

# Evaluation of Operational Effects of Joint Managed Lane Policies

Chih-Lin Chung, Ph.D.<sup>1</sup>; and Will W. Recker<sup>2</sup>

**Abstract:** This paper presents a method to evaluate the operational effects of managed lane policies—vehicle eligibility, access control, pricing, and the number of managed lanes—that form a policy combination set. Two macroscopic methods are developed to prescreen the set via simple criteria, followed by integer linear programming with multiple objectives and constraints to identify the noninferior policies among the downsized set. The approach is demonstrated on the Southern California SR-57 corridor. The application eliminates twelve of possible twenty policy combinations by the macroscopic methods, and generates four noninferior policies—the existing high-occupancy vehicle lane operation and three additional potential high-occupancy toll lane policies—in terms of maximum vehicle and passenger throughput, minimum vehicle hour traveled, and travel time variance. The prescreening efficiency of the macroscopic stage, ranging from 0 to 95%, is affected by the initial policies and traffic conditions. It is concluded that the approach can substantially assess a larger policy set and effectively identify the operational effects of joint managed lane policies. DOI: [10.1061/\(ASCE\)TE.1943-5436.0000385](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000385). © 2012 American Society of Civil Engineers.

**CE Database subject headings:** High occupancy vehicles; Tolls; Traffic management.

**Author keywords:** High-occupancy vehicle lane; High-occupancy toll lane; Managed lane; Policy evaluation; Highway operations.

## Introduction

Managed lane policies may refer to aspects of the operation that include vehicle occupancy, exempt vehicles, separation types, access types, directional regulations, hours of operation, and tolls. These policies are commonly categorized as vehicle eligibility, access control, and pricing. However, while attention has been paid to the effect of a single policy by holding or assuming all other policies unchanged, there currently is no effective way to comprehensively evaluate the joint effects of these policies. For example, Chu et al. (2007) evaluated operational performance of high-occupancy vehicle (HOV) lanes under various vehicle eligibilities; Nesamani et al. (2010) investigated the operational and air-quality impacts of allowing hybrid vehicles use of HOV lanes; Jang et al. (2009) contrasted empirical traffic-incident data regarding two types of HOV access control; FHWA (2003) suggested considerations in the pricing of managed lanes—high-occupancy toll (HOT) lanes specifically—including conditions in which HOV lane demand exceeds the capacity of a single lane but nonetheless cannot justify the expansion of the facility by adding an additional HOV lane, or conditions in which HOT might be added when demand for an existing HOV lane is below its capacity while the parallel general purpose (GP) lanes are congested during peak periods.

Other typical studies have focused on the evaluation of certain predetermined policy combinations (likely more comprehensive but less precise than studying a single policy). Loudon (2009) used

sketch-planning methods to assess such managed lane options as HOV, HOT, and truck-only lanes. The methods proposed enable a quick screen of the options, but the optimal policies may be forfeited without considering their potential variations, e.g., multiple access versus unlimited access, occupancy of 2+ versus 3+, and so on. Murray et al. (2001) proposed a methodology for assessing HOT lane configurations, pricing levels, carpooling attractiveness, and departure times. A specific set of preselected scenarios, instead of an overall screening, was analyzed. Eisele et al. (2006) built a decision-support tool for converting managed lanes of one form into another. Similar to solving a multiattribute decision-making problem, three categories with 22 attributes are scored. Based on default or user-defined weights, HOV/HOT conversion is suggested if the total score exceeds a specific threshold. The main concern in applying the tool rests in the validity of the weights as well as in the subjective scoring of the attributes.

The research presented here is directed toward evaluating the operational effects of joint managed lane policies. Distinct from prior studies focusing on a single policy or preselected policy combinations, this research proposes a two-stage methodology that first screens a complete policy set and then identifies noninferior options.

## Methodology

The research question addressed is formulated as a multiobjective binary integer linear problem. There are various techniques for solving multiobjective linear problems, e.g., the weighting method, constraint method, and so on [see the review of (Hsu 1994)]. For binary integer problems, the branch and bound algorithm is the most widely used method [see Ravindran et al. (1987) for details]. Simple enumeration with sensitivity analysis is applied here since the set of feasible solutions is first downsized from dozens (and potentially more) to a manageable scale via the elimination methods that are introduced in subsequent sections. Although the procedures are designed to identify the set of options that, regardless of

<sup>1</sup>Institute of Transportation Studies, 4000 AIRB, Univ. of California, Irvine, CA 92697 (corresponding author). E-mail: cchung@uci.edu

<sup>2</sup>Professor, Dept. of Civil and Environmental Engineering, 4000 AIRB, Univ. of California, Irvine, CA 92697.

Note. This manuscript was submitted on November 30, 2010; approved on December 12, 2011; published online on December 14, 2011. Discussion period open until December 1, 2012; separate discussions must be submitted for individual papers. This paper is part of the *Journal of Transportation Engineering*, Vol. 138, No. 7, July 1, 2012. ©ASCE, ISSN 0733-947X/2012/7-882-892/\$25.00.

the “weights” applied to the various aspects of operation, are “better” alternatives than any others (i.e., commonly referred to as the nondominated set of alternatives), it is recognized that, eventually, there will be only one policy combination actually adopted for each direction.

A policy vector  $\Omega$  that consists of tolls ( $\delta^{\text{toll}}$ ), vehicle eligibility ( $\delta^{\text{psgr}}$ ), access control ( $\delta^{\text{access}}$ ), and the number of managed lanes ( $\delta^{\text{lane}}$ ) is defined as:

$$\Omega = [\delta^{\text{toll}}, \delta^{\text{psgr}}, \delta^{\text{access}}, \delta^{\text{lane}}] \quad (1a)$$

$$\delta^{\text{toll}} = \begin{cases} 0, & \text{HOV} \\ 1, & \text{HOT} \end{cases} \quad (1b)$$

$$\delta^{\text{psgr}} = \begin{cases} 2, & \text{passengers } 2+ \\ 3, & \text{passengers } 3+ \end{cases} \quad (1c)$$

$$\delta^{\text{access}} = \begin{cases} 0, & \text{end access} \\ 1, & \text{multiple access} \\ 2, & \text{unlimited access } \forall \delta^{\text{toll}} = 0 \end{cases} \quad (1d)$$

$$\delta^{\text{lane}} = \begin{cases} 1, & \text{one managed lane} \\ 2, & \text{two managed lanes} \end{cases} \quad (1e)$$

For example,  $\Omega = [\delta^{\text{toll}}, \delta^{\text{psgr}}, \delta^{\text{access}}, \delta^{\text{lane}}] = [0, 2, 1, 1]$  represents a facility comprising a single multiaccess HOV2+ lane. Focusing on lane management instead of major capital construction, the number of managed lanes is adjusted within a fixed total lane number. As shown in Table 1, enumeration of the elements in the policy vector leads to 20 combinations for each highway direction, where

$$A_p^d = \begin{cases} 1, & \text{policy adopted} \\ 0, & \text{otherwise} \end{cases}, \quad \text{for direction } d = 1, 2 \quad \text{and} \\ p = 1, 2, \dots, 20 \quad (2)$$

