant, the magnitude can be insignificant for operational purposes, as in (*10*), for example.

As mentioned above, the main challenge in studying the effect of CMS is to distinguish its effect from the effect of other information sources. This challenge is illustrated in the following three aspects: the type of event studied, the factors considered to have an impact on diversion behavior, and the method to determine statistical significance of the factors, as categorized in TABLE 1.

For the type of event, out of the 14 studies listed in TABLE 1, one is for special events, six are for work zones. It is possible that drivers obtain some information about these pre-planned events before they begin their trips. Therefore, the diversion observed in the field is a combination of the effects of CMS and other information sources.

For the factors that affect diversion, most of the studies account for CMS, but only 4 out of 14 studies listed in TABLE 1 account for the effect of visible congestion. However, the effect of visible congestion on diversion is well-documented and is named the "natural diversion" phenomenon (*10*, *30*, *29*, *4*, *22*, *32*, *33*). Natural diversion refers to the empirical observation that many drivers change their routes when serious congestion is visible, and is potentially the explanation for the observation in (*13*) that sometimes the diversion rate changes before the message changes.

For the method to determine statistical significance, we can see mainly two types of meth-

TABLE 1 Summary of empirical studies on the effect of CMS on driver behavior.

Reference & year	Site location	Event type	Factors for diversion	Method for comparing diversion rates and determining statistical significance
(*11*) 1978	TX, USA	Special events	CMS *	Average with and without CMS message (alternating message and blank screen)
(*28*) 1978	TX, USA	Work zone	CMS *	5-min average before and after CMS message
(*10*) 1982	TX, USA	Accident	CMS * Congestion *	Average with and without CMS message
(*19*) 1996	Munich, Germany	Normal	CMS *	No explanation
(*34*) 1996	Paris, France	Normal	CMS *	5-min average before and after CMS message
(*14*) 2003	WI, USA	Work zone	CMS *	Average with and without CMS message
(*4*) 2004	NC, USA	Work zone	CMS ** Congestion *	Average with and without CMS message, controlling for congestion
(*9*) 2005	CA, USA	Work zone	CMS *	Average with and without CMS message
(*21*) 2006	CA, USA	Work zone	CMS *	Average with and without CMS message
(*15*) 2006	MN, USA	Accident	CMS *	10-min average before and after CMS message
(*13*) 2008	ON, Canada	Normal	CMS *	30-min average before and after CMS message
(*22*) 2011	WI, USA	Work zone	CMS *	Average with and without CMS message
(*32*) 2011	WA, USA	Normal	Congestion *	Threshold based method
(*33*) 2011	Shanghai, China	Accident	CMS * Congestion *	5-min average before and after CMS message

* The effect of this factor (either CMS or visible congestion) on the driver diversion behavior is statistically significant.
** The effect of CMS on the driver diversion behavior is statistically significant only with both delay and alternative route advisory.

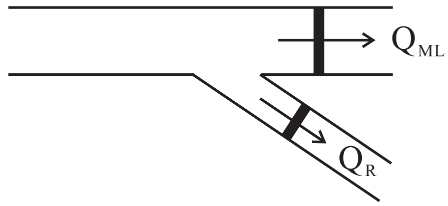


FIGURE 1 Definition of diversion rate

ods: comparing the diversion rate with and without CMS messages, and comparing the diversion rate some time period before and after certain CMS messages. The second method is more likely to reveal the real effect of CMS, while the first method is more likely to capture a mixed effect from various sources. We will describe the two methods in detail and compare their results later in this paper.

As a summary, there are a very limited number of empirical studies on the effect of CMS on drivers' diversion behavior. The effect of the CMS report by these studies varies greatly. Besides difference in site location, we think the mixed results can also be attributed to the targeted type of event, the factors considered that affect diversion, and the statistical method used. In this paper, to ascertain the real effect of CMS on driver diversion behavior, we will focus on accidents, account for both the effect of CMS and visible congestion, and compare result from the two types of statistical methods mentioned above.

EXPERIMENTAL DESIGN

First, we need a metric that can represent the drivers' choices of route. We do this by defining diversion rate $R(t)$ to be the proportion of total flow that heads to the off-ramp:

$$R(t) = \frac{Q_R(t)}{Q_{ML}(t) + Q_R(t)}, \tag{1}$$

where $Q_{ML}(t)$ and $Q_R(t)$ are mainline flow and off-ramp flow as shown in FIGURE 1. This diversion rate is generally determined by the destination of drivers and varies slowly with the time of day. Therefore, rapid change in the diversion rate is an indicator for behavior change in route choice. We will see whether drivers' change of route coincides in time with the provision of information, either from CMS or from visible congestion, and quantify the impact of these factors.

Note that flow data are aggregate, so they can only be used to study aggregate behavior, e.g., how many people take a certain exit. To study individual behavior, we need data that can trace individual travelers, like GPS data, which is not covered in this paper.

Data

The California Department of Transportation (Caltrans) provides a live feed of the CMS messages posted on all 771 freeway CMSs in its 12 districts (*5*). The messages are mostly updated once every minute. (No study site is from Caltrans District 2, where CMS messages are updated once every five minutes.) We archive all the messages into our database starting from November 2012. For this paper, we use data from November 2012 to April 2013.

Besides CMS messages, we use flow and occupancy data from both mainline and off-ramp loop detectors from Caltran's Performance Measurement System (PeMS) (*6*). The raw PeMS data are aggregated in 30-second intervals. We also use data from November 2012 to April 2013, consistent with the CMS data.

Sites

Because of the data-driven nature of this study, the study sites are restricted to those with suitable data. The availability of good data on off-ramp detectors is a significant constraint. It is extremely difficult to identify sites where all the off-ramps and the corresponding freeway mainline are equipped with functioning loop detectors. The study sites are selected where most of the detectors are functioning.

The site also needs to have a CMS nearby, and the CMS needs to display accident messages during the dates of interest. In this paper, we exclude CMS messages that are broadcast on multiple CMS signs and that involve directions on alternative routes. These messages have more potential to be effective but are left for future research. The difficulty with CMS messages on multiple CMS signs is that the timing when drivers receive the information is much more complicated. The difficulty with CMS messages with directions on alternative routes is that these messages are much less frequent, limiting the number of data samples.

