Effectiveness of variable speed limits considering commuters’ long-term response

Wei Liu, Yafeng Yin, Hai Yang

A Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
b Urban Transport Systems Laboratory (LUTS), École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
C Department of Civil and Coastal Engineering, University of Florida, 365 Weil Hall, Gainesville, FL 32611, USA

ABSTRACT

This paper examines the effectiveness of variable speed limits (VSLs) on improving traffic flow efficiency and reducing vehicular emissions in a stylized setting of morning commute where a fixed number of individuals commute from home to work through the freeway with a single recurrent bottleneck. The mechanism of interest is for a VSL system to prevent the bottleneck from being activated and thus avoid detrimental capacity drop that arises at the activated bottleneck. We firstly consider a VSL system installed along the freeway towards the bottleneck, which adjusts commuters’ cruising speeds in a continuous fashion and essentially regulates the upstream flow into the bottleneck. By investigating the resulting departure-time equilibrium of commuters, we find the VSL system can eliminate the efficiency loss caused by capacity drop, and further bound its improvements on various performance measures. We then turn to a more practical VSL system, which adjusts commuters’ cruising speeds in a discrete fashion. The conditions for such a system to improve various performance measures are established and its efficiencies are bounded. Based on empirical data, we conclude that the discrete VSL system can avoid or delay capacity drop associated with an active bottleneck and thus reduce queuing delay. It can help reduce the schedule delay cost and total emissions cost. However, it is unlikely for the system to reduce total travel time, individual travel cost and social cost in this particular setting. These results shed light on the effectiveness of VSL systems on realistic freeway networks.

1. Introduction

Imposing speed limits is a widely used and effective tool for speed regulation to reduce traffic accidents and prevent road casualties. As static speed limits designed for ideal or prevailing conditions may not be effective during adverse weather or heavy congestion, variable speed limits (VSLs) have been proposed to commend safe driving speeds during less than ideal conditions (see, e.g., Elefteriadou et al., 2012, for a recent review of implementations of VSLs in the U.S. and Europe). Considerable efforts have been made to evaluate the effects of VSLs on reducing crash potentials and enhancing safety (e.g., Boyle and Mannering, 2004; Lee et al., 2006; Abdel-Aty et al., 2006).

On the other hand, VSL control has emerged as a widespread scheme to improve traffic flow efficiency on freeways. A contemporary VSL system consists of a series of traffic sensors, VSL signs, variable message signs and a central processing...
unit, which integrates all system sensors and signs to determine time-varying speed limits and display them in a timely fashion (Lin et al., 2004). VSL systems can be designed to harmonize speeds across lanes to delay the onset of traffic breakdown and attain higher capacity and critical density, or reduce the probability of breakdown (e.g., Zackor, 1979; Smulders, 1990; Papageorgiou et al., 2008; Geistefeldt, 2011). More proactively, by forcing traffic to slow down in a controlled manner, VSL systems can be utilized as mainline metering devices to prevent traffic flow breakdown and avoid capacity drop at active bottlenecks (e.g., Hegyi et al., 2005a,b; Carlson et al., 2010). Although theoretically appealing, these VSL systems may not have significant impacts on traffic conditions if drivers do not comply with displayed speed limits for various reasons (e.g., Nissan and Koutsopoulos, 2011).

This study conducts a theoretical analysis of the effectiveness of VSL systems on reducing travel cost and vehicular emissions. The mechanism of interest is for VSL systems to regulate traffic inflow to prevent bottlenecks from being activated and thus avoid detrimental capacity drops that arise at activated bottlenecks (e.g., Carlson et al., 2010). In contrast to the abundant literature on VSLs that mainly focuses on their localized impacts on traffic flow and microscopic behaviors of drivers, this study examines VSL systems from an equilibrium perspective and considers individual travelers’ long-term response to VSL systems. Attempts have been made to investigate how speed limits reallocate traffic flow across a network in an equilibrium manner (e.g., Yang et al., 2012, 2013b; Wang, 2013). However, when variable or time-varying speed limits are concerned, the reallocation effect on traffic would no longer be limited to the spatial dimension, e.g., the choice of routes, but also the temporal dimension, e.g., departure-time choices. For instance, if a relatively low speed limit is imposed at a given time, some drivers may choose to depart at another time to avoid it. Such changes in departure times of travelers will yield different time-dependent traffic demands and inevitably affect the effectiveness of VSL systems.

As we seek to mitigate congestion closely associated with rush hour commute trips, we conduct our analysis in this paper based on Vickrey’s bottleneck model (Vickrey, 1969). In this model, a fixed number of individuals commute from home to work through a potential bottleneck. If the bottleneck is activated, a first-in-first-out (FIFO) queue will develop and, consequently, commuters incur a queuing delay, in addition to a free-flow cruising travel time to the bottleneck. Commuters also incur a schedule cost if they arrive at work earlier or later than desired arrival times. Assuming that each individual chooses his or her departure time to minimize a commute cost that combines (constant) cruising time, queuing delay and schedule cost, Vickrey’s analysis shows how individuals’ choice of departure time shapes traffic congestion and provides policy implications on how to mitigate morning commute traffic congestion. Given its analytical tractability, Vickrey’s bottleneck model has been used to study various transportation policies, such as congestion pricing, flexible work schedule, parking management and tradable mobility credits (e.g., Yang et al., 2013a; Liu et al., 2014a,b). For recent comprehensive reviews, see, e.g., Arnott et al. (1998) and de Palma and Fosgerau (2011). When a VSL system is introduced to Vickrey’s morning commute setting, it adjusts the free-flow cruising time to the bottleneck to influence commuters’ departure-time choices. A well-designed VSL system can control the traffic inflow rate to the downstream bottleneck to prevent its activation. This paper is to investigate the design of such a VSL system and bound its efficiency on reducing commuters’ travel cost, vehicular emissions and total social cost of morning commute.

The remainder of this paper is organized as follows. Section 2 reviews Vickrey’s bottleneck model and then extends it to capture the impact of capacity drop without VSL. In Section 3, an ideal VSL system is designed to prevent the onset of congestion and avoid capacity drop, and its effectiveness and efficiency are then discussed. Section 4 introduces a more practical VSL system operating in a discrete fashion, and Section 5 evaluates such a VSL system against a variety of performance measures, including total travel time, total schedule cost, total travel cost, total emission cost and total social cost. Finally, Section 6 concludes the paper.

2. Morning commute problem without VSL

2.1. Basic setting

Consider a continuum of mass \( N \) commuters traveling from home to work every morning, as shown in Fig. 1. All commuters are assumed to have identical preferences concerning the timing and cost of their trips. Departing at time \( t \), a commuter will experience the following travel cost, which includes both the travel time cost and the schedule delay cost:

\[
 c(t) = \alpha \cdot T(t) + \beta \cdot \max(0, t^* - t - T(t)) + \gamma \cdot \max(0, t + T(t) - t^*),
\]

where \( T(t) \) is the travel time at departure time \( t \); \( t^* \) is the desired arrival time at the destination, i.e., workplace; \( \alpha \) is the value of unit travel time, and \( \beta \) and \( \gamma \) are the schedule penalties for a unit time of early and late arrival respectively. The travel time,

![Fig. 1. Morning commute with a single bottleneck.](http://dx.doi.org/10.1016/j.trb.2014.12.001)
T(t), contains both the free-flow cruising time and the queuing delay at the bottleneck, namely, $T(t) = t_f + q(t)/s$, where $t_f$ is the free-flow travel time, $q(t)$ is the queue the commuter departing at time $t$ would face at the bottleneck and $s$ is the bottleneck capacity. Suppose the length of the freeway before the bottleneck is $l$ (as shown in Fig. 1), and thus the free-flow speed is $v_f = l/t_f$. Consistent with empirical observations, we assume that $\gamma > \alpha > \beta > 0$. We also denote $\delta = \beta\gamma/(\beta + \gamma)$.

Assuming that each commuter chooses his or her departure time to minimize the aforementioned travel cost, a Wardrop equilibrium is achieved if no commuter could reduce his or her cost by unilaterally changing his or her departure time. The departure-time equilibrium is depicted in Fig. 2(a), where the departure rates from home (also the arrival rates at the bottleneck) for commuters who arrive at the destination before and after $t^*$ are as follows:

$$r_1 = \frac{\alpha}{\alpha - \beta} s; r_2 = \frac{\alpha}{\alpha + \gamma} s. \quad (2)$$

The earliest departure time from home and the latest arrival time at the destination are given by:

$$t_s = t^* - \frac{\delta N}{\beta} - t_f; t_e = t^* + \frac{\delta N}{\gamma}.$$

respectively. The longest queue is $\delta N/\beta$. The equilibrium individual travel cost can be expressed as a function of total number of commuters, $N$, as follows:

$$c(N) = \alpha t_f + \frac{\delta N}{s}. \quad (4)$$

Note that the equilibrium is unique in the sense that the departure rate and queue profile are unique. Interested readers may refer to the related literature, e.g., Lindsey (2004).

2.2. Capacity drop

Many have discovered and discussed the phenomenon of capacity drop associated with active bottlenecks. Hall and Agyemang-Duah (1991) and Banks (1991) first suggested that the discharge flow at a bottleneck would diminish once
queues start forming upstream of the bottleneck. Empirical evidences have been provided in the literature. For example, Cassidy and Bertini (1999) reported that the average vehicle discharge rates from two freeway bottlenecks in and near Toronto, Canada can be 10% lower than the flows measured prior to the activation of the bottlenecks. Zhang and Levinson (2004) observed 3–12% capacity drop on freeway bottlenecks in the Twin Cities area, Minnesota. Various explanations have been suggested to the causes of these capacity drops, such as systematic lane changes near an active bottleneck or longer response times of drivers in reaction to disturbances around a bottleneck (e.g., Cassidy and Ahn, 2005; Laval et al., 2005; Laval and Daganzo, 2006; Treiber et al., 2006; Srivastava and Geroliminis, 2013; Chen et al., 2014). Others argued that freeway breakdown itself is intrinsically stochastic and thus the capacity can only be defined as a function of the breakdown probability (e.g., Brilon et al., 2005; Lorenz and Elefteriadou, 2001).