**Table 1.** Managed Lane Policy Combinations

Policy	$\delta^{\text{toll}}$	$\delta^{\text{psgr}}$	$\delta^{\text{access}}$	$\delta^{\text{lane}}$	Description
A1	0	2	0	1	One end-access HOV2+ lane
A2	0	2	0	2	Two end-access HOV2+ lanes
A3	0	2	1	1	One multiaccess HOV2+ lane
A4	0	2	1	2	Two multiaccess HOV2+ lanes
A5	0	2	2	1	One unlimited-access HOV2+ lane
A6	0	2	2	2	Two unlimited-access HOV2+ lanes
A7	0	3	0	1	One end-access HOV3+ lane
A8	0	3	0	2	Two end-access HOV3+ lanes
A9	0	3	1	1	One multiaccess HOV3+ lane
A10	0	3	1	2	Two multiaccess HOV3+ lanes
A11	0	3	2	1	One unlimited-access HOV3+ lane
A12	0	3	2	2	Two unlimited-access HOV3+ lanes
A13	1	2	0	1	One end-access HOT2+ lane
A14	1	2	0	2	Two end-access HOT2+ lanes
A15	1	2	1	1	One multiaccess HOT2+ lane
A16	1	2	1	2	Two multiaccess HOT2+ lanes
A17	1	3	0	1	One end-access HOT3+ lane
A18	1	3	0	2	Two end-access HOT3+ lanes
A19	1	3	1	1	One multiaccess HOT3+ lane
A20	1	3	1	2	Two multiaccess HOT3+ lanes

The set of values of  $A_p^d$  that form the nondominated set of policies considered for implementation is determined by solving an optimization problem according to the multiobjective considerations of (1) maximizing total vehicle throughput, (2) maximizing total passenger throughput, (3) minimizing total vehicle hours traveled, and (4) minimizing travel time variance. Prior to the optimization, the set of alternative policies is reduced to a manageable level by a macroscopic process that filters feasible options according to criteria based on equilibrium speed considerations and demand considerations.

## Macroscopic Policy Filters

To render the multiobjective integer optimization problem more manageable, certain policy options can be rejected a priori based on prevailing conditions. Here, we present two such filters to eliminate policies that can be expected to be suboptimal: (1) policy elimination by equilibrium speed and (2) policy elimination by demand.

### Policy Elimination by Equilibrium Speed

Conversion of the existing (before) managed lane policy to another (after) expands the combination set to a 20(before)  $\times$  20(after) matrix, as shown in Table 2, that contains four operational effects, as follows:

1. GP lane-favored effect ( $T_{\text{gpl}}$ ). Such an effect occurs when, given all other variables fixed, higher values of the toll or access variable increase the GP lane speed by introducing vehicles to the managed lane.  $T_{\text{gpl}}$  is suggested for when the operational performance of the existing managed lane is more than it warrants.
2. Managed lane-favored effect ( $T_{\text{ml}}$ ). Such an effect occurs when, given all other variables fixed, higher values of the eligibility or lane number variable increase the managed lane speed.  $T_{\text{ml}}$  is suggested for when the operational performance of existing managed lane is less than it warrants.
3. Uncertain effect ( $T_{\text{u}}$ ). Such an effect occurs when one policy change introduces vehicles to the managed lane but another takes some vehicles away.
4. No effect ( $T_{\text{s}}$ ). Both the managed and GP lanes remain at the same speed because the considered and existing policies are identical.

These effects can be identified through the concept of “speed equilibrium,” which designates a series of managed lane speeds with respect to GP lane speeds (from free flow to jam states) that lead directly to the favorability of one policy versus another. Here we present two candidate criteria—both based on travel utility concepts—for a priori selection: (1) the condition under which the ratio of marginal utilities of total passenger travel time (with respect to speed) is equal to unity, i.e., the respective speeds at which a decrease (increase) of travel time in the managed lane is compensated by an increase (decrease) of travel time in the GP lane and (2) the condition under which both lane types result in indifferent utility regarding passenger travel time, i.e., the aggregated travel time performances for the managed lane and the GP lane are identical.

Note that, by definition, space mean speed ( $S$ ) is determined by  $n$  vehicles spending individual travel times ( $t_i$ ) traversing stretch  $L$  (km), or  $S = nL / \sum t_i$ . For a 1-km stretch with density  $K$  (veh/km),  $\sum t_i = K/S$  in terms of total vehicle travel time, or, given average vehicle occupancy (AVO),  $\sum t_i^p = K \cdot \text{AVO} / S$  in terms of total passenger travel time. Let  $r = \text{AVO}_{\text{ml}} / \text{AVO}_{\text{gpl}}$  and assume a Greenshields speed-density ( $S$ - $K$ ) relationship for both lane types as  $S = \alpha K + ffs$ , where  $\alpha$  is a negative coefficient



and  $ffs$  is free flow speed. Then, total passenger travel time  $T^p$  for both lane types in the 1-km stretch is simply

$$T^p = \sum_{i \in \text{gpl}} t_i^p + \sum_{i \in \text{ml}} t_i^p = \frac{K_{\text{gpl}} \text{AVO}_{\text{gpl}}}{S_{\text{gpl}}} + \frac{K_{\text{ml}} \text{AVO}_{\text{ml}}}{S_{\text{ml}}} \\ = \frac{(S_{\text{gpl}} - ffs) \text{AVO}_{\text{gpl}}}{\alpha S_{\text{gpl}}} + \frac{(S_{\text{ml}} - ffs) \text{AVO}_{\text{ml}}}{\alpha S_{\text{ml}}} \quad (3)$$

$$T^* = \frac{\alpha T^p}{\text{AVO}_{\text{gpl}}} = 1 - \frac{ffs}{S_{\text{gpl}}} + r - \frac{ffs}{S_{\text{ml}}} r \quad (4)$$

The priority of the managed lane is founded on its higher average vehicle occupancy. The average vehicle occupancy ratio,  $r$ , is greater than 1 and about 2 (or 3) under HOV2+ (or HOV3+). This approximation is based on California HOV surveys in which (1) single-occupancy motorcycles and hybrid vehicles are allowed to use HOV lanes, (2) some HOV-eligible vehicles stay in the GP lanes, and (3) bus volumes are relatively low [California Department of Transportation (Caltrans) 2008, 2009].

Then, under the first criterion, i.e., the condition under which the ratio of marginal (dis)utilities of total passenger travel time (with respect to speed) is equal to unity, the condition of speed equilibrium can be represented as

$$\frac{\partial T^p / \partial S_{\text{gpl}}}{\partial T^p / \partial S_{\text{ml}}} = \frac{\partial T^* / \partial S_{\text{gpl}}}{\partial T^* / \partial S_{\text{ml}}} = \frac{\frac{ffs}{S_{\text{gpl}}^2}}{\frac{ffs \cdot r}{S_{\text{ml}}^2}} = \frac{S_{\text{ml}}^2}{r S_{\text{gpl}}^2} = 1, \quad \text{or} \quad S_{\text{ml}} = \sqrt{r} S_{\text{gpl}} \quad (5)$$

Under the second criterion, i.e., the condition under which both lane types result in indifferent (dis)utility regarding passenger travel time, the equilibrium condition is easily shown to be given as

$$\frac{\sum_{i \in \text{gpl}} t_i^p}{\sum_{i \in \text{ml}} t_i^p} = 1 \Rightarrow S_{\text{ml}}(S_{\text{gpl}} - ffs) = S_{\text{gpl}}(S_{\text{ml}} - ffs)r \quad \text{or} \\ S_{\text{ml}} = \frac{ffs \cdot r \cdot S_{\text{gpl}}}{(r - 1)S_{\text{gpl}} + ffs} \quad (6)$$

Given a free flow speed of  $120 \pm 8$  km/h, speed equilibrium lines associated with Eqs. (5) and (6) become equilibrium zones for different  $r$ , for each of the respective criteria, as shown in Figs. 1 and 2, by enumerating speed pairs  $(S_{\text{gpl}}, S_{\text{ml}})$  within a small range of speed equilibrium, i.e.,  $0.95 < \sqrt{r} S_{\text{gpl}} / S_{\text{ml}} < 1.05$  for criterion 1 and  $0.95 < S_{\text{gpl}}(S_{\text{ml}} - ffs)r / S_{\text{ml}}(S_{\text{gpl}} - ffs) < 1.05$  for criterion 2. Speed pairs located above, below, and within the specified equilibrium zone respectively denote a current operational performance of the managed lane of more than, less than, and equal to its warranted performance, respectively. The potential policies resulting in GP lane-favored effect ( $T_{\text{gpl}}$ ), managed lane-favored effect ( $T_{\text{ml}}$ ), and no effect ( $T_s$ ) will accordingly be considered to gain or maintain equilibrium. The logic provides an easy approach to screening the potential policies and reduces the policy set.