So far, we have identified three study sites that we think are appropriate for our exploration. The first site is along I-210E in Caltrans District 8 (San Bernardino / Riverside). A sketch of the site with mainline and off-ramp detectors is shown in FIGURE 2(a). The total distance from the CMS (indexed 808866) to Milliken is about 11.1 km (6.9 miles). The second site is along I-15N in Caltrans District 11 (San Diego / Imperial), and the third site is along I-15S, also in Caltrans District 11. Similar sketches are shown in FIGURE 2(b) and FIGURE 2(c). At site 2, the total distance from CMS 1106507 to Clairemont Mesa is about 12.6 km (7.9 miles). At site 3, the total distance from CMS 1106508 to Adams is about 13.8 km (8.6 miles).

There are a total of 18 accidents studied in this paper: seven from Site 1, seven from Site 2, and four from Site 3. Note that Euclid Avenue does not connect directly to the freeway, but is shown here because it is involved in one of the case studies. At site 2, there is another CMS indexed 1117651 at Aero and I-15N. This CMS displayed no accident message (only travel time) in our study and is thus ignored.

CASE STUDIES

We start our analysis with two case studies to draw insights into the effect of CMS on drivers' diversion behavior. The two case studies are selected such that in the first case drivers are already in congestion when they see the CMS accident message, and in the second case congestion has not yet reached the CMS location.

Setup

The first case occurs on the afternoon of Tuesday, November 27, 2012, when CMS 808866 showed information on an accident on I-210E around Euclid Avenue at Site 1. The only path to circumvent this accident is to take the Mountain Ave exit, so we focus on this off-ramp for the first case study. The second case occurs on the morning of Monday, December 3, 2012, when CMS 1106507 showed information on an accident on I-15N to the south of CA-52 at Site 2. Data are not available on the interchange from I-15N to I-805N, so we use the first downstream off-ramp with data available, University Avenue, for the second case study. The content and duration of the messages during the two case studies are shown in FIGURE 3.

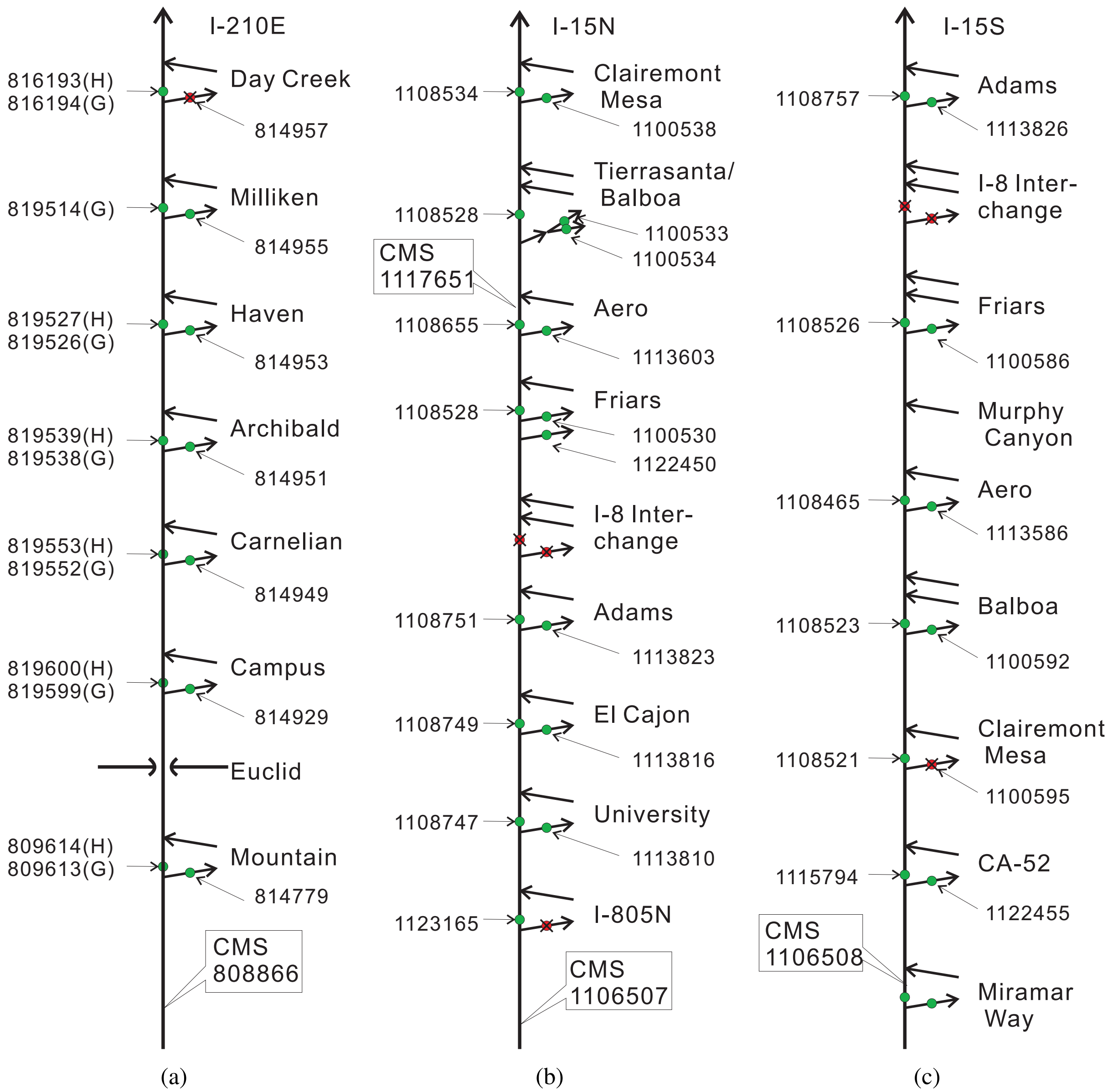
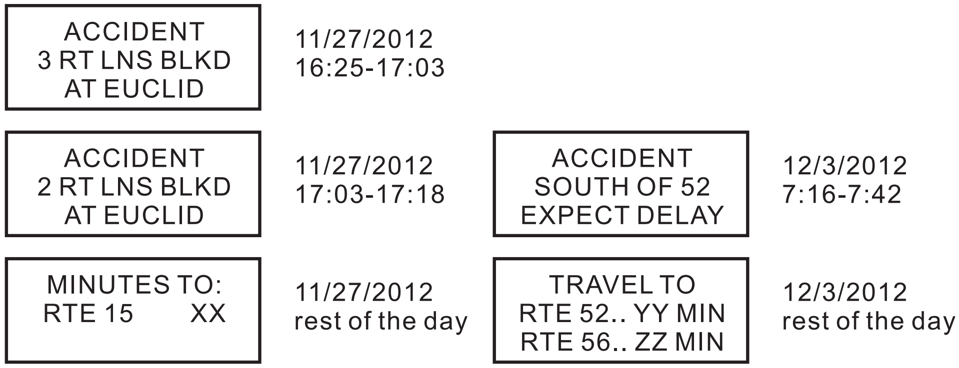