To be general, we assume in this paper when the queue at the bottleneck reaches a certain level, denoted as \( q_c \), the discharge rate or the service capacity of the bottleneck will drop by a certain amount, i.e., \( s = l_{c/q} \) where \( l < 1 \), representing the percentage of the remaining capacity. Without loss of generality, it is assumed that \( q_c < \frac{2}{3} N \), implying that capacity drop would arise in all the scenarios we consider in this paper if VSL is not introduced. Note that the queue in Vickrey’s analysis is a point queue. The critical queue length, \( q_c \), may correspond to a physical queue forming at a bottleneck that triggers the reduction in the discharge flow, as discussed in Cassidy and Rudjanakanoknad (2005). Alternatively, it may correspond to a certain traffic density, beyond which the bottleneck capacity would drop, as observed in Chung et al. (2007). We further assume that once the service capacity has dropped, it would not recover until the queue at the bottleneck clears.

Fig. 2(b) depicts the departure-time equilibrium when the bottleneck capacity drops after the queue length exceeds \( q_c \). The dashed lines represent the departure/arrival patterns when capacity drop is not considered, i.e., the same ones in Fig. 2(a). Note that in Fig. 2(b), \( e = -\frac{1}{\beta} \left( 1 - \frac{2}{3} \right) q_c \), which increases with the critical length \( q_c \). In Fig. 2(b), \( r_1 \) and \( r_2 \) are given by Eq. (2), and

\[
r_1' = \frac{\alpha}{\alpha - \beta} s', r_2' = \frac{\alpha}{\alpha + \gamma} s'.
\]

In the above, when \( s' \rightarrow s \), i.e., the capacity reduction approaches zero, we have \( r_1' \rightarrow r_1 \) and \( r_2' \rightarrow r_2 \). The earliest departure time from home and the latest arrival time at the destination are as follows:

\[
t'_s = t' - \frac{\delta N}{\beta s'} + \frac{\delta}{\beta} \left( \frac{q_c}{s'} - \frac{q_c}{s} \right) - t_f; t'_e = t' + \frac{\delta N}{\beta s'} - \frac{\alpha}{\beta} q_c + \frac{\alpha q_c}{s'},
\]

Note that \( t'_s < t_s \) and \( t'_e > t_e \), which implies that capacity drop extends the morning peak. Similarly, when \( s' \rightarrow s \), we have \( t'_s \rightarrow t_s \) and \( t'_e \rightarrow t_e \). Indeed, \( t_s \) and \( t_e \) determined by Eq. (3) can be viewed as a special case of \( t'_s \) and \( t'_e \) of Eq. (6) when \( s' = s \).

In Fig. 2(b), compared with the departure-time equilibrium without capacity drop, more commuters \((e > 0)\) will arrive earlier than \( t^* \), as commuters departing early enough could enjoy a higher capacity. Given the critical queue length, the equilibrium travel cost can be expressed as a function of total number of commuters:

\[
c(N) = \alpha t_f + \frac{N}{\alpha} \left( \frac{q_c}{s} - \frac{q_c}{s'} + \frac{q_c}{s} \right) + \frac{\alpha q_c}{\beta s}.
\]

It is worth mentioning that Eq. (7) reduces to Eq. (4) when \( s' = s \). For later use, we represent the longest queueing delay experienced by a commuter as follows:

\[
l_q(N) = \frac{\delta N}{\alpha} - \frac{\delta q_c}{\beta s'} + \frac{\delta q_c}{\beta s}.
\]

Note that commuters departing at different times may incur different queueing delays. However, when cruising towards the bottleneck, all commuters drive at the same free-flow speed \( v_f = l/t_f \) with the same cruising time \( t_f \).

As shown in Fig. 2, capacity drop re-shapes the equilibrium departure/arrival pattern. The existence of equilibrium requires the individual travel time over the time horizon to be continuous and concave (first increases until \( t^* \) and then decreases), since the schedule delay cost is continuous and convex (decreases first until \( t^* \) and then increases) over the time horizon. As can be seen in Fig. 2(b), the discontinuity of the bottleneck capacity does not affect the continuity and convexity of the travel time profile, and thus such an equilibrium can still exist. Similar to that depicted in Fig. 2(a) when capacity drop is not considered, the equilibrium pattern with capacity drop is also unique in the sense that the departure rates over time horizon are unique. A simple proof is provided in Appendix A. The intuition is that once commuters’ departure/arrival time interval is uniquely determined, the resulting departure rate (from home) and queue length profiles over time can be uniquely determined.

3. Effectiveness of continuous VSL systems

As previously considered, the service capacity of the bottleneck will drop once its queue exceeds a certain threshold. It is thus straightforward to observe that limiting the queue length at the bottleneck may avoid the capacity drop. One way to achieve this is to introduce a VSL system along the freeway towards the bottleneck. The system can regulate the upstream flow to the bottleneck to control its queue. In our deterministic setting with long-term equilibrium, a traffic-responsive VSL...
system functions similarly as pre-determined time-varying speed limits. Below we discuss how speed limits vary over time to avoid capacity drop. For the purpose of bounding the efficiency of VSL, we assume that drivers perfectly comply with speed limits.

3.1. Design of VSL

We first consider an ideal case where vehicles are equipped with an on-board unit that displays the advised speed limit; each individual driver entering the freeway at time $t$ is advised a speed limit $\bar{v}(t)$, which remains constant for the particular driver along the freeway segment towards the bottleneck. However, the speed limit varies continuously with time, implying that each individual driver may face a different speed limit. Fig. 3(b) presents a particular design of such a time-varying speed limit, described as follows:

$$\bar{v}(t) = \begin{cases} v_f & t < t_1 \text{ or } t > t_3 \\ v_f - v_1 \cdot (t - t_1) & t_1 \leq t \leq t_2 \\ \bar{v}_{\min} + v_2 \cdot (t - t_2) & t_2 < t \leq t_3 \end{cases}$$

where $t_1 = t' - \frac{N}{l_f} - t_f$, $t_2 = t' - \frac{N}{l_f} - t_f$ and $t_3 = t' + \frac{N}{l_f} - t_f$, which are indicated in Fig. 3(b). More specifically, the highest speed limit $v_f = l_f/t_f$ is assigned to the first and last commuters (it is not beneficial to the system to assign a lower speed limit to them), while the lowest speed limit $\bar{v}_{\min} = l_{\text{cr}}/(t_f + \frac{N}{l_f})$ is assigned to the commuter arriving on time, where $t_f + \frac{N}{l_f}$ is the longest cruising time under the proposed ideal VSL, which is identical to the longest travel time in Fig. 2(a). In addition,
\( v_1 = (v_y - \bar{v}_{\text{min}}) / \left( \frac{\delta N}{\beta} \right) \) and \( v_2 = (v_y - \bar{v}_{\text{min}}) / \left( \frac{\beta}{\lambda} \right) \). Note that the relationship \( v_1 > v_2 \) shown in Fig. 3(b) is only illustrative. Both \( v_1 > v_2 \) and \( v_1 \leq v_2 \) can arise, and \( v_1 > v_2 \) requires \( \beta > \frac{\delta N}{\lambda^2} \). Under such a continuous VSL, all the queuing delays are eliminated while the cruising time increases. The resulting patterns of arrival at destination and departure from home as shown in Fig. 3(a) are identical to those in Fig. 2(a) where capacity drop is not considered.

The basic idea behind the discussed ideal VSL is that, when the departure rate from home is higher than the bottleneck capacity, we slow down the traffic to reduce the inflow rate to the downstream bottleneck; when the departure rate is below the capacity, we gradually speed up the traffic to reduce the cruising time (Appendix B provides a proof that the FIFO principle is still satisfied in this case). The resulting equilibrium individual travel cost will be identical to the case specified by Eq. (4) when capacity drop is not considered. This implies that the VSL system can eliminate all deadweight losses caused by the capacity drop.

The resulting equilibrium flow pattern under the ideal VSL system, depicted in Fig. 3(a), is unique. The reasoning is similar to that for the case with capacity drop discussed in Section 2. Generally speaking, given the speed limit profile, the equilibrium departure/arrival time interval can be uniquely determined; and the travel times and departure rates (from home) over time can then be determined accordingly. The effectiveness of the VSL system in reducing travel time, travel cost, schedule delay cost, emission cost or social cost is discussed in Section 3.2.

### 3.2. Effectiveness of VSL

#### 3.2.1. Total travel time

Given the percentage of remaining capacity, i.e., \( \lambda \), total travel time under the flow pattern depicted in Fig. 2(b) can be expressed as a function of \( \lambda \), i.e.,

\[
\overline{T}(\lambda) = t_f \cdot N + \frac{1}{2} q_c \cdot \left[ \frac{\delta N}{\beta} + \frac{\alpha}{\beta + \gamma} (1 - \lambda) q_c \right] + \frac{1}{2} \left( N - \frac{\alpha}{\beta} q_c \right) \cdot \left( N - \frac{\alpha}{\beta} q_c \right).
\]

Note that “~” is used to represent the case when VSL is not introduced and capacity drop occurs. When capacity drop is not considered, the equilibrium flow pattern is depicted in Fig. 2(a) and total travel time is equal to \( \overline{T}(1) \). After implementing the ideal VSL in Fig. 3(b), total travel time reduces to \( \overline{T}(\lambda) \). It is easy to verify that the first-order derivative of Eq. (10) with respect to \( \lambda \) is negative, suggesting that total travel time reaches its minimum at \( \lambda = 1 \). Therefore, by implementing the VSL depicted in Fig. 3(b), total travel time can be reduced. Moreover, the reduction, defined as follows:

\[
\Delta T(\lambda) = \overline{T}(\lambda) - \overline{T}(1),
\]

can be bounded in relation to \( \lambda \) and \( \frac{\delta}{\beta} \). To provide a comparable measure, consider the total queuing delay at the equilibrium when capacity drop is not considered:

\[
\overline{T}(1) - N \cdot t_f = \left( \frac{1}{2} \frac{\delta N}{\beta} \right) \cdot N.
\]

The above measure excludes the total free-flow time (everyone drives at the highest speed). If it is included, the result would depend on the values of \( t_f \), i.e., the length of the freeway segment. Similar treatments will be done in the analyses of travel cost, emissions cost and social cost. The measure in Eq. (12) is often chosen for efficiency analysis of time-varying pricing since it is the maximum travel time that can be reduced when capacity drop is not considered. Here it can be interpreted as the total travel time under the ideal VSL system when the original total cruising time without VSL is deducted. Now, we are ready to present the following proposition.

**Proposition 3.1.** The total travel time reduction defined by Eq. (11) satisfies the following

\[
\sigma_1 < \frac{\Delta T(\lambda)}{\overline{T}(1) - t_f \cdot N} < \sigma_2,
\]

where \( \sigma_1 = \frac{\delta}{\beta} (1 - \lambda) \left( \frac{\alpha}{\beta} + \frac{\delta}{\beta} N \right) \) and \( \sigma_2 = \frac{\delta}{\beta} - 1 \).