### Policy Elimination by Demand

Among the four effects shown in Table 2, the uncertain effect accounts for 204 out of 400 results; the managed lane-favored effect and the GP lane-favored effect each accounts for 80; and no effect accounts for 20. The predominance of cases associated with the uncertain effect is because the uncertain effect may be either that the managed lane-favored effect raises speeds on the managed lane or the GP lane-favored effect does the opposite. The number of uncertain cases can be reduced by demand analysis, which may eliminate some HOV combinations. (HOT elimination requires more information on how motorists make lane choices and is discussed in the following sections.)

Consider a corridor network that consists of two ends,  $n$  interchanges,  $r$  traffic zones, one or two HOV lanes, and multiple GP lanes. Each interchange originally connects to several zones, which are now reidentified to one. Such reidentification is done on the demand level, whereas the network itself remains unchanged. The analysis requires current origin-destination matrixes available via a calibrated demand model or other data collection techniques. These  $(r \times r)$  matrixes will be downsized to  $(n + 2) \times (n + 2)$  after zone reidentification. The maximum possible throughput appearing on the HOV lane and the minimum possible throughput appearing on the GP lanes can be computed via the demand matrixes. If these throughputs are compared on a per-lane basis and the maximum

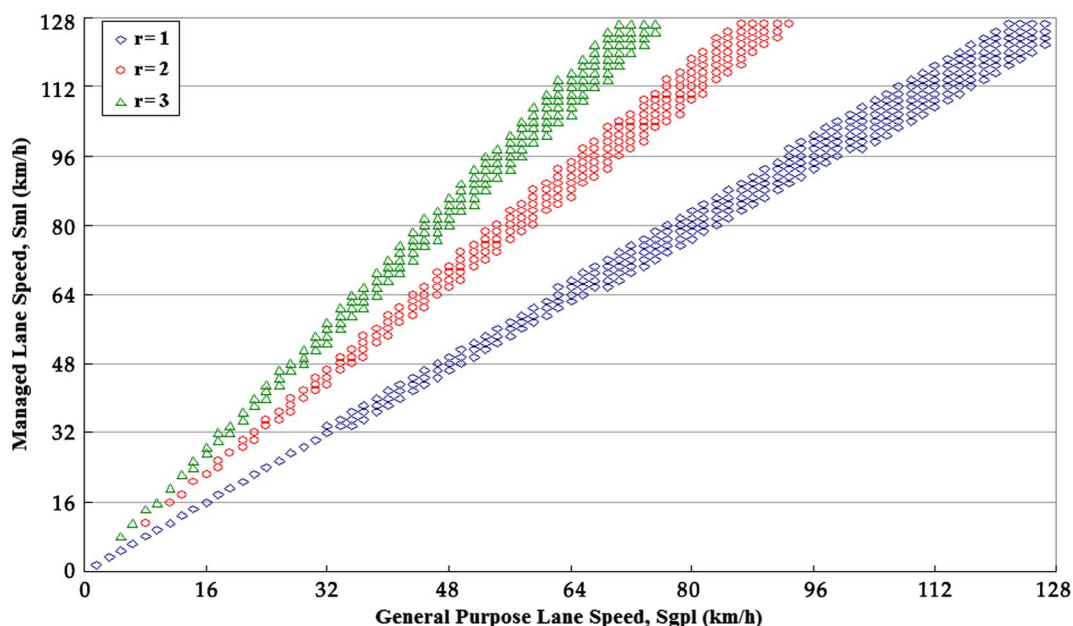
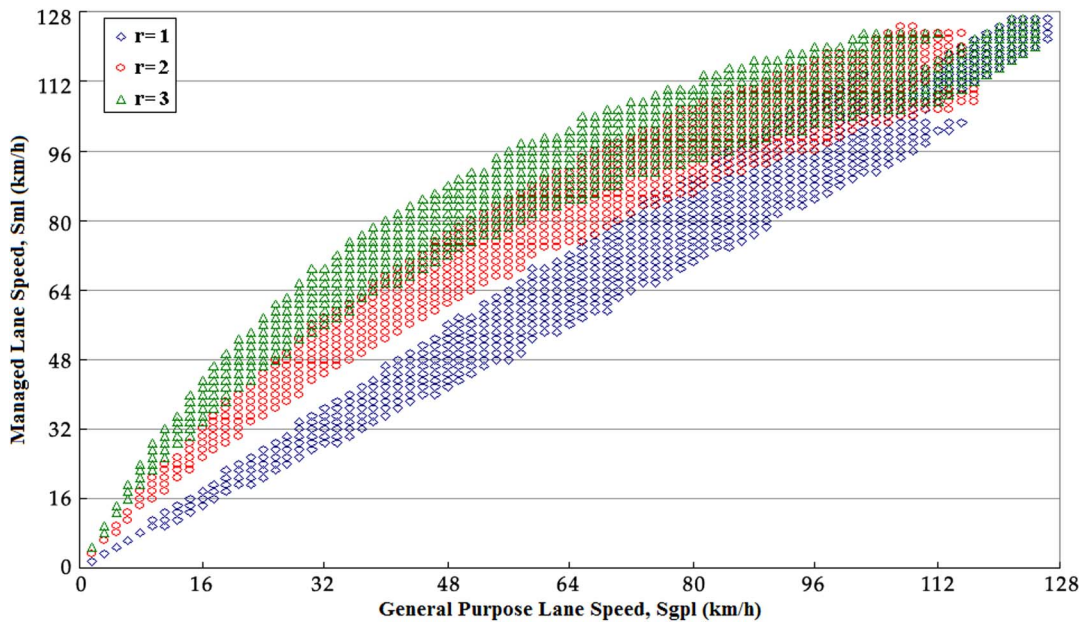


Fig. 1. Speed equilibrium status with respect to AVO ratio  $r$  associated with criterion 1: equal marginal (dis)utilities





**Fig. 2.** Speed equilibrium status with respect to AVO ratio  $r$  associated with criterion 2: equal (dis)utility of total travel time

HOV throughput is determined, as suggested by Caltrans (2003), to be within 800 to 1,650 vehicles per hour per lane (vphpl), additional HOV policy combinations may be eliminated.

### Optimization Problem Formulation

In this section, we present the details of the binary multiobjective optimization model used to identify the subset of noninferior lane-management policy options associated with the reduced set of options identified by the preceding macroscopic policy filter analysis.