FIGURE 2 Study sites. (a) Site 1: I-210E in Caltrans District 8 (San Bernardino / Riverside). (b) Site 2: I-15N in Caltrans District 11 (San Diego / Imperial). (c) Site 3: I-15S in Caltrans District 11 (San Diego / Imperial). The numbers are the detector ID, with H for high-occupancy-vehicle lanes and G for general purpose lanes. The circles and those with a cross indicate functioning and non-functioning mainline and off-ramps detectors. Only relevant detectors are shown, ignoring on-ramp and other mainline detectors.



(a) (b)

FIGURE 3 CMS messages for the case studies. (a) Message on CMS 808866 on Tuesday, November 27, 2012. (b) Message on CMS 1106507 on Monday, December 3, 2012. The "XX", "YY" and "ZZ" in the messages are numbers for real-time travel time.

Results

FIGURE 4 shows the diversion rate as a function of the time of day, for the two case studies. Also shown is the travel time on the CMS. The gap in CMS travel time is the period of time when accident information was displayed. First, we see that the diversion rate is fairly constant, except for some high values around the time of the accidents. However, closer inspection reveals that the timing of the high diversion rate does not overlap with the timing of the CMS message.

In the first case, the diversion rate starts to increase around 16:14, and the CMS accident message is not displayed until 16:25. Actually, at 16:14 the travel time on CMS is still the free flow travel time. Therefore, other factors are affecting drivers' route choice besides the CMS messages. Also, no obvious pattern can be observed between the diversion rate and the CMS travel time.

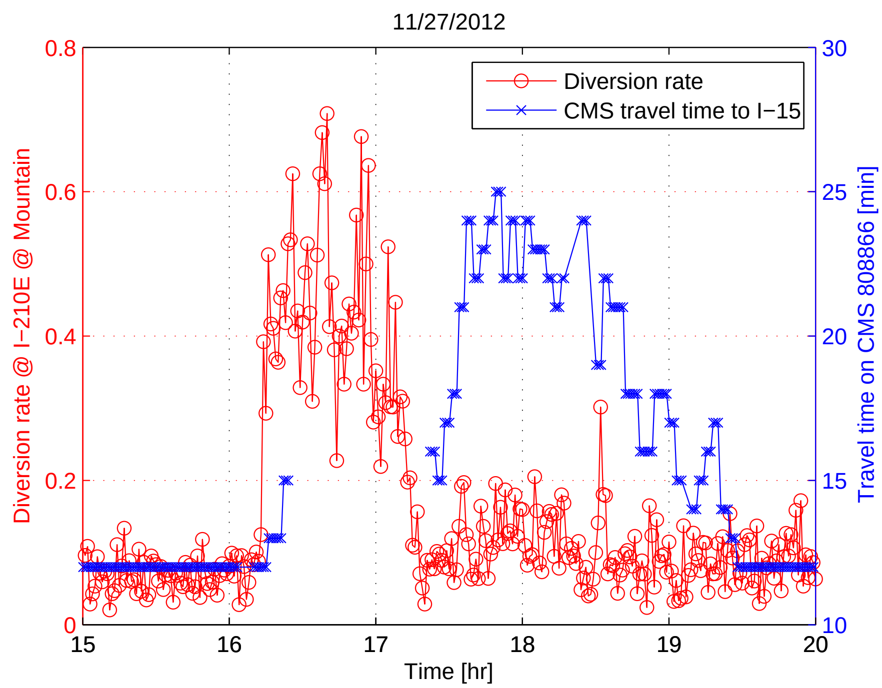
In the second case, the accident message appears at 7:16. Before that time, drivers may also get some clue from the increased CMS travel time to both CA-52 and CA-56. However, the diversion rate does not seem to increase until 7:30. Other off-ramps downstream of University Avenue exhibit similar patterns.

The case studies thus seem to suggest that the effect of the CMS messages (either the accident information or the travel time) on drivers' diversion behavior is minor, and the drivers seem to react to some other factors.

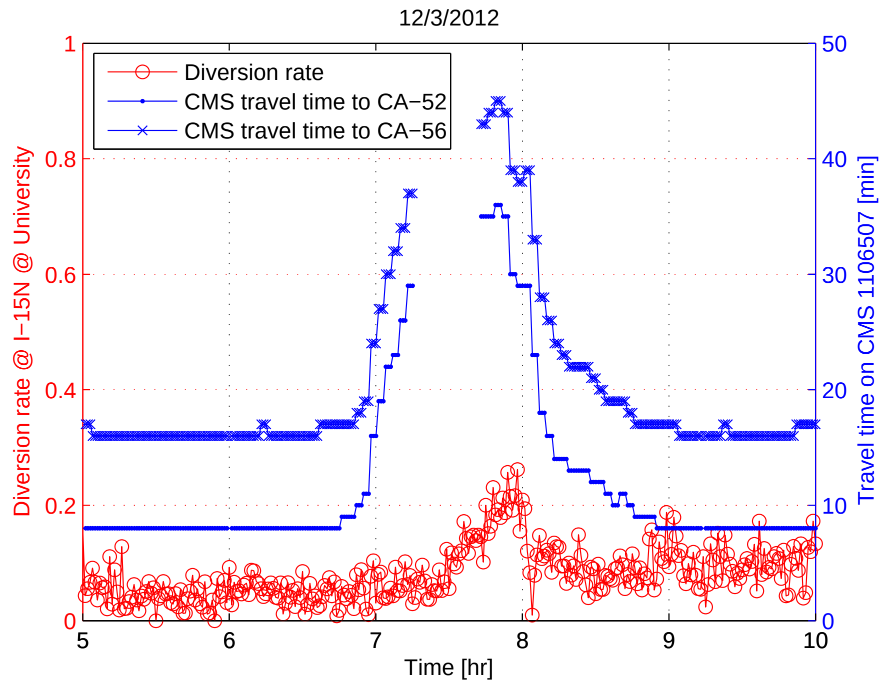
Based on the literature, we suspect that visible congestion is a possible cause for the increased diversion rate. This suspicion is supported by FIGURE 5, which shows both the diversion rate and mainline occupancy versus the time of day. The timing of visible congestion (as indicated by an increase in mainline occupancy) seems to be a better fit for the timing of the increase in diversion rate.