**Proof.** Observe that:

\[
\Delta T(\lambda) = \frac{1}{2} \frac{\delta N}{\beta} \left[ (1 - \lambda) \frac{\alpha}{\beta + \gamma} (q_c) \right] + \frac{1}{2} \left( N - \frac{\alpha}{\beta} q_c \right) \left( N - \frac{\alpha}{\beta} q_c \right).
\]

Checking the first-order derivative of Eq. (14) with respect to \( q_c \), we have

\[
\frac{\partial \Delta T(\lambda)}{\partial q_c} < 0,
\]
for \( 0 < q_c < \frac{1}{2}N \). Therefore,
\[
\frac{\delta}{\gamma} (1 - \lambda) \left( \frac{\delta}{\beta} \right) \cdot \left( \frac{1}{2} \frac{\delta N}{\alpha} \right) \cdot N < \Delta T_T (\lambda) < \left( \frac{1}{\alpha} \right) \cdot \left( \frac{1}{2} \frac{\delta N}{\beta} \right) \cdot N.
\]
Equation (13) compares the reduction in travel time achieved by implementing the ideal VSL with the total travel time under the ideal VSL system when the original total cruising time without VSL is deducted. Empirical studies (e.g., Tseng et al., 2005) show that a reasonable value of \( \lambda \) is in the range of \([2, 4]\), although it varies with the income level and other factors. As observed in previous empirical studies (e.g., Cassidy and Bertini, 1999; Zhang and Levinson, 2004; Bertini and Leal, 2005), the capacity drop is approximately 10%, and thus \( \lambda = 0.9 \). By assuming that \( \lambda = 0.9 \) and \( \frac{1}{2} \lambda \in [2, 4] \), we then have \( r_1 \in [2.04\%, 3.46\%] \) and \( r_2 = 11.11\% \).

3.2.2. Total travel cost
Similarly, total travel cost of the flow pattern depicted in Fig. 2(b) can be expressed as:
\[
\overline{T}_C (\lambda) = N \left( z t_f + \delta \left( \frac{1}{\alpha} \frac{N}{\lambda} \frac{1}{\beta} + \frac{2 q_c}{\beta} \right) \right).
\]
The ideal VSL leads to \( \overline{T}_C (1) \). The first-order derivative of Eq. (17) with respect to \( \lambda \) is negative, which suggests that the ideal VSL system reduces the total travel cost. The reduction, defined as follows, can be similarly bounded.
\[
\Delta \overline{T}_C (\lambda) = \overline{T}_C (\lambda) - \overline{T}_C (1),
\]
Proposition 3.2. The total travel cost reduction define by Eq. (18) satisfies the following
\[
\frac{\sigma_1}{\sigma_2} < \frac{\Delta \overline{T}_C (\lambda)}{\overline{T}_C (1) - z t_f \cdot N} < \frac{\sigma_2}{\sigma_3},
\]
where \( \sigma_2 = \frac{1}{2} - 1 \).

3.2.3. Total schedule delay cost
Utilizing the results in the above two subsections, total schedule delay cost of the flow pattern depicted in Fig. 2(b) can be expressed as:
\[
\overline{T}_S (\lambda) = \overline{T}_C (\lambda) - \alpha \cdot \overline{T}_T (\lambda),
\]
where \( \overline{T}_T (\lambda) \) and \( \overline{T}_C (\lambda) \) are given by Eqs. (10) and (17) respectively. The ideal VSL leads to \( \overline{T}_S (1) \). The first-order derivative of Eq. (20) with respect to \( \lambda \) is negative, suggesting that the ideal VSL system reduces the total schedule cost. The reduction, defined as follows, can be similarly bounded.
\[
\Delta \overline{T}_S (\lambda) = \overline{T}_S (\lambda) - \overline{T}_S (1).
\]
Note that \( \overline{T}_S (1) = \frac{1}{2} \frac{\alpha d_r}{\alpha} \), which is half of the total travel cost \( \overline{T}_C (1) \).

Proposition 3.3. The total schedule delay cost reduction define by Eq. (20) satisfies the following
\[
\frac{\sigma_1}{\sigma_2} < \frac{\Delta \overline{T}_S (\lambda)}{\overline{T}_S (1)} < \sigma_2,
\]
where \( \sigma_1 = \frac{1}{2} - 1 \) and \( \sigma_3 = \frac{1}{2} \left( \frac{1}{2} + \frac{1}{\alpha} - \frac{1}{r} \right) \).

3.2.4. Total emissions cost
We now investigate the impacts of VSL on total vehicular emissions, which are estimated using a modal emissions approach (e.g., Matzoros and Van Vliet, 1992; Frey et al., 2003; Zhang et al., 2013). In this approach, a vehicle's operation is divided into modes such as idle, acceleration, deceleration and cruise, and each mode is associated with a constant...
emissions factor for each primary pollutant. In our setting, there are two aggregate modes: cruising on freeway and queuing at the bottleneck. We use $e_o$ and $e_q$ to denote the monetary equivalents of emissions per unit time spent on cruising and queuing respectively and $e_q = \rho e_o, \rho > 1$. The ideal VSL scheme can eliminate all queuing and thus may reduce total emission cost.

With the above consideration, total emission cost can be expressed as follows:

$$
\bar{TE}(\lambda) = e_o \cdot N_{tf} + e_q \cdot \left( \frac{1}{2} \frac{\delta}{\beta} \frac{\lambda}{\beta + \gamma} N + \frac{\lambda (1 - \lambda)}{\beta + \gamma} \frac{q_c}{s} \right) + \frac{1}{2} \left( \frac{1}{\lambda} \frac{\delta}{\beta} \frac{N}{s} \right) \left( \frac{1}{\lambda} \frac{\delta}{\beta} \frac{q_c}{s} \right) \left( N - \frac{\lambda}{\beta} q_c \right). \tag{23}
$$

$\bar{TE}(1)$ represents the total emission cost under the equilibrium in the original setting without capacity drop. The ideal VSL system yield a total emission cost of $e_o \cdot \left( \frac{1}{2} \frac{\lambda}{\beta} \frac{N}{s} \right)$. Again, the first-order derivative of Eq. (23) with respect to $\lambda$ is negative, implying that the system reduces total emissions cost. The reduction defined below can be similarly bounded.

$$
\Delta T E(\lambda) = \bar{TE}(\lambda) - e_o \cdot \left( \frac{1}{2} \frac{\lambda}{\beta} \frac{N^2}{s} \right). \tag{24}
$$

**Proposition 3.4.** The reduction in total emissions cost define by Eq. (24) satisfies the following

$$
\sigma_1 + 1 - \rho^{-1} < \frac{\Delta T E(\lambda)}{T E(1) - e_o \cdot N_{tf}} < \chi^{-1} - \rho^{-1}, \tag{25}
$$

where $\sigma_1 = \frac{1}{\rho} (1 - \lambda) \left( \frac{\delta}{\beta} + \frac{\lambda}{\beta + \gamma} \right)$.

**Proof.** It can be verified that

$$
\Delta T E(\lambda) = e_q \cdot \left( \frac{1}{2} \frac{\delta}{\beta} \frac{\lambda}{\beta + \gamma} \frac{q_c}{s} \right)^2 + \frac{1}{2} \left( \frac{1}{\lambda} \frac{\delta}{\beta} \frac{N}{s} \right) \left( N - \frac{\lambda}{\beta} q_c \right)^2 + (e_q - e_o) \cdot \left( \frac{1}{2} \frac{\delta}{\beta} \frac{N^2}{s} \right) \cdot N. \tag{26}
$$

The remaining proof follows similarly to Proposition 3.1. $\square$

With $\chi = 0.9$ and $\frac{\delta}{\beta} \in [2, 4]$, we have $\sigma_1 \in [2.04\%, 3.46\%]$. For estimating $e_o$ and $e_q$, Table 1 shows the average emission rates for different driving modes of a gasoline powered vehicle (Frey et al., 2003).

<table>
<thead>
<tr>
<th>Mode</th>
<th>CO$_2$ (g/s)</th>
<th>NO (mg/s)</th>
<th>HC (mg/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>1.8</td>
<td>0.06</td>
<td>0.19</td>
</tr>
<tr>
<td>Acceleration</td>
<td>6.0</td>
<td>1.9</td>
<td>0.8</td>
</tr>
<tr>
<td>Deceleration</td>
<td>2.7</td>
<td>0.26</td>
<td>0.27</td>
</tr>
<tr>
<td>Cruise</td>
<td>3.9</td>
<td>0.72</td>
<td>0.42</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discount rate (%)</th>
<th>CO$_2$</th>
<th>N$_{2}$O</th>
<th>CH$_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>9.4</td>
<td>3500</td>
<td>370</td>
</tr>
<tr>
<td>3.0</td>
<td>33</td>
<td>13,000</td>
<td>810</td>
</tr>
<tr>
<td>2.5</td>
<td>52</td>
<td>20,000</td>
<td>1100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mode</th>
<th>Rate 5.0%</th>
<th>Rate 3.0%</th>
<th>Rate 2.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>0.068</td>
<td>0.239</td>
<td>0.377</td>
</tr>
<tr>
<td>Acceleration</td>
<td>0.251</td>
<td>0.886</td>
<td>1.392</td>
</tr>
<tr>
<td>Deceleration</td>
<td>0.105</td>
<td>0.368</td>
<td>0.579</td>
</tr>
<tr>
<td>Cruise</td>
<td>0.156</td>
<td>0.549</td>
<td>0.864</td>
</tr>
</tbody>
</table>

Please cite this article in press as: Liu, W., et al. Effectiveness of variable speed limits considering commuters’ long-term response. Transportation Research Part B (2014), http://dx.doi.org/10.1016/j.trb.2014.12.001
For the social cost of vehicular pollutants, Table 2 presents the estimates for marginal CO$_2$, N$_2$O, and CH$_4$ emissions in 2010 (Marten and Newbold, 2012). In the table, the discount rate is constant for the analysis horizon from 2010 to 2050. A higher discount rate means the pollutant’s damage on a future time diminishes faster.

Based on Tables 1 and 2, we calculate the monetary equivalents of vehicular emissions per unit time at different driving modes, shown in Table 3.