#### Exogenous Variables

To identify the noninferior policy combinations, a number of exogenous variables are introduced. Although, ostensibly, the values of these variables for any case existing in the field would be available, corresponding values for other options would not. However, as we show in the case study example, the values of these variables for such options could be obtained through simulation. We group these exogenous variables as a performance vector, vehicle hours traveled, and a coefficient of travel time variation.

$$\mathbf{D}_{\text{pwt}}^d = \begin{bmatrix} \mathbf{D}_{\text{MLpwt}}^d \\ \mathbf{D}_{\text{GPLpwt}}^d \end{bmatrix} \quad (7a)$$

The vector  $\mathbf{D}_{\text{pwt}}^d$  contains performance variables of the managed and GP lanes. The vector components are designed to capture aspects pertaining to throughput ( $Q$ ), passenger throughput ( $PT$ ), mean speed ( $S$ ), standard deviation of speed ( $SDS$ ), and occupancy ( $Occ$ ) under direction  $d$  (1 for south-/eastbound and 2 for north-/westbound), policy combination  $A_p$ , time segment  $t$ , and vehicle detector station  $w$ , i.e.,

$$\mathbf{D}_{\text{MLpwt}}^d = \left[ Q_{\text{MLpwt}}^d, PT_{\text{MLpwt}}^d, S_{\text{MLpwt}}^d, SDS_{\text{MLpwt}}^d, Occ_{\text{MLpwt}}^d | A_p^d \right] \quad (7b)$$

$$\mathbf{D}_{\text{GPLpwt}}^d = \left[ Q_{\text{GPLpwt}}^d, PT_{\text{GPLpwt}}^d, S_{\text{GPLpwt}}^d, SDS_{\text{GPLpwt}}^d, Occ_{\text{GPLpwt}}^d | A_p^d \right] \quad (7c)$$

In the methodology presented here, the aspects identified in Eqs. (7b) and (7c) are characterized by a corresponding set of four binary variables:

1. Valid managed lane (vehicle) throughput variable:  $Q_{\text{pwt}}^d$

$$Q_{\text{pwt}}^d = \begin{cases} 1, & \text{if } 1650 > Q_{\text{MLpwt}}^d > 800 \text{ vphpl} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

The function of this variable is to examine if the managed lane is well utilized based on the suggested vehicle throughput from the California HOV guideline.

2. Acceptable managed lane speed variable:  $S_{\text{pwt}}^d$

$$S_{\text{pwt}}^d = \begin{cases} 1, & \text{if } S_{\text{MLpwt}}^d > 45 \text{ mph} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

The function of this variable is to examine whether or not speed in the managed lane is maintained above the minimum required by the Safe, Accountable, Flexible, Efficient Transportation Equity Act.

3. Less managed lane speed dispersion variable:  $SD_{\text{pwt}}^d$

$$SD_{\text{pwt}}^d = \begin{cases} 1, & \text{if } \frac{SDS_{\text{MLpwt}}^d}{S_{\text{MLpwt}}^d} < \frac{SDS_{\text{GPLpwt}}^d}{S_{\text{GPLpwt}}^d} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

The function of this variable is to examine if the speed dispersion (in terms of the coefficient of variation) in the managed lane is less than that in the GP lanes.

4. Level-of-service equilibrium variable:  $LE_{\text{pwt}}^d$

The level of service of the managed and GP lanes can be classified into three treatment categories: “Do Nothing” (DN), “Lane Management” (LM), and “More Than Lane Management” (MTLM) (Chung and Recker 2009), where

- DN if the managed and GP lanes are operating at no worse than level-of-service *C* and *D*, respectively; the relationship is regarded as compatible;
- LM if either lane type is congested; the system could possibly be improved simply by lane management;
- MTLM if both lane types are congested or one is congested but the other cannot be manipulated enough to remove the congestion.

Contrasting levels of service of the managed and GP lanes will determine the equilibrium, as illustrated in Table 3.

$$\begin{aligned}
 (Occ_{MLpwt}^d, Occ_{GPLpwt}^d) &\Rightarrow (LOS_{MLpwt}^d, LOS_{GPLpwt}^d) \Rightarrow LE_{pwt}^d \\
 &= \{DN, LM, MTLM\} \\
 LE_{pwt}^d &= \begin{cases} 1, & \text{if DN} \\ 0, & \text{otherwise} \end{cases} \quad (11)
 \end{aligned}$$

5. Vehicle hours traveled variable:

$$VHT_p^d = \sum_v TT_{pv}^d \quad (12)$$

where  $TT_{pv}^d$  is the travel time of vehicle *v* under direction *d* and policy combination  $A_p$ .

6. Coefficient of travel time variation variable:

$$CTV_{pij}^d = \frac{\sqrt{\frac{\sum_v (TT_{pijv}^d - \overline{TT}_{pij}^d)^2}{z_{pij}^d}}}{\overline{TT}_{pij}^d} \quad (13)$$

where  $TT_{pijv}^d$  is the travel time of vehicle *v*,  $\overline{TT}_{pij}^d$  is the mean travel time, and  $z_{pij}^d$  is the demand from zone *i* to *j* under direction *d* and policy combination  $A_p$ .

### Objective Functions

We identify a set of four competing (and possibly conflicting) objectives that the managed lane policy seeks to obtain: (1) maximizing total vehicle throughput, (2) maximizing total passenger throughput, (3) minimizing total vehicle hours traveled, and (4) minimizing travel time variance.

1. Max total vehicle throughput  $VT^d$  [refer to Eqs. (7b) and (7c)]

$$VT^d = \sum_p \sum_w \sum_t A_p^d (Q_{MLpwt}^d + Q_{GPLpwt}^d) \quad \text{for } d = 1, 2 \quad (14)$$

2. Max total passenger throughput  $TPT^d$  [refer to Eqs. (7b) and (7c)]

$$TPT^d = \sum_p \sum_w \sum_t A_p^d (PT_{MLpwt}^d + PT_{GPLpwt}^d) \quad \text{for } d = 1, 2 \quad (15)$$

3. Min total vehicle hours traveled  $TVHT^d$  [refer to Eq. (12)]

$$TVHT^d = \sum_p A_p^d VHT_p^d \quad \text{for } d = 1, 2 \quad (16)$$

4. Min travel time variance  $TV^d$  [refer to Eq. (13)]

$$TV^d = \sum_p A_p^d \left( \frac{\sum_{(ij) \subset d} CTV_{pij}^d}{g_p^d} \right) \quad \text{for } d = 1, 2 \quad (17)$$

where  $g_p^d$  is the number of origin-destination pairs for direction *d* and policy combination  $A_p$ .

### Constraints

In achieving the objectives defined in the previous section, the following operational constraints must be satisfied:

1. *One combination constraint*: one and only one policy combination is applied to each direction.

$$\sum_p A_p^d = 1 \quad \text{for } d = 1, 2 \quad (18)$$

2. *Valid managed lane throughput constraint*: to ensure utilization, the sum of the valid managed lane throughput variables in Eq. (8) shall be no less than the predetermined threshold  $C2^d$ , which can be determined by the base model. Let

$$\mathbf{Qs}^d = \begin{bmatrix} A_1 \sum_w \sum_t Qs_{1wt}^d \\ A_2 \sum_w \sum_t Qs_{2wt}^d \\ \vdots \\ A_{20} \sum_w \sum_t Qs_{20wt}^d \end{bmatrix} \quad \text{and} \quad \mathbf{C2}^d = \begin{bmatrix} C2^d \\ C2^d \\ \vdots \\ C2^d \end{bmatrix} \quad (19)$$

Then  $\mathbf{Qs}^d \geq \mathbf{C2}^d$  for  $d = 1, 2$

3. *Acceptable managed lane speed constraint*: to ensure efficiency, the sum of the acceptable managed lane speed variables in Eq. (9) shall be no less than the threshold  $C3^d$  determined by the base model. Let

$$\mathbf{Ss}^d = \begin{bmatrix} A_1 \sum_w \sum_t Ss_{1wt}^d \\ A_2 \sum_w \sum_t Ss_{2wt}^d \\ \vdots \\ A_{20} \sum_w \sum_t Ss_{20wt}^d \end{bmatrix} \quad \text{and} \quad \mathbf{C3}^d = \begin{bmatrix} C3^d \\ C3^d \\ \vdots \\ C3^d \end{bmatrix} \quad (20)$$