STATISTICAL ANALYSIS

The case studies provide us with good insights, notably the seemingly minor effect of CMS messages on drivers' diversion behavior and the comparatively more dominant effect of visible congestion. But these observations have limited generality, due to the small sample size. The purpose of the statistical analysis is to generalize the observation across various sites and dates with more samples. Next, we will describe in detail and compare the two types of statistical analyses mentioned

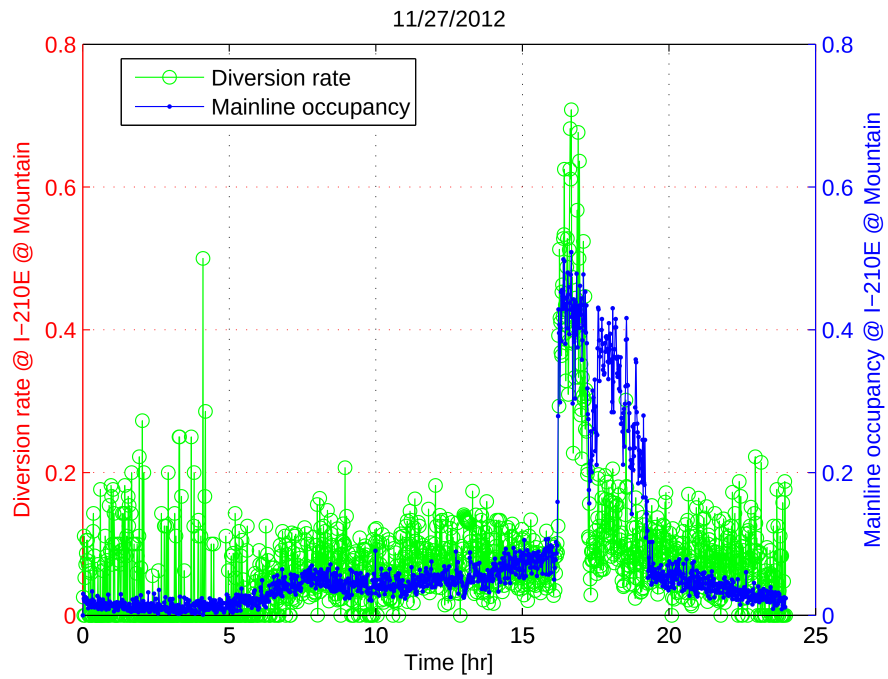


(a)

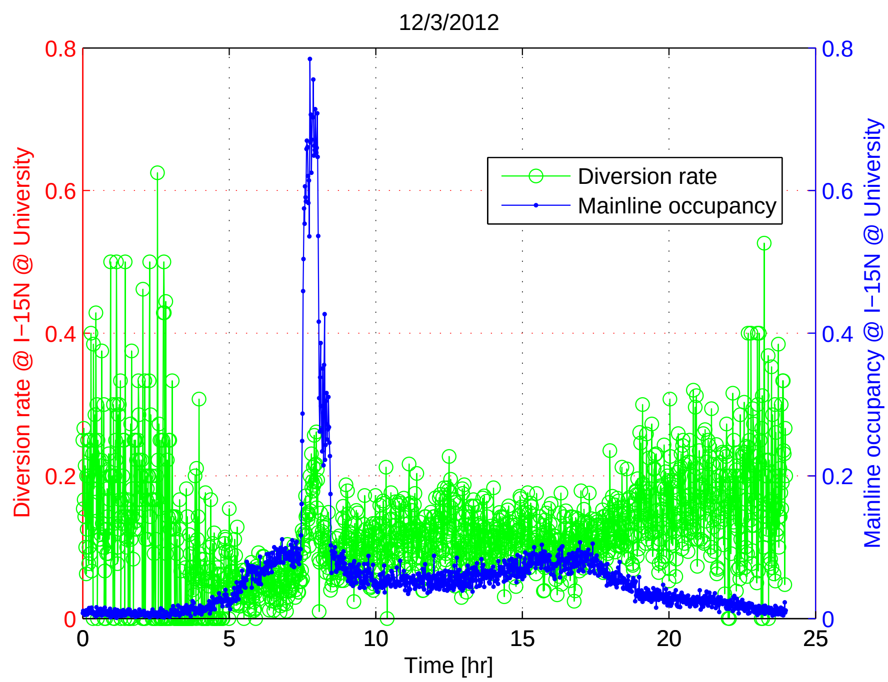


(b)

FIGURE 4 Diversion rate and CMS travel time versus the time of day, (a) on I-210E at Mountain Avenue, on November 27, 2012, (b) on I-15N at University Avenue, on December 3, 2012. The gap in CMS travel time is the period of time when accident information was displayed.



(a)



(b)

FIGURE 5 Diversion rate and mainline occupancy versus the time of day, (a) on I-210E at Mountain Avenue, on November 27, 2012, (b) on I-15N at University Avenue, on December 3, 2012.

in the literature review.

Correlation Analysis

This analysis compares the diversion rate with and without CMS accident messages, controlling for mainline occupancy. Through this analysis, we gain insights into the correlation between diversion rate and the existence of CMS accident messages. Therefore, the analysis is termed correlation analysis.

Setup

We compare the diversion rate R, controlling for mainline occupancy and the availability of CMS accident messages, both of which are the explanatory variables. The availability of CMS accident messages is a binary variable, x_1. For x_1 to be properly defined, we need to estimate the travel time from the CMS to the off-ramp of interest. This is because when drivers see the message, they have to wait until arriving at the off-ramp of interest to take any action if they so desire.

We use the flow and occupancy measured from all the loop detectors between the CMS and the off-ramp to estimate the travel time. The travel time estimation is similar to the G-factor method in (*16*), with small differences. Details on the estimation algorithm are not of interest here. The interesting part is that we validate the travel time estimates against travel time measurements from FasTrak (the electronic toll collection system in California) data. The validation on a 17.3-km (10.8-mile) section of US-101S in Caltrans District 4 shows a root mean square error (RMSE) of 1.3 minutes. The validation on a shorter 4.8-km (3.0-mile) section of I-80W also in Caltrans District 4 shows a RMSE of 1.0 minutes. Therefore, we expect the accuracy of the travel time estimation at the three study sites (with a length of 11.1-13.8 km, or 6.9-8.6 miles) to be approximately one minute.