For a cruising vehicle, it can be observed from Table 3 that $e_v$ varies from 0.156 to 0.864 (US$/hour) with different discount rates. In our setting, we consider queuing is a stop-and-go process and acceleration is a dominant mode. Consequently, $e_q$ changes from 0.251 to 1.392 (US$/hour) and thus $\rho$ is in the range of 1.609 to 1.614. Applying $\rho = 1.61$ yields $\sigma_1 + 1 - \rho^{-1} \in [40.0\%, 41.4\%]$ and $\lambda^{-1} - \rho^{-1} = 49.0\%$. The bounds are larger than those in the analysis of total travel time, schedule delay and travel cost.

### 3.2.5. Total social cost

Lastly, we consider total social cost that includes all the cost components previously considered, i.e., travel time, schedule delay and emissions, defined as follows:

$$\bar{TSC}(\lambda) = \bar{TC}(\lambda) + \bar{TE}(\lambda).$$

where $\bar{TC}$ and $\bar{TE}$ are given by Eqs. (17) and (23) respectively. With the above results, it is straightforward to conclude that the ideal VSL system can reduce total social cost. The reduction can be defined as follows:

$$ATSC(\lambda) = \Delta TSC(\lambda) = \bar{TSC}(\lambda) - \bar{TSC}(1).$$

As shown in Proposition 3.4, the reduction can be similarly bounded.

**Proposition 3.5.** The reduction of total social cost define by Eq. (27) satisfies the following inequality

$$\sigma_1 < \frac{\Delta TSC(\lambda)}{TSC(1) - (\alpha + e_v) \cdot Ntf} < \lambda^{-1} - \rho^{-1},$$

where $\sigma_1 = \frac{\lambda}{2}(1 - \lambda)\left(\frac{3}{2} + \frac{\lambda}{2}\right)$.

This above result is straightforward. Because total social cost includes travel time cost, schedule delay cost and emission cost, the lower bound is the smallest one among the lower bounds established in Propositions 3.1, 3.2 and 3.4, while the upper bound is the largest one among the upper bounds established in Propositions 3.1, 3.2 and 3.4. Alternatively, one may follow a similar process of the proof of Proposition 3.1 to obtain these two bounds. Empirical values of the bounds are the same as those in the previous analysis.

### 3.2.6. Summary

The ideal VSL system essentially trades queuing for cruising and does not change much the pattern of departure from home. In fact, the resulting departure pattern is the same as the equilibrium pattern when capacity drop is not considered. However, since it eliminates all queuing, the capacity drop associated with the highway bottleneck is avoided and thus the loss of efficiency due to capacity drop is then eliminated. As such, it reduces total travel time and individual travel cost. Moreover, as queuing is more environmentally hazardous than cruising, the VSL system yields more reduction in the emissions cost.

It is worth noting that the ideal VSL system is not as effective as the dynamic tolls described in, e.g., Arnott et al. (1998). The latter not only eliminates all queuing but also minimizes the total cost, by changing the departure pattern from home. For example, as per Proposition 3.1, the ideal VSL system can reduce travel time up to 11.11% of the original queuing delay (without considering capacity drop). In contrast, the saving achieved by the dynamic tolls is 100%.

### 4. Design of discrete VSL systems

We now focus on a more practical VSL scheme where the speed limit changes by time of day in a discrete fashion, similar to actual implementations of VSL systems. More specifically, it is assumed that roadside VSL signs display the speed limit, which will be the same for the same driver along the freeway segment towards the bottleneck. The speed limit can vary by time of day and a well-designed scheme is expected to limit the queue length to avoid capacity drop. From the analysis in Section 3, one may expect a VSL scheme depicted in Fig. 3(c), the discretization of the continuous scheme in Fig. 3(b), to be implementable and effective. Unfortunately, as explained later in the section, the FIFO principle will be violated when the speed limit increases as specified in Fig. 3(c) and a group of drivers who depart later will be blocked by early-departing ones. Below we discuss the design of a VSL system depicted in Fig. 4.

A desirable speed limit scheme should not starve the bottleneck and waste its capacity. This implies that commuters’ arrival to the destination needs to be continuous. Without loss of generality, we equally divide the duration of commuter’s arrival into $n$ intervals or steps and specify the same speed limit for the group of commuters arriving in the same interval or step. Therefore, all commuters can be characterized into two categories: the first $k$ groups whose total number of
commuters is $N_v$ face a speed limit that decreases in the order of groups while the last $n - k$ groups, whose total number is $N_c$ (and $N_v + N_c = N$), will encounter the same speed limit. The speed limit for the $i$-th group can be specified as follows:

$$v_{f_i} = \begin{cases} v_f \cdot \frac{t_f}{t_f + (i-1) \Delta t} & i \leq k \\ v_f \cdot \frac{t_f}{t_f + k \Delta t} & i \geq k + 1 \end{cases}$$

(29)

Fig. 4. Commuting equilibrium with proposed VSL scheme.

Please cite this article in press as: Liu, W., et al. Effectiveness of variable speed limits considering commuters' long-term response, Transportation Research Part B (2014), http://dx.doi.org/10.1016/j.trb.2014.12.001
where $\Delta t_k$ is defined as follows:

$$
\Delta t_k = \frac{1}{2} \frac{\beta}{N_v} \frac{N_v}{s}. 
$$

(30)

Note that, for different groups of commuters, the lengths of their departure intervals may be different, although their arrival intervals have the same length. An appropriate combination of $N_v$, $N_c$, and $k$ can be selected to optimize a particular system performance measure, such as minimizing the travel time cost or the total social cost. However, it is socially preferable to set $k > \frac{N_c}{N_v}$, which guarantees that the first $k$ groups of commuters face a queue length less than the critical length $q_c$ and capacity drop does not arise. When $k \to \infty$ (as well as $n$), all commuters in the first $k$ groups will experience no queue.

Fig. 4 depicts the time-varying speed limits profile prescribed by Eqs. (29) and (30), and the corresponding equilibrium flow patterns. In Fig. 4(b), $\varepsilon = \frac{1}{2} \left( 1 - \frac{k}{N_v} \right) q_c$ and $N_c = \frac{1}{2} q_c$. It can be observed in Fig. 4 that the speed limit given by Eq. (29) ensures no waste of the bottleneck capacity, because, for the first $k$ groups of commuters, the first commuter in the $i$-th group will arrive at the bottleneck just after the last commuter in the $(i-1)$-th group. It is worth mentioning that, the speed limit scheme specified by Eqs. (29) and (30) uniquely corresponds to a particular combination of $N_v$ and $N_c$ for a given $k$. Note that under the proposed VSL system, capacity drop may still occur, which is depicted in Fig. 4(b) where $N_v$ is larger than a threshold $N_c = \frac{1}{2} q_c$, i.e., $N_v > N_c$. In this case, the longest queue formed by $N_v$ commuters will go beyond the critical one $q_c$ and capacity drop arises. The departure-time equilibrium depicted Fig. 4(b) includes the one in Fig. 2(b) as a special case where $N_c = N > N_v$.

We now discuss how the VSL scheme described in Fig. 3(c), which closely mimics its continuous counterpart, may fail. We begin with the last group of commuters. If they face a higher speed limit than the second last group, they will choose to depart right after the second last group in order to reduce their late-arrival penalty. With a higher speed, they will be inevitably blocked by the commuters of the second last group. It is thus not beneficial for the VSL system to advise the commuters in the last group with a higher speed limit. On the other hand, it is not beneficial either to recommend a lower speed limit. Otherwise, all commuters will be pushed to depart earlier, and none will choose to depart in the last group. Indeed, this is equivalent to implement another speed limit scheme with a smaller $N_c$. By repeating this analysis in a reverse order of groups, one can obtain the design and results shown in Fig. 4.

The equilibrium flow pattern under the VSL profile in Fig. 4 is unique. Given the time-varying speed limit and the schedule delay cost over time, the equilibrium departure/arrival time interval can be uniquely determined, and the travel times and departure rates (from home) can then be determined accordingly. The uniqueness can be explained by Fig. 5 where the scheme depicted in Fig. 4(a) is considered. In the figure, the dashed lines represent the schedule delay cost against the arrival time at the destination, and the solid lines above represent the summation of the corresponding schedule delay cost and cruising time cost. Since the time-varying speed limit is discrete (for demonstration, five short intervals are depicted in Fig. 5), the solid line is discrete.

Suppose that in Fig. 4(a), total travel demand $N = N_v$. Then in Fig. 5, the dotted line segment $x$ with a length of $N_v/s$ corresponds to the unique arrival time interval at equilibrium under the VSL in Fig. 4(a). Note that the dotted line segment $x'$ corresponds to the equilibrium arrival time interval if capacity drop is not considered. If $N$ becomes larger, e.g., $N_v$, then line segment $y$ corresponds to another new unique arrival time interval at equilibrium. To facilitate our presentation and analysis, we still consider that capacity drop will be avoided by VSL. Once the arrival time interval is given, the equilibrium queue length and departure rate profile can be uniquely determined.

As long as $N = N_v \geq N_c$, under the VSL scheme depicted in Fig. 4(a), the equilibrium is unique (note that the longest queue is still less than the critical one). There is a one-to-one correspondence between $N_v$ and line segments $y$. Given $N = N_v \geq N_c$, the arrival time interval is determined. Also note that $N = N_v \geq N_c$ implies that line $y$ is above all the disjoint short segments of the solid line. And line $x$ represents a boundary case where the line is just above the corresponding disjoint short segments of the solid line. The difference between line $x$ and the solid line is indeed the equilibrium queueing delay cost with respect to arrival time. The queue profile can be subsequently and uniquely determined.

![Fig. 5. Uniqueness of the equilibrium under the proposed VSL scheme.](http://dx.doi.org/10.1016/j.trb.2014.12.001)
When \( N \) becomes smaller, i.e., \( N < N_c \), but still \( N > N_x \) holds, the departure/arrival equilibrium may not exist, or at least not unique, because one may not be able to determine two unique intersections of a shorter line \( x \) and the solid line. For the case with \( N < N_x \), the departure/arrival equilibrium is unique (but all the commuters experience the lowest speed limit and largest cruising time). The VSL in Fig. 4(a) is evidently not well designed for such a demand level of \( N < N_x \).

5. Effectiveness of discrete VSL systems

This section examines the effectiveness of the VSL system proposed in the last section. To simplify the analysis, we consider the case where \( k \to \infty \) (and \( n \to \infty \)). Note that this simplification may only affect our analyses where the cruising time and queuing delay need to be differentiated, such as total emission cost. When we consider, e.g., total travel time, because the time commuters spend on driving or queuing makes no difference, as long as \( k > \frac{s}{N_c} \) to ensure no capacity drop for the first \( k \) groups of commuters, the analysis results will not be affected. However, with this simplification, a specific VSL scheme can be represented by its corresponding combination of \( N_v \) and \( N_c \) or simply \( N_c \).