Then  $\mathbf{Ss}^d \geq \mathbf{C3}^d$  for  $d = 1, 2$

4. *Less managed lane speed dispersion constraint*: to ensure travel reliability, the sum of the less managed lane speed dispersion variables in Eq. (10) shall be no less than the threshold  $C4^d$  determined by the base model. Let

**Table 3.** Relationships between Managed Lanes and GP Lanes

Level of service	Managed lane						
	A	B	C	D	E	F	
GP lane	A	DN	DN	DN	LM	LM	LM
	B	DN	DN	DN	LM	LM	LM
	C	DN	DN	DN	LM	LM	LM
	D	DN	DN	DN	MTLM	MTLM	MTLM
	E	LM	LM	MTLM	MTLM	MTLM	MTLM
	F	LM	LM	LM	MTLM	MTLM	MTLM

Note: Level of service is determined by occupancy converted to density. See the density criteria table (Exhibit 23-2, [Transportation Research Board 2000](#)).

$$SD^d = \begin{bmatrix} A_1 \sum_w \sum_t SD_{1wt}^d \\ A_2 \sum_w \sum_t SD_{2wt}^d \\ \vdots \\ A_{20} \sum_w \sum_t SD_{20wt}^d \end{bmatrix} \quad \text{and} \quad C4^d = \begin{bmatrix} C4^d \\ C4^d \\ \vdots \\ C4^d \end{bmatrix} \quad (21)$$

Then  $SD^d \geq C4^d$  for  $d = 1, 2$

5. *Level-of-service equilibrium constraint:* to sustain level-of-service balance, the sum of the level-of-service equilibrium variables in Eq. (11) shall be no less than the threshold  $C5^d$  determined by the base model. Let

$$LE^d = \begin{bmatrix} A_1 \sum_w \sum_t LE_{1wt}^d \\ A_2 \sum_w \sum_t LE_{2wt}^d \\ \vdots \\ A_{20} \sum_w \sum_t LE_{20wt}^d \end{bmatrix} \quad \text{and} \quad C5^d = \begin{bmatrix} C5^d \\ C5^d \\ \vdots \\ C5^d \end{bmatrix} \quad (22)$$

Then  $LE^d \geq C5^d$  for  $d = 1, 2$

6. *Binary decision variable constraint:*

$$A_p^d \in \{0, 1\} \quad \text{for } d = 1, 2; \quad p = 1 \text{ to } 20. \quad (23)$$

### Noninferior Solutions

The solution to the multiobjective optimization problem presented above can be obtained by any of the classical techniques (e.g., weighting method, constraint method) producing options that cannot be improved in any one of the objectives without sacrificing something on one or more of the other objectives. Alternatively, if a priori weights are applied to the objectives, the multiple objectives in the noninferior solution set  $\{A_p^d\}$  collapse to a single-objective formulation. Whether through negotiation or the assignment of weights, since the decision variables are binary and only one policy is adopted for each direction, the noninferior solution set

$\{A_p^d\}$  ultimately leads to selection of the “best” compromise solution

$$\{A_p^d\} = \{0, 0, 0, \dots, 1, \dots, 0, 0\} \quad \text{for } d = 1, 2; \quad p = 1 \text{ to } 20 \quad (24)$$

### Case Study

To further demonstrate the procedures presented here, the methodology is applied to a real-world case study. The study site is SR-57 in Orange County, California—a 19.2-km long freeway that currently comprises one full-time HOV2+ lane and from four to five GP lanes per direction. The southbound HOV lane has 12 ingress/egress sections, whereas the northbound has 10. The configuration corresponds to  $A_3$  in Table 1. The study period is from 6:00 to 8:00 a.m. with southbound/inbound as the peak direction.

#### Policy Elimination by Equilibrium Speed

Loop detector data on a typical weekday were collected to examine the existing traffic conditions with respect to the speed equilibrium zone. There are 22 vehicle detector station sets in the southbound and 23 in the northbound direction paired at the same mileposts on the HOV and GP lanes. Scatter plots (Figs. 3 and 4) show that the majority of the speed observations in both directions were located below the equilibrium zone under both criteria, indicating that the HOV lane performance is currently less than it is capable of. Table 2 shows one unlimited and one multiaccess HOV2+ lane ( $A_5$  and  $A_{15}$ ) that will result in the removal of the GP lane-favored effect, leaving 18 noneliminated policy combinations for further consideration.

#### Policy Elimination by Demand

The existing configuration of SR-57 was coded into a simulation network serving as the base model. The network has 122 traffic zones and calibrated demand matrixes based upon observed data in 2007 through 2009 (CLR Analytics 2009). The 122 zones are reidentified into 12 new zones that consist of 2 ends and 10 interchanges, as shown in Fig. 5. Originally there are six

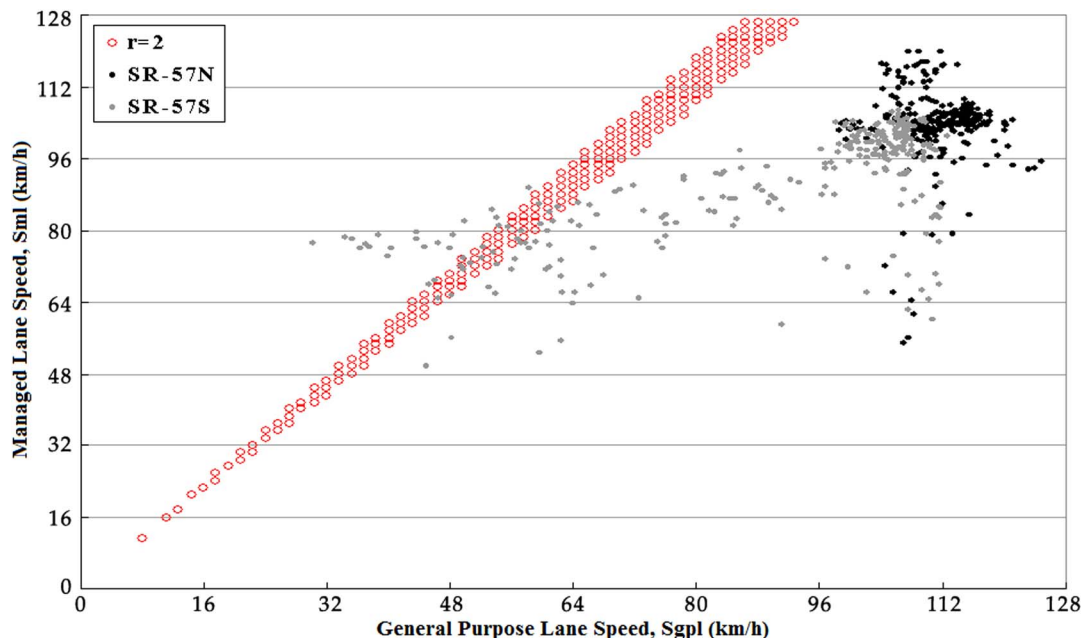


Fig. 3. SR-57 speed scatter plot regarding speed equilibrium associated with criterion 1: equal marginal (dis)utilities