Assume that a CMS accident message starts at time t_1 and stops at time t_2 and that the estimated travel time from CMS to the off-ramp is T_1 and T_2 if drivers arrive at the CMS at time t_1 and t_2, $[t_1 + T_1, t_2 + T_2]$ is the period of time when the CMS message is effective at the location of the off-ramp. Therefore, we label $x_1 = 1$ when time $t \in [t_1 + T_1, t_2 + T_2]$ and $x_1 = 0$ otherwise.

We use the mainline occupancy as the indicator for visible congestion. Instead of using the occupancy as a continuous variable, we will use two binary variables x_2 and x_3 to describe the qualitative level of visible congestion: $x_2 = 0$, $x_3 = 0$ when mainline occupancy $0 < O_{ML} < 0.15$; $x_2 = 1$, $x_3 = 0$ when $0.15 \leq O_{ML} < 0.35$; $x_2 = 0$, $x_3 = 1$ when $O_{ML} \geq 0.35$. So, x_2 is an indicator of medium congestion, x_3 is an indicator of heavy congestion, and x_2 and x_3 will not be 1 at the same time.

N-way ANOVA

To see if the average diversion rate is different when x_1, x_2, x_3 take different values, we first carry out an N-way analysis of variance (ANOVA). The diversion rate R is the response variable and x_1, x_2, x_3 are the explanatory variables. The model will have all linear terms x_1, x_2, x_3, interaction terms $x_1 x_2$, $x_1 x_3$, and the constant term.

TABLE 2 shows the result at Mountain Avenue on I-210E. The average diversion rate is significantly different across all the groups, at least at this off-ramp location. This means all the terms involved, including the linear terms and the interactions terms, may have a potential effect on diversion rate, which warrants further analysis.

TABLE 2 N-way ANOVA for diversion rate (R) versus CMS accident messages (x_1) and mainline occupancy (x_2, x_3) at Mountain Avenue on I-210E.

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob > F
x_1	0.9802	1	0.98019	810.18	0
x_2	0.5311	1	0.53109	438.98	0
x_3	4.3548	1	4.35479	3599.48	0
x_1x_2	0.1209	1	0.12088	99.92	0
x_1x_3	1.0375	1	1.03752	857.57	0
Error	78.0432	64507	0.00121		
Total	89.1941	64512			

TABLE 3 Nonparametric regression for diversion rate (R) versus CMS accident messages (x_1) and mainline occupancy (x_2, x_3).

Term	Estimate	S.E.	tStat	pValue
Constant	0.057975	0.00014116	410.72	0
$x_1(1-x_2)(1-x_3)$	0.00051013	0.0011830	0.43123	0.6663
x_2	0.036795	0.00072682	50.625	0
x_3	0.10542	0.0015939	66.139	0
x_1x_2	0.067651	0.0066119	10.232	1.4935e-24
x_1x_3	0.20155	0.0067623	29.804	7.139e-194

Nonparametric Regression
Given that all the coefficients in the N-way ANOVA are statistically significant, we further carry out a nonparametric regression, to determine how much effect each of these factors contributes to the diversion rate. The diversion rate R is the response variable and $x_1(1-x_2)(1-x_3)$, x_2, x_3, x_1x_2, x_1x_3 are the explanatory variables.

TABLE 3 shows the result at Mountain Avenue on I-210E. The coefficient of $x_1(1-x_2)(1-x_3)$ is the effect of CMS accident messages without congestion, which is very small and insignificant. The rest of the terms are all statistically significant. The constant term is the diversion rate without congestion or CMS accident messages, whose value 0.058 is consistent with our observation in the case studies. The coefficients of x_2 and x_3 are the effect of medium and heavy congestion, respectively, without CMS accident messages. The coefficients of x_1x_2 and x_1x_3 are the effect of CMS accident messages in medium and heavy congestion, respectively.

Results from other ramps show mixed results. Four other ramps (I-210E at Milliken, I-15N at University, El Cajon, and Adams) show similar results: The coefficients for both x_1x_2 and x_1x_3 are significant. At the rest of the ramps, either there is no data for statistical inference, or the coefficients for x_1x_2 and x_1x_3 are not statistically significant.

Findings

Although site-dependent, there is evidence for positive correlation between the existence of CMS accident messages and higher diversion rate. However, we cannot conclude that the higher diversion rate is due to CMS accident messages. There could be other factors, such as drivers being more likely to seek out traffic information (through radio, for example) when congestion becomes visible. The correlation analysis may still be useful, though, for the purpose of prediction, if all the other factors (beyond congestion and CMS) that affect the diversion behavior remain the same.

Causality Analysis

This analysis compares the diversion rate right before and after CMS accident messages are turned on or off, also controlling for mainline occupancy. It is termed causality analysis because it is more likely to reveal the real effect of CMS on drivers' diversion behavior. The underlying assumption, similar in nature to regression discontinuity, is that in a small time window before and after CMS accident messages are turned on or off, other factors that have an effect on the diversion rate will not change significantly. Therefore, the change in diversion rate should be attributed to CMS accident messages. The caveat is that if some other sources of information happen to occur at exactly the same time when CMS accident messages are turned on or off, their effects would be captured as the effect of CMS, and our estimate would be biased.

Setup

We will compare the diversion rates before and after time $t_1 + T_1$ and $t_2 + T_2$ at all the off-ramps between the CMS and the accident location, assuming the same definition of t_1, t_2, T_1, and T_2 as in the correlation analysis. We need to do some local smoothing of the data (diversion rate and mainline occupancy) with kernel functions (or weighting functions) to obtain the average diversion rate before and after time $t_1 + T_1$ (or $t_2 + T_2$), because simply using a single data point before and after would be very noisy. Rectangular kernel carries equal weights and will yield the arithmetic mean, which is a reasonable choice. But it also makes sense for data points to carry less weight as they get further away from time $t_1 + T_1$ (or $t_2 + T_2$), because other factors are more likely to affect diversion rate. We would use a unilateral triangular kernel for this purpose.