5.1. Total travel time

5.1.1. Theoretical analysis

The proposed VSL system may reduce the total travel time as it can increase the total throughput of the bottleneck by completely avoiding capacity drop (\( N_c \leq N_x \)) or delaying it (\( N_c > N_x \)). The resulting total travel time can be expressed as a function of \( N_c \):

\[
TT(N_c) = TT_o(N_c) + TT_q(N_c),
\]

where \( TT_o \) is the total cruising time on the freeway segment, which is given by:

\[
TT_o(N_c) = (N - N_c) \cdot \left( t_f + \frac{1}{2} \frac{N - N_c}{s} \right) + N_c \cdot \left( t_f + \frac{\beta}{2} \frac{N - N_c}{s} \right).
\]

and \( TT_q \) is the total queuing delay at the bottleneck, which is given by:

\[
TT_q(N_c) = \begin{cases} 
N_c \cdot \left( \frac{1}{2} \frac{N}{s} \right) & \text{when } N_c \leq N_x \\
\frac{1}{2} \frac{N}{s} + \frac{\beta}{2} N_c \left( 1 - \frac{\beta}{\beta + 1} \right) + \frac{1}{2} l_q(N_c) \cdot [N_c - \frac{\beta}{\beta + 1} q_c] & \text{when } N_c > N_x
\end{cases}
\]

Note that \( l_q(N_c) \) is determined by Eq. (8). It is worth mentioning that \( TT_q \) and thus \( TT \) are discontinuous at \( N_c = N_x \). Regarding the total cruising time in Eq. (32) and the total queuing delay in Eq. (33), we have the following two lemmas.

**Lemma 5.1.** Under the proposed VSL scheme, total cruising time decreases with \( N_c \).

**Proof.** From Eq. (32), we have

\[
\frac{dTTo(N_c)}{dN_c} = -\frac{\beta}{2} \frac{N_c}{s} < 0.
\]

This completes the proof. \( \square \)

**Lemma 5.2.** Under the proposed VSL scheme, total queuing delay increases with \( N_c \).

**Proof.** From Eq. (33), we have

\[
\frac{dTq(N_c)}{dN_c} = \begin{cases} 
\frac{1}{2} \frac{N}{s} > 0 & \text{when } N_c \leq N_x \\
\frac{1}{2} \frac{N}{s} + \frac{\beta}{2} N_c - \frac{\beta}{\beta + 1} q_c > 0 & \text{when } N_c > N_x
\end{cases}
\]

Also note \( TT_q(N_c) > TT_q(N_x) \) if \( N_c > N_x \). Thus **Lemma 5.2** is proved. \( \square \)

**Lemma 5.1** implies that total cruising time would reach its minimum when \( N_c = N_x \), i.e., no VSL scheme is introduced. **Lemma 5.2** implies that the total queuing delay would reach its minimum when \( N_c = 0 \). These two lemmas suggest a tradeoff for reducing the total travel time.

To facilitate our analysis, we define the following function:

\[
\theta(N_c) = \frac{\delta}{\beta} \left( \frac{s}{s - \beta (s - 1) \frac{q_c}{N_c}} \right).
\]
Note that $\theta(N) > 0$, i.e., $\theta(N_c)$ increases with $N_c$. It is evident that, given $N_c$, Eq. (36) depends on $\lambda = \frac{q_c}{\beta}$. When $\lambda = \lambda_1$ where $\lambda_1$ is given below, $\theta(N) = 1$.

$$\lambda_1 = \left(1 - \frac{\alpha q_c}{\beta N}ight) \left(\frac{\alpha}{\beta} - \frac{\alpha q_c}{\beta N}\right)^{-1}.$$  \hspace{1cm} (37)

For total travel time when $N_c = N$, i.e., $TT(N)$, $\frac{dTT(N)}{dN_c} < 0$. When $\lambda \to 0$, $TT(N) \to \infty$. When we consider $\lambda = 1$, it can be verified that $TT(N) \to (t_f + \frac{1}{2} + \frac{q_c}{\beta}) \cdot N$. For $q_c < \frac{\beta}{\alpha} N$, since $TT(N_c) = N \cdot t_f + \frac{1}{2} + \frac{q_c}{\beta} \cdot N$, one can verify that $(t_f + \frac{1}{2} + \frac{q_c}{\beta}) \cdot N < TT(N_c) \to \infty$. We can then conclude that there exists a unique percentage of the remaining capacity, denoted by $\lambda_2$, that solves:

$$TT(N) = TT(N_c).$$  \hspace{1cm} (38)

Note that $\lambda_2$ can be either larger than or less than $\lambda_1$, depending on $\frac{q_c}{\beta}$. We are now ready to present the following proposition concerning total travel time.

**Proposition 5.1.** An appropriate design of the proposed VSL scheme can help reduce total travel time if and only if

$$\lambda < \min\{\lambda_1, \lambda_2\}.$$  \hspace{1cm} (39)

**Proof.** From the proofs of Lemmas 5.1 and 5.2, and Eq. (31), we have

$$\frac{dTT(N_c)}{dN_c} = \begin{cases} \frac{\alpha}{\beta} \cdot \left(\frac{\alpha}{\beta} - 1\right) & \text{when } N_c \leq N_c \\ \frac{\alpha}{\beta} \cdot (\theta(N_c) - 1) & \text{when } N_c > N_c, \end{cases}$$

where $\theta(N_c)$ is determined by Eq. (36).

If $\lambda < \lambda_1$, it can be shown that $\theta(N) > 1$, and thus $\frac{dTT(N_c)}{dN_c} \mid_{N_c=N} > 0$. This means, compared with $N_c = N$, a small decrease in $N_c$ would lead to a decrease in total travel time, i.e., an appropriate VSL scheme can reduce total travel time. If $\lambda > \lambda_1$, then $\theta(N) \leq 1$. From Eq. (40), we know that $\frac{dTT(N_c)}{dN_c} < 0$. Note that when $N_c > N_c$, travel time reaches minimum at $N_c = N$. Therefore, the VSL scheme can reduce total travel time only if $TT(N) > TT(N_c)$. As mentioned, $\frac{dTT(N_c)}{dN_c} < 0$ and $TT(N) = TT(N_c)$ when $\lambda = \lambda_2$, then we know that $TT(N) > TT(N_c)$ if and only if $\lambda < \lambda_2$.

It is possible that $\lambda_1 < \lambda_2$ or $\lambda_1 < \lambda_2$. When $\lambda_1 < \lambda_2$, the VSL can reduce total travel time if and only if $\lambda < \lambda_2 = \max\{\lambda_1, \lambda_2\}$. When $\lambda_1 > \lambda_2$, the VSL can reduce total travel time if and only if $\lambda < \lambda_1 = \min\{\lambda_1, \lambda_2\}$. The conditions are equivalent to $\lambda < \max\{\lambda_1, \lambda_2\}$. This completes the proof.  \hspace{1cm} \Box

From Proposition 5.1 and Eqs. (36)–(38), we know that whether the VSL scheme can reduce total travel time depends on the values of $\alpha$, $\beta$, $\gamma$, $\lambda$, $q_c$, and $N$. Considering $\frac{q_c}{\beta} \to 0$ as practically $q_c$ is very small as compared to $N$, $\lambda_1 \to \frac{\alpha}{\beta}$ and $\lambda_2 \to \frac{\alpha}{\beta}$. We thus have that the VSL system can reduce total travel time if and only if $\lambda < \frac{\alpha}{\beta}$. This implies if the capacity drop is relatively large, the VSL is more likely to reduce total travel time. On the contrary, if $\frac{q_c}{\beta} \to \frac{\beta}{\alpha}$, we see that $\lambda_1 \to \frac{\beta}{\alpha}$ and $\lambda_2 \to \frac{\beta}{\alpha}$. This implies that the VSL scheme is very likely to reduce total travel time, if $N$ is relatively small. However, in this case, the loss caused by capacity loss is expected to be small as well. The critical queue length $q_c$ defined in this paper is not a physical one, it is difficult, if not impossible, to estimate its value. Below we present a stronger condition that does not require knowing the value of $q_c$ (and $\alpha$ and $N$).

**Proposition 5.2.** An appropriate design of the proposed VSL scheme can help reduce total travel time if

$$\lambda \leq \frac{\gamma}{\beta + 2\gamma}.$$  \hspace{1cm} (41)

**Proof.** As assumed in Section 2 that $0 < \frac{q_c}{\beta} < \frac{\beta}{\alpha}$, we have:

$$\frac{\gamma}{\beta + 2\gamma} < \lambda_1 \leq \lambda_2 = \frac{\alpha}{\beta}.$$ \hspace{1cm} (42)

If $\lambda \leq \frac{\gamma}{\beta + 2\gamma}$, according to Proposition 5.1, the VSL scheme can reduce total travel time.  \hspace{1cm} \Box

Proposition 5.2 can be reflected in a two-dimension domain, $\left(\frac{q_c}{\beta}, \lambda\right)$ in Fig. 6 where two dash lines, i.e., Lines 1 and 2, represent $\lambda = \frac{\gamma}{\beta + 2\gamma}$ and $\lambda = \frac{\alpha}{\beta}$ respectively. These two lines divide the domain into three regions, i.e., Regions 1, 2 and 3. In Region 3, $\lambda \leq \frac{\gamma}{\beta + 2\gamma}$ and thus total travel time can be reduced regardless of $\frac{q_c}{\beta}$. This region spreads as $\lambda$ increases, indicating

Please cite this article in press as: Liu, W., et al. Effectiveness of variable speed limits considering commuters' long-term response. Transportation Research Part B (2014), http://dx.doi.org/10.1016/j.trb.2014.12.001
the VSL system is more likely to help reduce total travel time if late arrivals are more heavily penalized than early arrivals. We also notice that, if $\frac{q_cN}{\gamma} \to 0$, in Region 1, $\lambda \geq \frac{1}{\beta+\gamma}$ and total travel time cannot be reduced, while in Regions 2 and 3, $\lambda < \frac{1}{\beta+\gamma}$ and total travel time can be reduced.