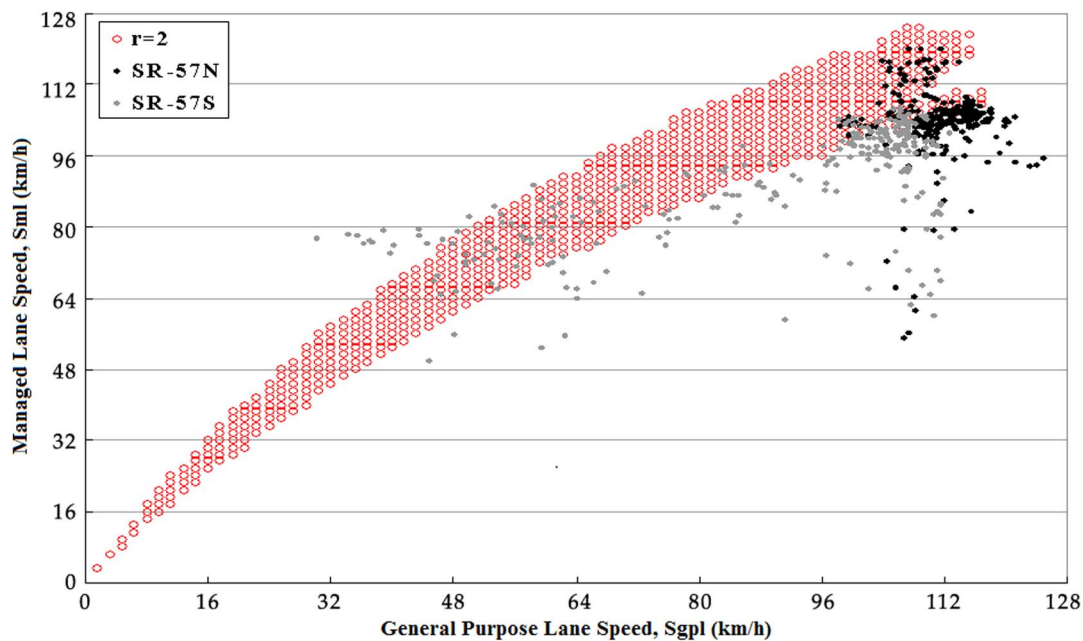


Fig. 4. SR-57 speed scatter plot regarding speed equilibrium associated with criterion 2: equal (dis)utility of total travel time

122 × 122 matrixes with respect to three time periods (one warm-up period from 5:45 to 6:00 a.m. and two hourly study periods from 6:00 to 8:00 a.m.) and two vehicle occupancies corresponding to single-occupancy vehicle and HOV2+ that respectively account for 89 and 11% of the total demand. The hourly 122 × 122 HOV demand matrixes were converted into a 12 × 12 matrix based on the reidentification of zones.

The maximum possible throughput obtained via the downsized demand matrixes appearing on the end-access HOV2+ policies— $A_1$  and  $A_2$  in Fig. 5—are either below or just above the lower bound during the peak periods. These two policies are thus removed. The unlimited-access HOV2+ lanes ( $A_6$ ) are also removed for two reasons. First, an additional southbound HOV lane converted from the adjacent GP lane not only causes the majority of HOV throughput to be around or below the lower bound but results in some sections with non-HOV throughput close to or over capacity that severely impacts the southbound GP lane system. Second,  $A_6$  northbound would generate HOV throughput significantly below the HOV lower bound and non-HOV throughput consistently at a heavy level. If, for any reason, not all HOV-eligible vehicles were to stay in the HOV lanes, the traffic on the HOV and GP lanes would become more imbalanced in both directions. For the multiaccess HOV2+ lanes ( $A_3$  and  $A_4$ ),  $A_3$  is the base model and reserved for further analysis;  $A_4$  is removed because it would cause even lower HOV throughput and greater non-HOV throughput than  $A_6$  as a result of reduced HOV accessibility.

Analysis for HOV3+ lanes ( $A_7$  through  $A_{12}$ ) follows the same procedures for HOV2+. The HOV2+ demand matrixes are divided for HOV3+ by a multiplier of 5% based on the HOV surveys regarding SR-57 (Caltrans 2005, 2008) that show HOV2 accounting for 95% of the HOV demand and the rest for HOV3+. HOV3+ hourly demand was found to be less than 500 vehicles in the whole network and HOV lane throughput far below the criteria.  $A_7$  through  $A_{12}$  are thus removed. In summary, 12 out of 20 policy combinations are eliminated before conducting optimization.

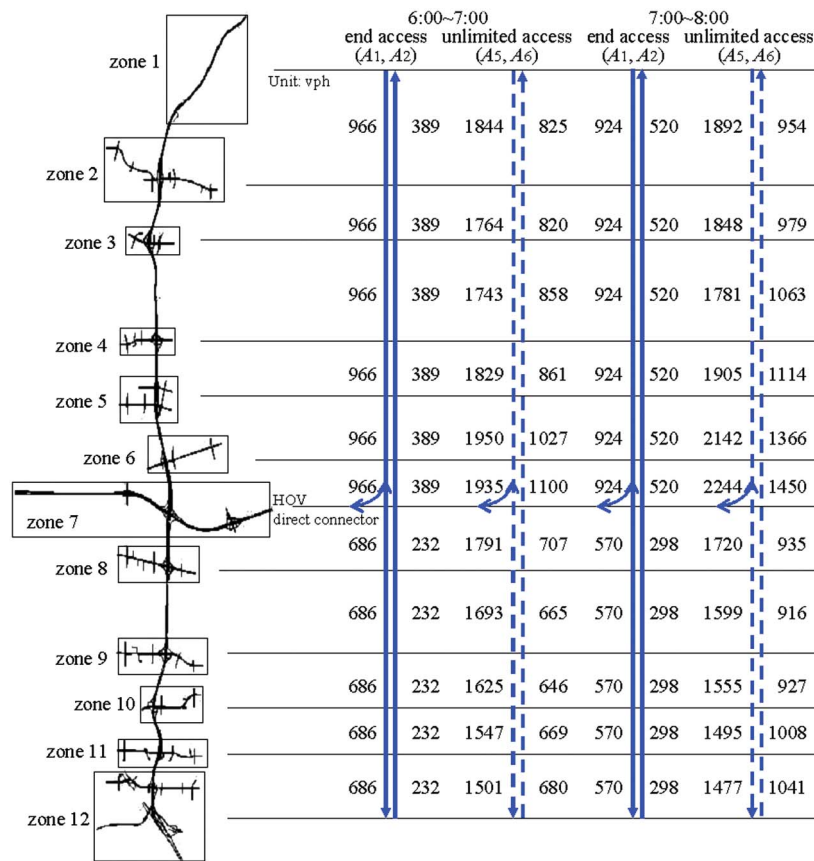
#### Optimization of the Noneliminated Policy Combinations

The noneliminated policy combinations, including one multiaccess HOV (base model), three multiaccess HOT, and four end-access

HOT, are simulated by Paramics 6.6 (2009) with identical demand matrixes. Each eligible vehicle type is assigned a value of time that averages \$25 per hour. If the monetary value of HOT time savings is greater than the toll, motorists will choose the HOT lane, otherwise the GP lanes. In keeping with current HOT practice, the simulated tolling scheme has a starting toll and tolling boundary. Once the HOT speed is less (more) than the lower (upper) tolerance speed, the toll will change by an increment (decrement) in the next feedback period. For the simulated end-access HOT lanes, there is only one tolling segment with tolls starting from \$1.50 and up to \$10 per entry to keep the HOT speed greater/less than 55/65 mph. For the multiaccess HOT lanes, there are 9 southbound and 7 northbound tolling segments along the corridor with the same ingress/egress sections as the base model; starting from \$0.15, single-occupancy vehicles will be charged from \$0.15 to \$1 in each segment with the same speed boundary as the end-access HOT. The toll increment/decrement is \$0.50 for the end-access HOT and \$0.15 for the multiaccess HOT.

