**Tradeoffs among Free-flow Speed, Capacity, Cost, and Environmental Footprint in Highway Design**

Chen Feng Ng and Kenneth A. Small

August 28, 2008

**Abstract**

This paper investigates differentiated design standards as a source of capacity additions that are more affordable and have smaller aesthetic and environmental impacts than expressways. We consider several tradeoffs, including narrow versus wide lanes and shoulders on an expressway of a given total width, and high-speed expressway versus lower-speed arterial. We quantify the situations in which off-peak traffic is sufficiently great to make it worthwhile to spend more on construction, or to give up some capacity, in order to provide very high off-peak speeds even if peak speeds are limited by congestion. We also consider the implications of differing accident rates. The results support expanding the range of highway designs that are considered when adding capacity to ameliorate urban road congestion.

Keywords:
highway design, capacity, free-flow speed, parkway

*Journal of Economic Literature* codes:
L91, R42

Contacts:
Chen Feng Ng: Department of Economics, California State University at Long Beach; cng3@csulb.edu
Kenneth Small: Department of Economics, University of California at Irvine; ksmall@uci.edu

An earlier version was presented to Third Kuhmo-Nectar Conference, Amsterdam, July 2008.

We are grateful to the University of California Transportation Center for financial support, and to many colleagues for comments on earlier drafts, especially Pablo Durango-Cohen, Robin Lindsey, Robert Noland, Robert Poole, Peter Samuel, Ian Savage, and Erik Verhoef. Of course, all responsibility for facts and opinions expressed in the paper lies with the authors.

## 1. Introduction

Many analysts and policy makers have argued that building more highways is an ineffective response to congestion: specifically, that it is infeasible to add enough highway capacity in large urban areas to provide much relief. The argument is supported, for example, by examining the funding requirements for a highway system that is estimated to accommodate travel at acceptable levels of service. Such requirements come to hundreds of billions of dollars for the entire US (US FHWA 2006b). This observation, along with impediments to various other measures, has led Downs (2004) to conclude that there is little chance for a resolution to congestion problems in the US.

Much of the expense envisioned in such lists of needs is for new or expanded expressway routes.[1] Expressways are very intensive in land and structures, requiring great expense and disruption to existing land uses. Furthermore, the resulting road has significant environmental spillovers on surrounding neighborhoods. These impacts, including air pollution, noise, and visual impact, are closely related to the size of the road and even, in the case of noise and nitrogen-oxide emissions, directly to speed itself. Thus there are significant tradeoffs that need to be considered when deciding what kind of free-flow speed should be provided by a given highway investment.

The highest-capacity roads built in the US are usually expressways, and they generally conform to very high design standards. The most rigorous of these standards are those of the federal Interstate system, which specify lane width, sight distance, grade, shoulders, and other characteristics (AASHTO 2005). These standards are mainly dictated by two underlying assumptions: the road must be safe for travel at high speeds (typically 55-70 mi/h in urban settings), and it must be able to carry mixed traffic including large trucks.

But does it make sense to build high-speed roads that will be heavily congested for large parts of the day, so that only a minority of vehicles experience those high speeds? And does it make sense, in an area served by a grid of expressways, to design every major route to accommodate large trucks?

[1] See US FHWA (2006b), ch 7. Of the $84.5 billion in investments in urban arterial and collector roads meeting defined cost-benefit criteria in this report, 53 percent is for freeways and expressways – of which about half is for expansion, half for rehabilitation or environmental enhancement (Exhibit 7-3).

One way to look at the problem is in terms of equilibration of travel times. In a heavily congested urban area, higher-quality roads become congested more severely than others, so that levels of service tend to be equalized (Pigou 1920, Downs 1962). When this occurs, the extra expense incurred to raise the design speeds on major roads has no payoff during congested periods, whereas anything to improve capacity has a huge payoff. Clearly, then, one key factor determining the ideal highway design will be the ratio of peak to off-peak traffic.

We can illustrate the tradeoffs by considering lane width. The standard 12-foot-wide lanes of US interstate highways provide safety margins for mixed traffic at high speeds, often under difficult conditions of weather and terrain. On most urban commuting corridors, there are fewer trucks and speeds are low during much of the day; thus the need for such safety margins is smaller. Indeed, urban expressway expansions are sometimes carried out by converting shoulders to travel lanes and restriping all lanes to an 11-foot width. These have a disadvantage of slower free-flow travel, and perhaps of higher accident rates – although as we shall see the evidence on safety is mixed. The point here is that by squeezing lanes and shoulders, more capacity can be obtained at the expense of some other desirable features. Hence, there is a tradeoff.

In this paper, we examine just a few of the tradeoffs involved by considering examples of pairwise comparisons between two urban highway designs, in which as many factors as possible are held constant. We first consider different lane and shoulder widths for a given highway type (expressway or signalized arterial). In the case of the expressway, this really amounts to a reconsideration of the "parkway" design that prevailed in the US prior to 1950, except we do not attempt to account quantitatively for the truck restrictions, tighter curves, or nicer landscaping that may further enhance this option. We then compare expressways with high-performance unsignalized arterials. In each comparison, we characterize the range of conditions under which the more modest design (narrower lanes, lower design speed) provides for greater travel-time savings or involves the least total cost including user costs.