5.1.2. Empirical analysis

Empirical studies suggest that a reasonable value of $c_b$ varies between [2, 4], which places $c_b + 2c$ in the range of $\frac{2}{C_2/C_3}$ and $\frac{4}{C_2/C_3}$. Consistent with field observations, we assume that the capacity drop is approximately 10%, and thus $\hat{\lambda} = 0.9$. Consequently, $\lambda > \frac{1}{\beta+\gamma}$ holds, which can also be observed in Fig. 6 where the solid line, i.e., $\lambda = 0.9$, falls in Region 1 when $\hat{\lambda}$ varies between 2 and 4. Considering practically $q_c$ is very small as compared to $N$, i.e., $\frac{q_cN}{\gamma} \to 0$, all these suggest that the proposed VSL system is unlikely to reduce total travel time even if capacity drop can be completely avoided.

5.2. Total travel cost

5.2.1. Theoretical analysis

We now examine the impact of the VSL scheme on total travel cost, which includes total travel time and commuters’ schedule cost and can be expressed as a function of $N_c$ as follows:

$$TC(N_c) = \begin{cases} N \cdot (\alpha tf + \delta N_c + \beta N_c - N \frac{N_c}{\gamma}) & \text{when } N_c \leq N_c^c, \\ N \cdot \left(\alpha tf + \delta N_c - \frac{\hat{\lambda} q_c}{\gamma} + \frac{\gamma}{\beta+\gamma} \frac{q_cN}{\gamma} + \beta N_c - N \frac{N_c}{\gamma}\right) & \text{when } N_c > N_c^c. \end{cases}$$  \hfill (43)

Similarly, the above function is neither differentiable nor continuous at the point $N_c = N_c^c$. To facilitate our further analysis, we define the following critical percentage of the remaining capacity:

$$\hat{\lambda}_3 = \left(1 - \frac{\alpha q_c}{\beta N}\right) \left(\frac{\beta}{\gamma} - \left(\frac{\alpha}{\beta} + \frac{\alpha}{\gamma} \frac{q_c}{N}\right)^{-1}. \right. \hfill (44)$$

The above critical value yields $TC(N) = TC(N_c^c)$. Examining Eq. (37), we know $\hat{\lambda}_3 > \hat{\lambda}_1$. In addition, it can be verified that $\frac{\partial TC}{\partial N_c} > 0$, $\hat{\lambda}_3 \to \frac{1}{\beta+\gamma}$ when $\frac{q_c}{N} \to 0$ while $\hat{\lambda}_3 \to 1$ when $\frac{q_c}{N} \to \frac{\gamma}{\beta}$.

**Proposition 5.3.** The proposed VSL scheme can help reduce total travel cost if and only if $\lambda < \hat{\lambda}_3$. If this condition holds, to minimize total travel cost, we should set $N_c = N_c^c$. 

![Fig. 6. The regions where total travel time can or cannot be reduced by VSL.](image-url)
Proposition 5.4. Regarding the efficiency of the proposed VSL scheme, (i) if \( N > \frac{1}{k} \), \( \phi > \frac{1 - \frac{1}{k} N}{N - \frac{1}{k} q} > 0 \); (ii) if \( \frac{1}{k} < \lambda < \lambda_3 \), the efficiency \( \phi < 1 \). Below we provide a proposition to bound the efficiency of the VSL scheme.

Proposition 5.4. Regarding the efficiency of the proposed VSL scheme, (i) if \( \lambda < \frac{1}{k} \), \( \phi > \frac{1 - \frac{1}{k} N}{N - \frac{1}{k} q} > 0 \); (ii) if \( \frac{1}{k} < \lambda < \lambda_3 \), the efficiency \( \phi < 1 \). Below we provide a proposition to bound the efficiency of the VSL scheme.

Proof. From Eqs. (4), (43) and (46), we have

\[
\phi = \frac{\alpha t_f + \delta \left( \frac{N - s}{\gamma} + \frac{q}{\beta} \right) - \alpha t_f - \delta \frac{N}{\gamma} + \delta \frac{N - s}{\gamma}}{\alpha t_f + \delta \left( \frac{N - s}{\gamma} + \frac{q}{\beta} \right) - \alpha t_f - \delta \frac{N}{\gamma}} = \frac{\left( \frac{1}{\beta} - \frac{1}{\gamma} \right) N - \frac{q}{\gamma} + \frac{q}{\beta} + \frac{s}{\gamma} - \frac{s}{\beta}}{\frac{1}{\beta} - \frac{1}{\gamma} \left( N - \frac{q}{\beta} \right)}.
\]

We note that

\[
\frac{\gamma}{\beta + \gamma} > 1 \iff -\alpha q_c + \alpha q_c + \alpha \frac{q_c}{\beta} - \alpha \frac{1}{\beta} N > 0.
\]

Therefore, from Eqs. (47) and (48), and \( \lambda < \frac{1}{k} \), we have

\[
\phi > \frac{\left( \frac{1}{\beta} - \frac{1}{\gamma} \right) N - \frac{q}{\gamma} + \frac{q}{\beta} + \frac{s}{\gamma} - \frac{s}{\beta}}{\frac{1}{\beta} - \frac{1}{\gamma} \left( N - \frac{q}{\beta} \right)} = \frac{1 - \frac{s}{\beta}}{1 - \lambda} > 0.
\]

This completes the proof of part (i).

Rearranging Eq. (47), by noting Eq. (44), we have

\[
\phi = \frac{\left( \frac{1}{\beta} - \frac{1}{\gamma} \right) N - \frac{q}{\gamma} + \frac{q}{\beta} + \frac{s}{\gamma} - \frac{s}{\beta}}{\frac{1}{\beta} - \frac{1}{\gamma} \left( N - \frac{q}{\beta} \right)} = \frac{1 - \frac{s}{\beta}}{1 - \lambda}.
\]
For $\gamma / \beta \leq \lambda < \lambda_3 < 1$, we then have
\[
\phi = 1 - \frac{1 - \frac{1}{\lambda_3}}{1 - \lambda} = 1 - \left(1 - \frac{1}{\lambda_3}\right) \frac{\gamma}{\beta} < 1.
\]
This completes the proof of part (ii).

The lower bound of the efficiency established in Proposition 5.4 provides a meaningful indicator on how the efficiency varies with the input parameters. The lower bound decreases with $\lambda$ but increases with the ratio of $c_b$. Intuitively, if all commuters have rigid work start times and are heavily penalized for being late, i.e., $\gamma \to \infty$, the bound approaches 100%, implying that the inefficiency due to capacity drop would be fully eliminated. On the other hand, if the capacity drop approaches 100%, i.e., $\lambda \to 0$, the bound approaches 100% as well. In the same domain of $\left(\frac{\gamma}{\beta}, \lambda\right)$ as Fig. 6, Fig. 7 shows the contours of the efficiency lower bound for $\gamma / \beta < \lambda_3$.

For $\gamma / \beta \leq \lambda < \lambda_3$, the upper bound of the efficiency established in Proposition 5.4 still depends on $q_c$ due to the fact that $\lambda_3$ depends on $\frac{\gamma}{\beta}$. When $\frac{\gamma}{\beta} \to 0$, $\lambda_3 \to \frac{\gamma}{\beta}$, then the upper bound $1 - \left(1 - \frac{1}{\lambda_3}\right) \frac{\gamma}{\beta} = 0$, indicating the efficiency would approach zero for $\gamma / \beta \leq \lambda < \lambda_3$. When $\frac{\gamma}{\beta} \to \frac{\gamma}{\beta} / \lambda_3 \to 1$, then the upper bound $1 - \left(1 - \frac{1}{\lambda_3}\right) \frac{\gamma}{\beta} = 1$.

Before moving to the empirical analysis, we look at more closely the total schedule delay cost, which is a part of the total travel cost as defined below:
\[
TS(N_c) = TC(N_c) - \alpha \cdot TT(N_c),
\]
where $TC$ is the total travel cost given by Eq. (43), and $TT$ is the total travel time given by Eq. (31). We have the following proposition:

**Proposition 5.5.** There always exists a design of the VSL scheme that reduces total schedule delay cost. When $N_c < \bar{N}_c$, i.e., capacity drop can be fully avoided, the cost decreases with $N_c$.

**Proof.** From Eqs. (43) and (31), differentiating Eq. (52) yields:
\[
\frac{dTS(N_c)}{dN_c} = \begin{cases} 
\frac{N_c - N_c}{\tilde{N}_c} (\delta - \beta) & \text{when } N_c \leq \bar{N}_c \\
\frac{N_c - N_c}{\tilde{N}_c} (\delta - \beta) - \frac{\beta}{\bar{N}_c} (\beta(N_c) - \tilde{N}_c) & \text{when } N_c > \bar{N}_c
\end{cases}
\]

When $N_c < \bar{N}_c$, it is straightforward to see that $\frac{dTS(N_c)}{dN_c} < 0$ since $N > N_c$ and $\delta < \beta$. Therefore, the total schedule delay cost decreases with $N_c$.

When $N_c > \bar{N}_c$, below we show that $\left.\frac{dTS(N_c)}{dN_c}\right|_{N_c = \bar{N}_c} > 0$. This implies that a VSL system will reduce total schedule delay cost because a small decrease in $N_c$ would lead to a reduction in the cost. From Eq. (53), we have:

![Efficiency Lower Bound](image_url)
\[
\frac{d\text{TS}(N_c)}{dN_c}\bigg|_{N_c=N} = -\frac{N}{\beta} \cdot \left(\frac{\beta \theta(N) - \delta}{\delta}\right) .
\]

Note that \( \theta(N) = \frac{1}{\beta} \left( \lambda - \frac{S}{\beta} \left( \frac{S}{\beta} - 1 \right) \right) \), as defined in Eq. (36). It can be verified that:

\[
s > s' \Rightarrow -\frac{1}{\beta} \left( \frac{S}{\beta} - 1 \right) \frac{q_c}{N} < 0 \Leftrightarrow -\frac{\delta}{\lambda} - \frac{\lambda}{\beta} \left( \frac{S}{\beta} - 1 \right) \frac{q_c}{N} < 0 \Leftrightarrow \theta \left( \frac{1}{\beta} \left( \lambda - \frac{S}{\beta} \left( \frac{S}{\beta} - 1 \right) \right) \right) - \frac{\delta}{\lambda} < 0 \Leftrightarrow \theta(N) = \frac{\beta \theta(N) - \delta}{\delta} < 0.
\]

This completes the proof. □

Intuitively, because a VSL system increases the total throughput of the bottleneck, it shortens the duration of the morning peak and makes commuters' arrivals more concentrated at their desired arrival time, thereby reducing the total schedule delay cost. However, once the capacity drop is fully avoided, i.e., \( N_c = N \), a VSL scheme with a lower \( N_c \) cannot further shorten the duration to reduce the total schedule cost.