Five representative detector stations, about 3.2 km apart along both directions, are set to record the performance measures on a 5-min basis. To identify the noninferiors by direction, the trips following traffic assignment are divided into southbound and northbound, whereas trips not using SR-57 (those that appear on only the local streets of the simulation network) are discarded via external computations. The simulation results are compiled to fit the optimization problem formulation. Constraint 5, which refers to HOV speed equilibrium, is omitted since the base model is the only qualification. To retain the base model in the feasible solution sets, the thresholds of the constraints are set to be a ratio  $\gamma$  of the sums of the base model's performance, e.g., the threshold is  $\gamma x_y^d$  if direction  $d$  of the base model has qualified data points  $x_y^d$  for constraint  $y$ . By setting the threshold ratio as 0.5, the constraints of valid throughput, acceptable speed, and less speed dispersion do not eliminate any policy combinations, whereas the constraint of level-of-service equilibrium removes southbound two end-access HOT2+ /HOT3+ ( $A_{14}$  and  $A_{18}$ ), two multiaccess HOT2+ ( $A_{16}$ ), and northbound end-access HOT2+ /HOT3+ ( $A_{13}$ ,  $A_{14}$ ,  $A_{17}$ , and  $A_{18}$ ) from consideration. Those inactive constraints result from either poor performance of the base model or performance of





**Fig. 5.** Maximum possible HOV2+ throughput appearing on the HOV lane(s)

other policy combinations that is almost as good as or is better than that of the base model. Raising the ratio by an increment of 0.05 would respectively activate the constraints of valid throughput, acceptable speed, and less speed dispersion at 0.55, 0.75, and 0.95 for at least one direction.

Given the one combination and binary constraints, the remaining combinations are examined under each objective. Table 4 illustrates that the base model ( $A_3$ ) is consistently optimal/noninferior for the southbound direction regardless of the objectives; the northbound noninferior set includes two multiaccess HOT2+ lanes ( $A_{16}$ ) for maximizing total vehicle throughput and passenger throughput, one multiaccess HOT3+ lane ( $A_{19}$ ) for minimizing total vehicle

hour traveled, and the base model for minimizing travel time variance. Sensitivity analysis indicates that the southbound optimal policy combination does not change with the threshold ratio, whereas the northbound noninferior set converges to the base model when the ratio increases from 0.5 to 0.9, as shown in Table 5.

The selected objectives are incommensurable and shown to be conflicting in certain circumstances, as revealed in the northbound direction of Tables 4 and 5. Passenger throughput is the occupancy count of individual vehicles passing the detector stations and seems to be nonconflicting with vehicle throughput. As vehicles on the managed lane may carry 1 (for HOT), 2, 3, and more persons, higher vehicle throughput very likely, but not necessarily, has

**Table 4.** Checklist for Policy Combinations under Various Objectives

Direction	Remaining policy combination	Max vehicle throughput $VT$	Max passenger throughput $TPT$	Min vehicle hours traveled $TVHT$	Min travel time variance $TV$
Southbound	One multiaccess HOV2+ lane: $A_3$	<b>86,480</b>	<b>106,294</b>	<b>6,395</b>	<b>0.121</b>
	One end-access HOT2+ lane: $A_{13}$	83,040	102,228	7,899	0.202
	One end-access HOT3+ lane: $A_{17}$	81,328	100,196	8,398	0.233
	One multiaccess HOT3+ lane: $A_{19}$	82,495	101,493	8,108	0.167
	Two multiaccess HOT3+ lanes: $A_{20}$	70,529	87,094	10,329	0.197
Northbound	One multiaccess HOV2+ lane: $A_3$	70,824	81,625	4,509	<b>0.086</b>
	Two multiaccess HOT2+ lanes: $A_{16}$	<b>73,515</b>	<b>86,763</b>	4,813	0.102
	One multiaccess HOT3+ lane: $A_{19}$	73,029	85,339	<b>4,489</b>	0.094
	Two multiaccess HOT3+ lanes: $A_{20}$	72,617	85,423	4,682	0.102

Note: The bold numbers are the optimums with respect to the objectives. The unit of  $VT$  is vehicles in  $2\text{ h} \times 5$  detector stations  $\times 5$  lanes, the unit of  $TPT$  is passengers in  $2\text{ h} \times 5$  detector stations  $\times 5$  lanes, the unit of  $TVHT$  is hour, and  $TV$  does not have a unit.

**Table 5.** Sensitivity Analysis Regarding Different Constraint Thresholds

Threshold ratio $\gamma$	Southbound				Northbound					
	Qualifications under constraints 2, 3, 4, 5	Noninferior set objective				Qualifications under constraints 2, 3, 4, 5	Noninferior set objective			
		1	2	3	4		1	2	3	4
0.50	A3, A13, A17, A19, A20	A3	A3	A3	A3	A3, A16, A19, A20	A16	A16	A19	A3
0.75	A3, A19	A3	A3	A3	A3	A3, A19, A20	A19	A20	A19	A3
0.85	A3	A3	A3	A3	A3	A3, A19	A19	A19	A19	A3
0.90	A3	A3	A3	A3	A3	A3	A3	A3	A3	A3

higher passenger throughput, causing occasional conflicts between these two objectives.

### Discussion of Case Study Results

Elimination by speed equilibrium zones is greatly affected by the existing policies and the speed pairs' locations with respect to the zones. For example, given speeds falling below the zone, only two policy combinations will be eliminated under one end-access HOV3+ ( $A_7$ ) but eight under one end-access HOT2+ ( $A_{13}$ ). As can be computed via Table 2, the overall elimination efficiency is 5.4 out of 20 policy combinations for speeds falling off the zone. As needed, the proposed 20 combinations could be extended to represent more scenarios by further segmenting the decision variables, for example, bus for eligibility in addition to occupancy of 2+ and 3+, use of numbers of ingress/egress sections for accessibility instead of multiaccess, and adoption of multiple levels for tolls. The whole analysis process remains valid but at a larger scale. An expanded Table 2 would be renewed based on the same principles. Accordingly, the elimination efficiency would be different.

Although HOT could theoretically fill the excess capacity with tolled non-HOV traffic, an issue about HOT configuration that affects the system performance was apparent in the simulations: the ingress/egress sections involve more lane-changing behavior than for a regular HOV lane, which in turn causes queue buildup (and resulting congestion) upstream of both HOT and GP lanes. The multiaccess HOT combinations, in particular, have nine southbound and seven northbound ingress/egress sections along the 19.2-km study corridor. Such intensive accessibility with significant lane-changing behavior potentially deteriorates traffic operations. An associated concern is that vehicles would be "trapped" awaiting gaps for lane changing even though the lane ahead is empty for quite a distance. This situation is noticeable at the end of the ingress/egress sections with vehicles trying to get in or out of the HOT lane at the last possible point. This helps explain why the base HOV model is favored over the HOT policies in the southbound direction. For the northbound SR-57 with unused HOV lane capacity during the study period, adoption of HOT does effectively raise the vehicle or passenger throughputs and shortens the total vehicle hours traveled, although the base HOV model is better in terms of lower travel time variance.

Another HOT simulation issue is the setting of value of time. To prevent the all-or-nothing assignment for HOT/GP lane choice that brings serious congestion in the HOT simulations, each eligible vehicle type is designated a value of time drawn from a distribution. While a valid distribution of value of time by vehicle type is not available, a pseudo one ranging from \$11 to \$56 per hour with an average of \$25 is adopted to ensure smoother HOT operations. The distribution is determined via trial and error as well as references from studies on California HOT lanes that show a median value of

time of between \$20 and \$45 (with the majority between \$20 and \$30) under the revealed preference survey (Brownstone and Small 2005).

### Conclusions

Evaluation of the operational effects of joint managed lane policies was conducted in two parts. First, two macroscopic screening devices were developed to prescreen the policy set via simple evaluation criteria. Second, a multiobjective optimization technique was introduced to find noninferior policies given multiple objectives and constraints. The case study showed that the existing HOV policy and three potential HOT policies on SR-57 are superior in terms of maximum vehicle and passenger throughput, and minimum vehicle hour traveled and travel time variance. The procedure is comprehensive and uncomplicated.