Assume $\Delta T = 30$ seconds is the sampling interval, $Q_{ML}(t)$ and $Q_R(t)$ are the mainline flow and ramp flow at time t. Define $Q(t) = Q_{ML}(t) + Q_R(t)$, $R(t) = Q_R(t)/Q(t)$. The average diversion rate before time $t_i + T_i$, $\bar{R}_{ib}$ and that after time $t_i + T_i$, $\bar{R}_{ia}$ can be expressed as:

$$\bar{R}_{ib} = \frac{\sum_{k=1}^{N} R(t_i + T_i - k\Delta T)Q(t_i + T_i - k\Delta T)w(k)}{\sum_{k=1}^{N} Q(t_i + T_i - k\Delta T)w(k)}, \tag{2a}$$

$$\bar{R}_{ia} = \frac{\sum_{k=1}^{N} R(t_i + T_i + k\Delta T)Q(t_i + T_i + k\Delta T)w(k)}{\sum_{k=1}^{N} Q(t_i + T_i + k\Delta T)w(k)}, \tag{2b}$$

where $i = 1$ for CMS accident messages turned on, $i = 2$ for CMS accident messages turned off. The kernel function $w(k)$ is defined as:

$$w(k) = \begin{cases} 1, & \text{for rectangular kernels } k \in [1, N] \\ N - k + 1. & \text{for unilateral triangular kernels } k \in [1, N] \end{cases} \tag{3}$$

Note that when calculating the average diversion rate, the weight should include total flow as well as the kernel function. Otherwise, we would be favoring time periods with low flow, and the average diversion rate would be biased.

There is a parameter for window size, $N\Delta T$, which is the time period to perform the local smoothing over. We will perform sensitivity analysis with both the rectangular and unilateral triangular kernels and a range of window sizes to eliminate the artifacts introduced by our choice of parameters.

We need to control for mainline occupancy, which is known to have an effect on diversion, so we also calculate the average occupancy before and after in a similar manner:

$$\bar{O}_{ib} = \frac{\sum_{k=1}^{N} O(t_i + T_i - k\Delta T)w(k)}{\sum_{k=1}^{N} w(k)}, \tag{4a}$$

$$\bar{O}_{ia} = \frac{\sum_{k=1}^{N} O(t_i + T_i + k\Delta T)w(k)}{\sum_{k=1}^{N} w(k)}. \tag{4b}$$

At the end, we define the difference in diversion rate and mainline occupancy:

$$\Delta R = \begin{cases} \bar{R}_{ia} - \bar{R}_{ib}, & \text{if } i = 1 \\ \bar{R}_{ib} - \bar{R}_{ia}, & \text{if } i = 2 \end{cases} \tag{5a}$$

$$\Delta O = \begin{cases} \bar{O}_{ia} - \bar{O}_{ib}, & \text{if } i = 1 \\ \bar{O}_{ib} - \bar{O}_{ia}. & \text{if } i = 2 \end{cases} \tag{5b}$$

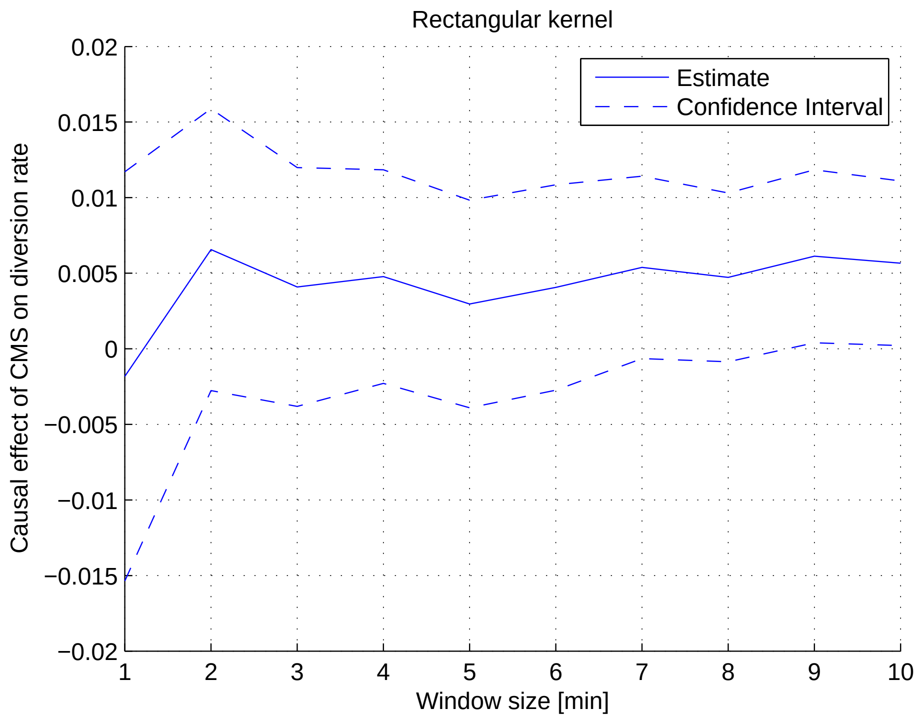
We perform a linear regression with ΔO being the explanatory variable and ΔR being the response variable. The causal effect of CMS accident messages is represented by the constant term, while the coefficient of ΔR captures the average effect of mainline occupancy.