5.2.2. Empirical analysis

As previously mentioned, active bottlenecks suffer approximately 10% capacity drop. Thus \( \lambda = 0.9 \). With \( \lambda \) varying in the range of \([2, 4]\), \( \frac{S}{\beta} \) falls in the range of \([\frac{2}{\beta}, \frac{4}{\beta}]\), which is always less than 0.9. Considering \( \lambda_1 = \frac{S}{\beta} \) and \( \lambda_2 = \frac{S}{\beta} \), when \( \frac{S}{\beta} \to 0 \), we thus conclude that the proposed VSL scheme is unlikely to reduce the total travel cost of morning commuters. This can be also observed in Fig. 6, where the solid line is always above Line 1 (representing \( \frac{S}{\beta} \)) when \( \lambda \) is between 2 and 4.

5.3. Total emission cost

5.3.1. Theoretical analysis

As shown in Section 5.1, a VSL scheme can yield a tradeoff between queueing delay and cruising time, and thus may reduce total emissions cost. Under a specific VSL scheme, the total emissions can also be expressed as a function of \( N_c \):

\[
\text{TE}(N_c) = \text{TT}_o(N_c) \cdot e_o + \text{TT}_q(N_c) \cdot e_q,
\]

where \( \text{TT}_o \) is the total cruising time determined by Eq. (32), and \( \text{TT}_q \) is the total queueing delay determined by Eq. (33). To facilitate our analysis, we then define two critical values:

\[
\rho_1 = \min\{\theta(N)\}^{-1}, \rho_2 = \frac{\beta}{\delta},
\]

where \( \theta(\cdot) \) is defined by Eq. (36), and \( \rho_o \) solves

\[
\text{TE}(N) = \text{TE}(N_c).
\]

for given \( \frac{S}{\beta} \) and \( \lambda \), and \( \rho < \rho_o \iff \text{TE}(N) \leq \text{TE}(N_c) \). It can be shown that \( \rho_o \) is uniquely determined. As assumed, \( N > N_c = \frac{S}{\beta} q_c \).

Note that \( \theta(\cdot) > 0 \), it can then be verified that \( \rho_1 < \rho_2 \). We are now ready to present the following proposition.

**Proposition 5.6.** To minimize total emissions cost, under the proposed VSL scheme: (i) If \( \rho \leq \rho_1 \), then \( N_c = N \); (ii) If \( \rho_1 < \rho < \rho_2 \), then \( 0 < N_c < N \); (iii) If \( \rho \geq \rho_2 \), then \( N_c = 0 \).

**Proof.** Differentiating Eq. (55) with respect to \( N_c \), we have:

\[
\frac{d\text{TE}(N_c)}{dN_c} = \begin{cases} 
\frac{6}{\beta} \cdot \left( \frac{S}{\beta} \cdot e_q - e_o \right) & \text{when } N_c < N_c \\
\frac{6}{\beta} \cdot \left( \theta(N) \cdot e_q - e_o \right) & \text{when } N_c > N_c
\end{cases}
\]

where \( \theta(\cdot) \) is defined by Eq. (36). It can be verified that:

\[
\rho < \rho_1 \Rightarrow \frac{d\text{TE}(N_c)}{dN_c} < 0 \text{ and } \text{TE}(N) \leq \text{TE}(N_c) \\
\rho_1 < \rho < \rho_2 \Rightarrow \frac{d\text{TE}(N_c)}{dN_c} < 0 \text{ for } N_c < N \text{ and } \frac{d\text{TE}(N_c)}{dN_c} > 0 \text{ for } N_c = N \\
\rho \geq \rho_2 \Rightarrow \frac{d\text{TE}(N_c)}{dN_c} < 0
\]

This completes the proof. □

5.3.2. Empirical analysis

From Tseng et al. (2005), we know \( \frac{S}{\beta} \in [2, 4] \). The capacity drop is approximately 10% and thus \( \lambda = 0.9 \). Considering \( \frac{S}{\beta} \to 0 \), we then have \( \rho_1 \in [1.08, 1.35] \), \( \rho_2 \in [1.2, 1.5] \). As shown in Section 3.2.4, \( \rho \) is in the range of 1.609–1.614. Comparing the conditions in Proposition 5.6, we have \( \rho > \rho_2 = 1.5 \). This implies the proposed VSL scheme can help reduce total emissions cost and minimize it by completely avoiding capacity drop.
5.4. Total social cost

5.4.1. Theoretical analysis

Lastly, we consider total social cost that includes all the cost components previously considered, i.e., travel time, schedule delay and emissions. It can be expressed as a function of \( N_c \):

\[
TSC(N_c) = TC(N_c) + TE(N_c),
\]

where \( TC(\cdot) \) and \( TE(\cdot) \) are the total travel cost and total emissions cost defined by Eqs. (43) and (55) respectively. Again, we define another critical value as follows:

\[
\lambda_4 = \left( \frac{2 \delta + \alpha \delta}{\beta} - \frac{\alpha \delta}{\beta} \right) \left( \frac{1}{N_T} \right) - \frac{1}{N_T}.
\]

We assume that \( \alpha > e_q \) and \( \frac{\alpha \delta}{\beta} < \frac{\alpha \delta}{\beta} \) to avoid tedious elaboration of unlikely scenarios such as \( \alpha \ll e_q \) (empirically, \( \frac{\alpha \delta}{\beta} \approx 0 \)). Under this setting, one can verify that \( 0 < \lambda_4 < 1 \). Also note when \( N_c = N \), total social cost \( TSC(N) \) depends on \( \lambda \). It can be shown that \( \frac{dTSC(N)}{dN} < 0 \) since \( \frac{dTSC(N)}{dN} < 0 \) and \( \frac{dTSC(N)}{dN} < 0 \). In addition, \( TSC(N) \rightarrow \infty \) if \( \lambda \rightarrow 0 \), while \( TSC(N) \rightarrow \left[ \alpha(e_q - e_o) + (1 + \frac{\alpha \delta}{\beta}) \right] N \) when \( \lambda \rightarrow 1 \). For \( q_c < \frac{\alpha}{\beta} N \) and \( \alpha > e_q \), by noting \( \left( \frac{\alpha}{\beta} N \right) \cdot e_q \), it can be verified from Eq. (36) that for \( TSC(N) > TSC(N_c) \), a unique percentage of remaining capacity, denoted by \( \lambda_\beta \), solves:

\[
TSC(N) = TSC(N_c).
\]

Note that \( \lambda < \lambda_\beta \iff TSC(N) > TSC(N_c) \) and vice versa. Below is a proposition on whether the proposed VSL scheme can reduce total social cost.

**Proposition 5.7.** An appropriate design of the proposed VSL scheme can help reduce total social cost if and only if

\[
\lambda < \max\{\lambda_4, \lambda_5\}.
\]

**Proof.** From Eq. (60), we have

\[
\frac{dTSC(N_c)}{dN_c} = \begin{cases} 
\frac{\alpha (\delta - \beta)}{\beta} + \frac{\frac{\alpha}{\beta} N_c}{\beta} \cdot (\frac{\alpha}{\beta} e_q - e_o) \quad \text{when } N_c \leq N_c \\
\frac{\alpha (\delta - \beta)}{\beta} + \frac{\frac{\alpha}{\beta} N_c}{\beta} \cdot (\delta e_q - e_o) \quad \text{when } N_c > N_c.
\end{cases}
\]

It can be verified from Eq. (36) that for \( N_c > N_c, \theta(N_c) > \frac{\alpha}{\beta} \). Also note that \( \theta(\cdot) > 0 \). Thus \( \frac{dTSC(N)}{dN_c} \) increases with \( N_c \).

If \( \lambda < \lambda_4, \frac{\alpha}{\beta} \delta - \beta > -\frac{\alpha}{\beta} (\theta(N_c) \cdot e_q - e_o) \), then \( \frac{dTSC(N_c)}{dN_c} \bigg|_{N_c=N_c} > 0 \). This implies that, when \( N_c = N \), a small decrease in \( N_c \) would lead to a decrease in total social cost, i.e., introducing VSL can reduce total social cost.

If \( \lambda > \lambda_4, -\frac{dTSC(N_c)}{dN_c} < 0 \). The VSL scheme can reduce total social cost only if \( TSC(N) > TSC(N_c) \), which requires \( \lambda < \lambda_5 \). Note that \( \lambda_5 \) can be either larger than or less than \( \lambda_4 \), depending on \( \frac{\alpha}{\beta} \).

Similar to Proposition 5.1, the conditions can be summarized as \( \lambda < \max\{\lambda_4, \lambda_5\} \). This completes the proof. □

It can be observed that, when \( \lambda \) is smaller, \( \frac{\alpha}{\beta} \) or \( \frac{\alpha}{\beta} \) is larger, the condition \( \frac{\alpha}{\beta} \delta - \beta > -\frac{\alpha}{\beta} (\theta(N_c) \cdot e_q - e_o) \) is more likely to hold. Generally, VSL is more likely to reduce total social cost when capacity drop is larger, and late arrival is more heavily penalized as compared with early arrival, and queuing is also more environmentally hazardous than queuing. We also note that when \( \frac{\alpha}{\beta} \delta \rightarrow 0, \lambda_4 \rightarrow \frac{\alpha}{\beta} e_q \) and \( \lambda_5 \rightarrow \frac{\alpha}{\beta} e_q \). Considering \( e_q > e_o \), we then have \( \lambda_4 > \lambda_5 \), indicating the VSL can reduce total social cost if and only if \( \lambda < \frac{\alpha}{\beta} e_q \). When \( \frac{\alpha}{\beta} \delta \rightarrow 0, \lambda_4 \rightarrow \alpha(e_q - e_o) - \frac{\alpha}{\beta} e_q \frac{\alpha}{\beta} \delta \) and \( \lambda_5 \rightarrow 1 \). We then have \( \lambda_5 > \lambda_4 \), indicating the VSL can reduce total social cost if and only if \( \lambda < \lambda_5 \). Similarly, below we present a stronger condition without requiring knowing \( q_c \) and \( N \).

**Proposition 5.8.** The proposed VSL scheme can reduce total social cost if

\[
\lambda < \alpha \left( \frac{\alpha + e_o - \frac{\alpha \delta}{\beta} e_q}{\beta} \right)^{-1} \frac{\alpha}{\beta}.
\]

**Proof.** As assumed, we have \( 0 < \frac{\alpha}{\beta} < \frac{\alpha}{\beta} \), then

\[
\alpha(e_q - e_o) - \frac{\alpha}{\beta} e_q \left( \frac{\alpha}{\beta} e_q \right)^{-1} \frac{\alpha}{\beta} \leq \lambda_5 \leq \frac{\alpha + e_o}{\alpha + e_o} \frac{\alpha}{\beta}.
\]

From Eqs. (65) and (66), it can be verified that

\[
\lambda < \lambda_5 \leq \max\{\lambda_4, \lambda_5\}.
\]
According to Proposition 5.7, the VSL scheme can reduce total social cost.