For highways with significant directional traffic variation (e.g., inbound morning peak in one direction and outbound evening peak in the other), separating directional traffic enables distinct optimal management lane policies instead of a single one based on aggregated data. Such separation, however, assumes operations in the two directions to be independent. If, for example, southbound off-ramp traffic blocks signalized local streets and causes delay of vehicles heading for the northbound on-ramp, then southbound and northbound traffic are not independent. Fortunately, no significant sign of dependent directional traffic was found during the SR-57 simulations. Although the two directions result in different noninferior policies, identical configurations are applied when preparing the networks, i.e., for the eight simulated policy combinations, there are 8 instead of  $8 \times 8$  networks.

Highways with distinct policies by direction can better correspond to traffic variation than those with a single policy, but a single policy typically is adopted in practice for a given highway and even an entire urban area (with a few exceptional corridors) for the sake of consistency. In such a case, the macroscopic prescreening methods are still valid, as is the optimization process to identify an optimal policy for the highway without separating directional traffic. If one highway adopts a different policy from others, decision makers may either consider a larger highway system analysis for a consistent policy or accept the exception.

The concept of speed equilibrium could be expressed in other forms, which would lead to different results. Thus speed equilibrium should be used with caution. Finally, use of identical demand and mode choice for all the policy combinations, albeit not real, is acceptable because it enables comparison of the policies on the same demand basis. Such multiobjective binary integer linear programming problems may be treated as a multiattribute decision-making problem since the optimal "product mix" is to pick one policy combination out of a set. The attributes may consist of the measures of the proposed objectives and constraints.

Assigning a weight to each attribute yields the optimal policy combination with the highest weighted score. Nonetheless, the weight setting could be debated.

Such considerations as aggregate tolls charged and percentage of tolled vehicles in HOT lanes could be added to multiobjective programming in future research.

## Notation

The following symbols are used in this paper:

- $A_p^d$  = binary variables regarding whether a policy combination is chosen;  
 $C2, C3, C4, C5$  = constraint thresholds;  
 $C2 - C5$  =  $C2$ -,  $C3$ -,  $C4$ -, and  $C5$ -based vectors;  
 $CTV_{pij}^d$  = coefficient of travel time variation variables;  
 $D_{pwt}^d, D_{MLpwt}^d, D_{GPLpwt}^d$  = performance vectors;  
 $g_p^d$  = number of origin-destination pairs that occurs in simulation hours;  
 $K_{ml}, K_{gpl}$  = density variables (veh/ln/km);  
 $LE_{pwt}^d$  = binary variables regarding level-of-service equilibrium;  
 $LE^d$  =  $LE_{pwt}^d$ -based vectors;  
 $LOS_{MLpwt}^d, LOS_{GPLpwt}^d$  = level-of-service variables;  
 $Occ_{MLpwt}^d, Occ_{GPLpwt}^d$  = occupancy variables;  
 $PT_{MLpwt}^d, PT_{GPLpwt}^d$  = passenger throughput variables (persons/5 min);  
 $Q_{MLpwt}^d, Q_{GPLpwt}^d$  = vehicle throughput variables (veh/ln/5 min);  
 $Q_{pwt}^d$  = binary variables regarding valid managed lane throughput;  
 $Q_{pwt}^d$  =  $Q_{pwt}^d$ -based vectors;  
 $(R)r$  = (adjusted) average vehicle occupancy ratio;  
 $S_{ml}, S_{gpl}, S_{MLpwt}^d, S_{GPLpwt}^d$  = mean speed variables (km/h);  
 $SD_{pwt}^d$  = binary variables regarding less managed lane speed dispersion;  
 $SD^d$  =  $SD_{pwt}^d$ -based vectors;  
 $SDS_{MLpwt}^d, SDS_{GPLpwt}^d$  = variables of standard deviation of speed (km/h);  
 $SE_{pwt}^d$  = binary variables regarding speed equilibrium for HOV lanes;  
 $SS_{pwt}^d$  = binary variables regarding acceptable managed lane speed;  
 $SS^d$  =  $SS_{pwt}^d$ -based vectors;  
 $T_{gpl}, T_{ml}, T_u, T_s$  = effects of converting existing policy to another;  
 $TPT^d$  = total passenger throughput variables (persons/2 h);  
 $TT_{pv}^d, TT_{pijv}^d$  = travel time variables (hr);  
 $\overline{TT}_{pij}^d$  = mean travel time variables (h);  
 $TV^d$  = travel time variance variables;  
 $TVHT^d$  = total vehicle hour traveled variables (h);

- $VHT_p^d$  = vehicle hour traveled variables (h);  
 $VT^d$  = total vehicle throughput variables (veh/2 h);  
 $x_y^d$  = qualified data points for base model;  
 $z_{pij}^d$  = 2-h volume (veh/2 h); and  
 $\gamma$  = ratio of sums of base model's performance for constraint thresholds.

## References

- Brownstone, D., and Small, K. A. (2005). "Valuing time and reliability: Assessing the evidence from road pricing demonstrations." *Transp. Res. A*, 39, 279–293.
- California Department of Transportation (Caltrans). (2003). *High-occupancy vehicle guidelines*, Sacramento, CA.
- California Department of Transportation (Caltrans). (2005). *Annual HOV Report*, District 12, Orange County, CA.
- California Department of Transportation (Caltrans). (2008). *Annual HOV Report*, District 7, Los Angeles/Ventura Counties, CA.
- California Department of Transportation (Caltrans). (2009). *Annual HOV Lane Report*, District 4, Bay Area, CA.
- Chu, L., Kalandiyur, N., and Benouar, H. (2007). "Priority-based high-occupancy-vehicle lane operation." *Proc., Transportation Research Board 86th annual meeting*, Washington, DC.
- Chung, C. L., and Recker, W. W. (2009). "An approach to assessing freeway lane management hot spots." *Transportation Research Record 2099*, Transportation Research Board, Washington, DC, 141–150.
- CLR Analytics. (2009). *Model calibration report for SR-57 CSMP*, Irvine, CA.
- Eisele, W., et al. (2006). *Evaluating criteria for adapting HOV lanes to HOT lanes: Development and application of HOT START software tool*, Texas Transportation Institute, College Station, TX.
- Federal Highway Administration (FHWA). (2003). "A guide for HOT lane development." Washington, DC.
- Hsu, J. Y. (1994). *Multi-objective decision making*, Wunan Culture Enterprise, Taipei, Taiwan.
- Jang, K., Ragland, D., and Chan, C. Y. (2009). *A comparative safety study of limited versus continuous access HOV facilities*, Univ. of California, Berkeley, CA.
- Loudon, W. (2009). "Sketch planning assessment of managed-lane options for I-5 freeway corridor master plan." *Proc., Transportation Research Board 88th annual meeting*, Washington, DC.
- Murray, P., Mahmassani, H., and Abdelghany, K. (2001). "Methodology for assessing HOT lane usage and network performance." *Transportation Research Record 1765*, Transportation Research Board, Washington, DC, 8–15.
- Nesamani, K. S., Chu, L., and Recker, W. W. (2010). "Policy implications of incorporating hybrid vehicles into high-occupancy vehicle lanes." *J. Transp. Syst. Eng. Inf. Technol.*, 10(2), 30–41.
- Paramics 6.6. (2009). Quadstone Paramics, Edinburgh, UK.
- Ravindran, A., Phillips, D. T., and Solberg, J. J. (1987). *Operations research: Principles and practice*, 2nd Ed., Wiley, New York.
- Transportation Research Board. (2000). *Highway capacity manual*, Washington, DC.