Linear Regression

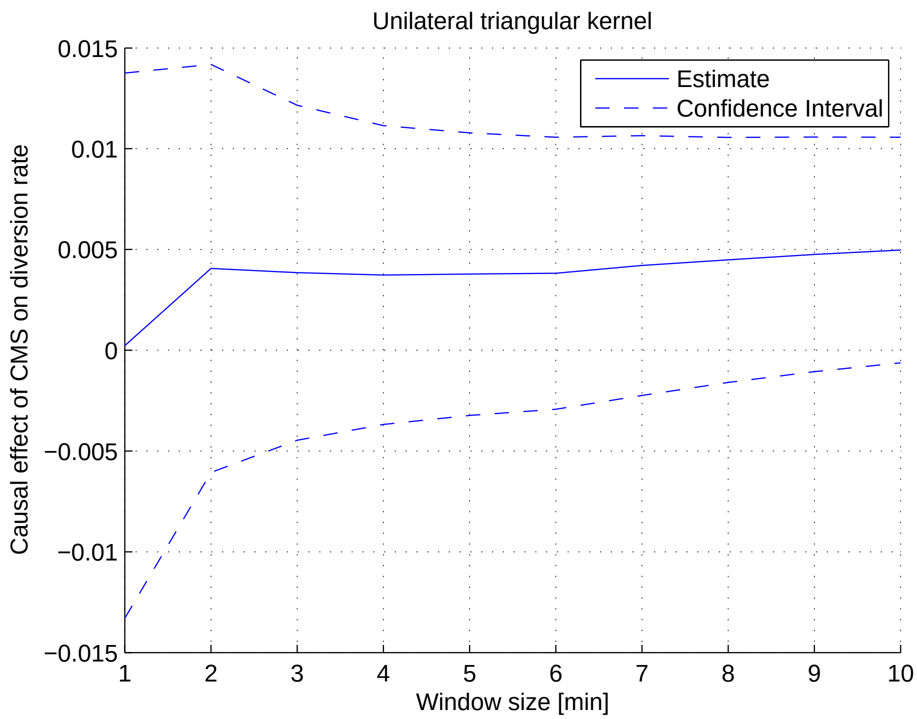
TABLE 4 shows the result of the aforementioned linear regression for the rectangular kernel with a 10-minute window size. Over all the data points, the estimate of the constant term is small, about 0.006, and not very significant, with a p-value of 0.04. When we distinguish the cases when CMS accident messages are turned on versus off, the result is similar: The estimate is small and not significant. The statistics indicate that there is no immediate effect when CMS accident messages are turned on or off.

To make sure the result is not sensitive to our choice of kernel functions and window sizes, we repeat the same analysis for both the rectangular kernel and the unilateral triangular kernel, and a range of window sizes from 1 to 10 minutes. The estimate of the constant term, which is the real effect of CMS accident messages, and its 95% confidence interval are shown in FIGURE 6. The results are similar and do not seem to be affected too much by our choice of kernel functions and window sizes.

It is recognized that drivers will not use all the off-ramps to reroute themselves. Therefore, we double-checked all the data points, especially those with high ΔR values. Very often, high ΔR value happens in the off-peak period, when mainline and ramp flow is low and granularity produces large variation on diversion rate, as shown in FIGURE 5. A larger window size would have smoothed out more noise. We have checked all the data points and did not find any obvious pattern in diversion rate that can be attributed to CMS accident messages.



(a)



(b)

FIGURE 6 Estimate of the causal effect of CMS accident messages on diversion rate and its 95% confidence interval. (a) Rectangular kernel with varying window size. (b) Unilateral triangular kernel with varying window size.

TABLE 4 Linear regression for difference in diversion rate ($\triangle R$) versus difference in mainline occupancy ($\triangle O$). Data for the linear regression are generated with the rectangular kernel and a 10-minute window size.

		Estimate	S.E.	tStat	pValue
All data	Constant	0.0056585	0.0027506	2.0572	0.041948
	$\triangle O$	0.16244	0.036121	4.4971	1.6666e-05
CMS accident messages turned on	Constant	0.00264	0.0035381	0.74617	0.45869
	$\triangle O$	0.15752	0.037236	4.2303	8.7254e-05
CMS accident messages turned off	Constant	0.0085968	0.0044	1.9538	0.055724
	$\triangle O$	0.16944	0.086808	1.9519	0.055958

Findings

We think it is safe to conclude that CMS accident messages do not seem to have much causal effect on the diversion rate. Note that this finding is consistent with some of the previous studies that are carried out carefully (*10, 4*). Also note that the conclusion here is limited to the data used in this study.

Another important conclusion is that the misuse of correlation analysis to draw causal inferences, as in some of the previous studies, will lead to false conclusions. Note that from the same dataset, the correlation analysis suggests positive correlation between diversion rate and the existence of CMS accident messages, while causality analysis suggests little effect of the CMS accident messages. Also note that the effect of visible congestion has been accounted for in both analyses. A possible explanation for the different results between the two analyses is that some other information sources (other than CMS or visible congestion, e.g., radio or smartphone app) play a role as well, prompt more drivers to change their routes, and yield a higher diversion rate. It is unclear what they are and how it works.

CONCLUSION AND FUTURE RESEARCH

There are two purposes for this paper: to explore the real effect of CMS accident messages on drivers' route choice behavior, and to compare the results from two commonly used statistical analyses. The change in drivers' behavior is indicated by the change in the diversion rate, i.e., the proportion of total flow heading to the off-ramp at a diverging location. Results from two case studies show that CMS accident messages have only minor effect on the diversion rate, and visible congestion plays a much more dominant role. Accounting for the effect of visible congestion, results from the second statistical analysis show that the causal effect of CMS accident messages on the diversion rate is insignificant. The first statistical analysis, on the other hand, shows positive correlation between CMS accident messages and diversion rate. It is likely that other factors played a role in changing the diversion rate. Through the comparison of the two analyses, we see that the misuse of the correlation analysis to draw causal inferences will lead to false conclusions. Again, due to the data-driven nature of the study, the conclusions are only valid with respect to the data used in this study.

There are a few possible directions for future research:

1. The study obviously should be expanded to include more study sites, if data are available.

2. The CMS accident messages in this study are descriptive (i.e., where and what happened) in nature. There is some evidence, in the literature and through our investigation, that prescriptive messages (i.e., CMS tells drivers to take certain routes) are more effective. Although prescriptive CMS messages are much less frequent, further investigation of them is needed.

3. The flow data used in this study only provide insights into aggregated behavior represented by the change of diversion rates. GPS data carry much richer information on individual routes, and can help us better understand how drivers change their routes.

ACKNOWLEDGMENT

The authors would like to thank Zhen Sun (UC Berkeley) for constructive discussions. This research is funded by the California Department of Transportation under the Connected Corridors program.

REFERENCES

[1] Abdel-Aty, M., Abdalla, M., 2004. Modeling drivers' diversion from normal routes under atis using generalized estimating equations and binomial probit link function. Transportation 31, 327–348.

[2] Al-Deek, H., Kanafani, A., 1993. Modeling the benefits of advanced traveler information systems in corridors with incidents. Transportation Research Part C: Emerging Technologies 1, 303 – 324.

[3] Benson, B., 1996. Motorist attitudes about content of variable-message signs. Transportation Research Record: Journal of the Transportation Research Board 1550, 48–57.

[4] Bushman, R., Berthelot, C., Chan, J., 2004. Effects of a smart work zone on motorist route decisions, in: 2004 Annual Conference of the Transportation Association of Canada, pp. 1–11.

[5] Caltrans, 2013a. Changeable message signs (CMS) feed. http://www.dot.ca.gov/cwwp2/documentation/cms/cms.htm.