5.4.2. Empirical analysis

Previous empirical studies, e.g., Small (1982), Bates et al. (2001), Lam and Small (2001), Tseng and Verhoef (2008) and Devarasetty et al. (2012), have estimated the values of $a$, $b$ and $c$. For recent comprehensive reviews, see, e.g., Carrion and Levinson (2012). Here we consider the average values of $a$, $b$ and $c$ are 13.7, 6.4 and 20.0 US$ per hour (Tseng et al., 2005), and thus $d = 4.8$ US$ per hour. Similar as Section 3.2.4, we consider that $e_q$ and $e_o$ are 1.392 and 0.864 US$ per hour, with a discount rate of 2.5%. Consequently, the condition in Eq. (65) yields $\lambda < 0.768$, implying that if the capacity drops by more than 23.2%, the VSL scheme can help reduce total social cost. Considering $\lambda$ and total demand $N$ are not in need.

As can be seen in Fig. 8, $\max\{\lambda_1, \lambda_2\} < \lambda_3 < \max\{\lambda_4, \lambda_5\}$, implying that the VSL scheme is less likely to reduce total travel time (due to increased cruising time), followed by total travel cost and then total social cost. The scheme is more likely to reduce total social cost, because the emission cost can be reduced by eliminating queuing and capacity drop. Considering practically $\frac{e_q}{e_o} \to 0$, it can be observed that the VSL scheme can help reduce all the above three items if capacity drop is higher than 26.4%, i.e., $\lambda < 0.736$. It does not help if capacity drop is less than 21.5%, i.e., $\lambda \geq 0.785$. Empirically, capacity typically does not drop by such a large percentage.

6. Concluding remarks

This paper has examined the effectiveness of VSL to improve traffic flow efficiency and reduce vehicular emissions in a stylized setting of the morning commute, where commuters choose their departure times based on a combination of factors: the chance of running into bottleneck congestion, the possible schedule delay, and the cruising travel time regulated by VSL.

We firstly consider a VSL system that adjusts commuters’ cruising speeds in a continuous fashion and conclude that the system eliminate all queuing and avoid capacity drop. It reduces travel time, travel cost, emissions cost and thus the total social cost. Although the system helps improve all performance measures, it is not as effective as the first-best dynamic tolls.

For a more practical VSL system that changes the speed limit in a discrete fashion, we show that a well-designed VSL scheme can prevent or delay the activation of the bottleneck and thus reduce queuing delay. It can help reduce the schedule delay cost and total emissions cost. However, it is unlikely for such a VSL system to reduce individual travel cost and social

Please cite this article in press as: Liu, W., et al. Effectiveness of variable speed limits considering commuters’ long-term response. Transportation Research Part B (2014), http://dx.doi.org/10.1016/j.trb.2014.12.001
cost. More specifically, we show that the discrete VSL scheme considered in this paper always increases total cruising time as it slows down vehicles. As commuters’ arrival to the bottleneck is well controlled, it reduces the queuing delay. However, the reduction does not necessarily overweight the increase in the cruising time. Consequently, total travel time does necessarily decrease. We provide conditions when the discrete VSL scheme can reduce total travel time. Using available empirical data, we conclude that the discrete VSL system is unlikely to reduce total travel time even though capacity drop can be completely avoided. Similarly, we establish conditions for the discrete VSL system to reduce total travel cost that includes commuters’ schedule delay cost. It is found that if capacity drop is sufficiently large or late arrivals are heavily penalized, the discrete VSL system can reduce total travel cost. Unfortunately, empirical data suggest otherwise. As far as vehicular emissions are concerned, because queuing is much more environmentally hazardous than cruising, VSL likely helps reduce total emissions cost. When we consider total social cost that includes all the cost components previously considered, the discrete VSL system is more of help, if capacity drop is larger, and late arrival is more heavily penalized as compared with early arrival, and queuing yields comparably higher emissions cost than cruising. However, the proposed VSL system once again fails in the empirical test and we conclude that it unlikely reduces total social cost.

We conclude the paper by providing a few caveats regarding our analysis. Firstly, the social cost analysis in this paper does not consider safety benefit, which a VSL system can significantly contribute (e.g., Smulders, 1990). If the safety benefit is considered, it is more likely for even the discrete VSL systems to reduce total social cost. Secondly, this paper considers a recurrent bottleneck, while VSL systems can be of much help in non-recurrent bottlenecks caused by, e.g., incidents and work zones. Moreover, the mechanism of interest is for VSL to prevent or delay capacity drop. If capacity drop does not exist, the VSL systems considered in this paper will not offer any help with performance measures except the emissions cost. However, other mechanisms of VSL may work even if capacity drop does not exist. Therefore, our observations in this paper do not necessarily apply to all practical VSL systems, which may be designed upon other mechanisms and are applied to freeway networks with various types of bottlenecks and geometric configurations. Lastly and more importantly, we assume that commuters react to VSL systems and the system achieves a long-term equilibrium. Whether this behavioral assumption is plausible remains to be seen. In short, our analysis only suggests that there are conditions where even a well-designed VSL scheme does not improve traffic flow efficiency. Caution thus should be exercised to identify those where VSL can be of help.

Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive and helpful comments. This paper also benefits from discussions with Profs. Jorge Laval at Georgia Institute of Technology, Gang-Len Chang and Lei Zhang at University of Maryland, Michael Zhang at University of California, Davis and Yinhai Wang at University of Washington. The work described in this paper was partly supported by the U.S. National Science Foundation (CNS-1239364), Hong Kong’s Research Grants Council (HKUST620913) and National Natural Science Foundation of China (71371020). Yafeng Yin would also like to thank the support from Lloyd’s Register Foundation (LRF). LRF helps to protect life and property by supporting engineering-related education, public engagement and the application of research.

Appendix A. Proof of uniqueness of equilibrium pattern in Fig. 2(b)

Proof. In Fig. 2(b), the earliest and latest arrival times at the bottleneck are \( t'_s + t_f \) and \( t'_e \) respectively, and the corresponding departure times from home are \( t'_s \) and \( t'_e - t_f \). Since the departure curve from home and the arrival curve at bottleneck are parallel, without loss of generality we herein consider \( t_f = 0 \).

At equilibrium, there cannot be unused capacity during interval \([t'_s, t'_e]\), and there cannot be queue at the time of the last departure and that the travel costs of the first and last commuters to depart are the same. Then one can verify that the departure (from home) interval starting at \( t'_s \) and ending at \( t'_e \) is uniquely determined by

\[
\frac{\nu}{s} (\nu t'_e - t'_s) = \gamma (\nu t'_s - t'_e); \quad \nu t'_e - t'_s = \frac{\nu q_e}{s} + \frac{N - \frac{\nu}{s} q_e}{s'},
\]

The time when capacity drops from \( s \) to \( s' \) is then uniquely determined since \( t_d = t'_s + \frac{\nu}{s} q_e / s \).

Given that the above critical time points are unique, similar to the case without capacity drop, the queue length and departure rate profiles over the time horizon ensuring the constant individual travel cost can be uniquely determined accordingly (due to the fact that the schedule delay cost over time is continuous and convex). This completes the proof. \( \square \)

Appendix B. Proof of FIFO in Fig. 3(a)

Proof. In Fig. 3(a), note that the earliest and latest commuter experience the same cruising time \( t_r \) with the same speed \( \nu_r \), where \( l = t_f \cdot t_r \). For early-arrival commuters, it is evident that FIFO is satisfied since the speed limit gradually decreases. Now we focus on late-arrival commuters. For ease of presentation, we define that

\[
\Delta t' = \frac{1}{s} - \frac{1}{t_1}; \quad \Delta t'' = \frac{1}{s} - \frac{1}{t_2}.
\]

Please cite this article in press as: Liu, W., et al. Effectiveness of variable speed limits considering commuters’ long-term response. Transportation Research Part B (2014), http://dx.doi.org/10.1016/j.trb.2014.12.001
The cruising time experienced by the commuter arriving at destination on time is
\[ t_f = t_f + \frac{\delta}{\beta} N \Delta t'' = t_f + \frac{\delta}{\gamma} N \Delta t'' = t_f + \frac{\delta}{\alpha} \frac{N}{s}. \] (70)

The corresponding speed limit is
\[ \bar{v}_{\min} = \frac{l}{t_f}. \] (71)

Arranging late-arrival commuters in the order of their departure times, we can denote the x-th commuter’s cruising time, speed limit and departure time by \( t_f(x), \bar{v}(x) \) and \( t(x) \) respectively. Under the above setting, we have
\[ t_f(0) = t_f, \bar{v}(0) = \bar{v}_{\min}. \] (72)

The cruising time of the x-th commuter is
\[ t_f(x) = t_f - x \Delta t'', \] (73)
and the speed limit imposed on him or her is
\[ \bar{v}(x) = \frac{l}{t_f - x \Delta t''}. \] (74)

and his or her departure time is
\[ t(x) = t(0) + \frac{x}{t_2}. \] (75)

where \( x < \frac{l}{N} \). Without loss of generality, denote the entrance of the freeway as the origin of the location coordinate, and then at time \( t \), the x-th commuter will arrive at the location
\[ h(x, t) = (t - t(x)) \cdot \bar{v}(x), \] (76)

where \( t(x) \leq t \leq t(x) + t_f(x) \). Note that, \( h(x, t) \) is shown in Fig. 1, which satisfies \( h(x, t) \leq l \). Differentiating Eq. (76) with respect to \( x \) yields
\[ \frac{\partial h(x, t)}{\partial x} = \bar{v}(x) \left( \frac{1}{r_2} + (t - t(x)) \frac{\Delta t''}{t_f(x)} \right). \] (77)

For the x-th commuter, before he or she arrives, i.e., \( t < t(x) + t_f(x) \), we have,
\[ \frac{\partial h(x, t)}{\partial x} < \bar{v}(x) \left( \frac{1}{r_2} + t_f(x) \frac{\Delta t''}{t_f(x)} \right) = \bar{v}(x) \left( \frac{1}{r_2} + \frac{1}{r_2} - \frac{1}{s} \right) < 0. \] (78)

This implies that a commuter who departs later will not overtake an earlier-departure one. □

References