[6] Caltrans, 2013b. Performance measurement system (PeMS). http://pems.dot.ca.gov/.

[7] Chatterjee, K., Mcdonald, M., 2004. Effectiveness of using variable message signs to disseminate dynamic traffic information: Evidence from field trails in european cities. Transport Reviews 24, 559–585.

[8] Chorus, C.G., Molin, E.J., van Wee, B., 2006. Traveler information as an instrument to change car-drivers' travel choices: a literature review. European Journal of Transport and Infrastructure Research 6, 335–364.

[9] Chu, L., Kim, H.K., Chung, Y., Recker, W., 2005. Evaluation of effectiveness of automated work zone information systems. Transportation Research Record: Journal of the Transportation Research Board 1911, 73–81.

[10] Dudek, C.L., Stokton, W.R., Hatcher, D.R., 1982. Real-time freewayto-freeway diversion: The san antonio experience. Transportation Research Record: Journal of the Transportation Research Board 841, 1–14.

[11] Dudek, C.L., Weaver, G.D., Hatcher, D.R., Richards, S.H., 1978. Field evaluation of messages for real-time diversion of freeway traffic for special events. Transportation Research Record: Journal of the Transportation Research Board 682, 37–45.

[12] Emmerink, R.H., Nijkamp, P., Rietveld, P., Ommeren, J.N.V., 1996. Variable message signs and radio traffic information: An integrated empirical analysis of drivers' route choice behaviour. Transportation Research Part A: Policy and Practice 30, 135 – 153.

[13] Foo, S., Abdulhai, B., Hall, F., 2008. Impacts on traffic diversion rates of changed message on changeable message sign. Transportation Research Record: Journal of the Transportation Research Board 2047, 11–18.

[14] Horowitz, A., Weisser, I., Notbohm, T., 2003. Diversion from a rural work zone with traffic-responsive variable message signage system. Transportation Research Record: Journal of the Transportation Research Board 1824, 23–28.

[15] Huo, H., Levinson, D., 2006. Effectiveness of VMS Using Empirical Loop Detector Data. Technical Report UCB-ITS-PWP-2006-4. University of Minnesota.

[16] Jia, Z., Chen, C., Coifman, B., Varaiya, P., 2001. The pems algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors, in: Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE, pp. 536–541.

[17] Kanafani, A., Al-Deek, H., 1991. A simple model for route guidance benefits. Transportation Research Part B: Methodological 25, 191 – 201.

[18] Khattak, A.J., Schofer, J.L., Koppelman, F.S., 1993. Commuters' enroute diversion and return decisions: Analysis and implications for advanced traveler information systems. Transportation Research Part A: Policy and Practice 27, 101 – 111.

[19] Kuhne, R., Langbein-Euchner, K., Hilliges, M., Koch, N., 1996. Evaluation of compliance rates and travel time calculation for automatic alternative route guidance systems on freeways. Transportation Research Record: Journal of the Transportation Research Board 1554, 153–161.

[20] Lappin, J., Bottom, J., 2001. Understanding and predicting traveler response to information: a literature review. Technical Report. John A. Volpe National Transportation Systems Center.

[21] Lee, E.B., Kim, C., 2006. Automated work zone information system on urban freeway rehabilitation: California implementation. Transportation Research Record: Journal of the Transportation Research Board 1948, 77–85.

[22] Liu, Y., Horowitz, A., Effinger, J., 2011. Development of a Traffic Diversion Estimation Model for Freeway Construction Work Zones. Technical Report TPF-5(081) & DOT Contract 12063. Center for Urban Transportation Studies, University of Wisconsin - Milwaukee.

[23] Madanat, S., Yang, C., Yen, Y., 1995. Analysis of stated route diversion intentions under advanced traveler information systems using latent variable modeling. Transportation Research Record: Journal of the Transportation Research Board 1485, 10–17.

[24] Mannering, F., Kim, S.G., Barfield, W., Ng, L., 1994. Statistical analysis of commuters' route, mode, and departure time flexibility. Transportation Research Part C: Emerging Technologies 2, 35–47.

[25] Peeta, S., Ramos, J., Pasupathy, R., 2000. Content of variable message signs and on-line driver behavior. Transportation Research Record: Journal of the Transportation Research Board 1725, 102–108.

[26] Polydoropoulou, A., Ben-Akiva, M., Khattak, A., Lauprete, G., 1996. Modeling revealed and stated en-route travel response to advanced traveler information systems. Transportation Research Record: Journal of the Transportation Research Board 1537, 38–45.

[27] Robinson, E., Jacobs, T., Frankle, K., Serulle, N., Pack, M., 2012. Deployment, Use, and Effect of Real-Time Traveler Information Systems. Technical Report NCHRP W192 (NCHRP Project 08-82). Westat and University of Maryland-CATT. Rockville, MD.

[28] Turner, J.M., Dudek, C.L., Carvell, J.D., 1978. Real-time diversion of freeway traffic during maintenance operations. Transportation Research Record: Journal of the Transportation Research Board 683, 8–10.

[29] Ullman, G., 1996. Queuing and natural diversion at short-term freeway work zone lane closures. Transportation Research Record: Journal of the Transportation Research Board 1529, 19–26.

[30] Ullman, G.L., 1992. Natural Diversion at Temporary Work Zone Lane Closure on Urban Freeways in Texas. Technical Report FHWA/TX-92/1108-6. Texas Transportation Institute.

[31] Wardman, M., Bonsall, P., Shires, J., 1997. Driver response to variable message signs: a stated preference investigation. Transportation Research Part C: Emerging Technologies 5, 389 – 405.

[32] Wu, Y.J., Hallenbeck, M.E., Wang, Y., 2011. Impacts of freeway traffic congestion on en-route traveler's diversion, in: Yin, Y., Wang, Y., Lu, J., Wang, W. (Eds.), 11th International Conference of Chinese Transportation Professionals (ICCTP) 2011, pp. 4077–4090.

[33] Xu, T., Sun, L., Peng, Z.R., 2011. Empirical analysis and modeling of drivers' response to variable message signs in shanghai, china. Transportation Research Record: Journal of the Transportation Research Board 2243, 99–107.

[34] Yim, Y., Ygnace, J.L., 1996. Link flow evaluation using loop detector data: Traveler response to variable-message signs. Transportation Research Record: Journal of the Transportation Research Board 1550, 58–64.