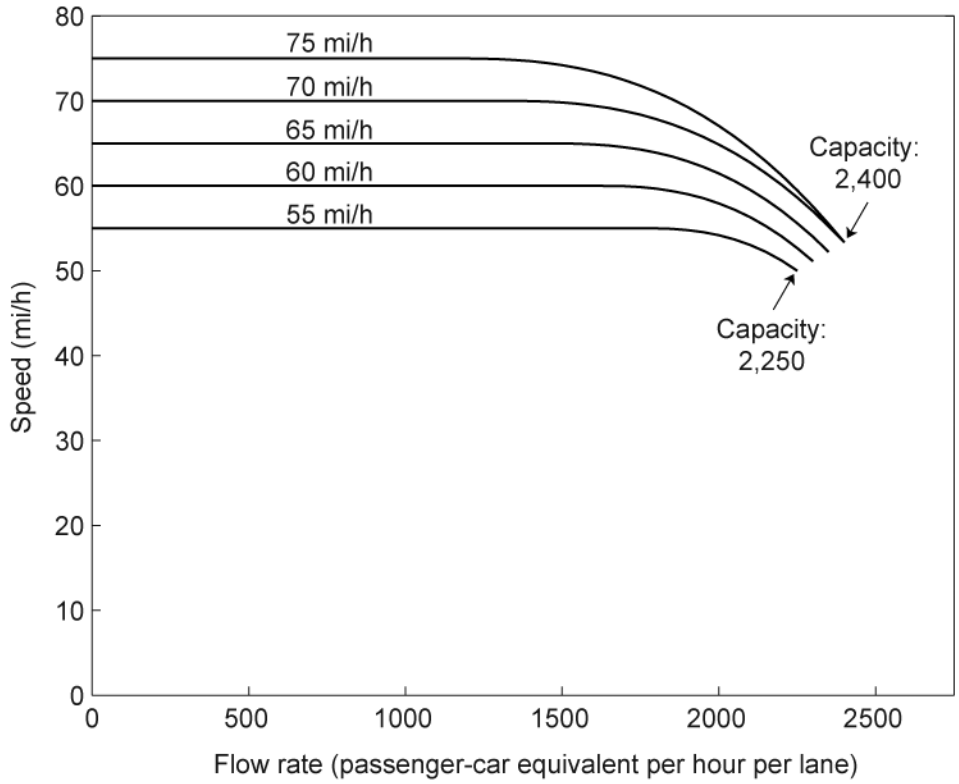
## 2. Congestion Formation, Capacity, and Travel Time

We consider four determinants of travel time on a highway. First is free-flow speed, which for expressways is specified as a function of highway design including lane and shoulder widths. Second is congestion delay due to the inflow of traffic at rates less than the highway's capacity; this is described by speed-flow curves. Third is further congestion delay due to queuing when inflow exceeds capacity; this is described by a simple bottleneck queuing model. Fourth is control delay, which applies only to arterials with signalized intersections. In this section, we derive peak and off-peak travel times from basic relationships involving these four determinants. In Section 3, we relate the parameters of these relationships numerically to highway design parameters, and carry out the comparisons described earlier.

The first two determinants just described are summarized by a speed-flow function, $S(V)$, giving speed as a function of inflow. We define it for $0 \leq V \leq V_K$, where $V_K$ is the highway's capacity defined as the highest sustainable steady flow rate.[2] For expressways, speed decreases as the flow rate approaches the expressway's capacity. Figure 1 shows the speed-flow curves for various free-flow speeds for expressways based on the 2000 edition of the Highway Capacity Manual (Transportation Research Board 2000). The free-flow speed is $S^0 \equiv S(0)$. For signalized urban arterials, the speed-flow curve is flat, i.e., the speed on the urban street remains at the free-flow speed for all flow rates up to the road's capacity $V_K$. This capacity is determined by the saturation flow rate at intersections (the maximum flow rate while the signal is green) multiplied by the proportion of the signal's cycle time during which it is green.

[2] Cassidy and Bertini (1999) suggest that the highest observed flow, which is larger, is not a suitable definition of capacity because it generally breaks down within a few minutes – although Cassidy and Rudjanakanoknad (2005) hold out some hope that this might eventually be overcome through sophisticated ramp metering strategies. Our speed-flow function does not include the backward-bending region, known as congested flow in the engineering literature and as hypercongested flow in the economics literature, because flow in that region leads to queuing which we incorporate separately. See Small and Verhoef (sect 3.3.1, 3.4.1) for further discussion of hypercongestion.

**Figure 1. Speed-flow curves for different free-flow speeds for expressways**




The third determinant, queuing delay, may be approximated by deterministic queuing of zero length behind a bottleneck (Small and Verhoef 2007, sect 3.3.3). For the sake of concreteness, we assume the bottleneck occurs at the entry to the section of road under consideration.[3] Suppose traffic wishing to enter the road arrives at rate $V_o$ during an off-peak period of total duration $F$, and at rate $V_p$ during a peak period of duration $P$ and starting at time $t_p$. We describe here the case $V_o$<$V_K$<$V_p$ so that a queue forms during the peak period and vehicles leave the queue at the rate $V_K$. We also assume $F$ is long enough that the queue disappears by the end of the off-peak period. The number of vehicles in the queue, $N(t)$, builds up at rate $V_p$-$V_K$ starting at time $t_p$, causing queuing delay $D(t)$=$N(t)/V_K$ to a vehicle entering at time $t$. At time $t_{p'} = t_p+P$, the end of the peak period, this delay has reached its maximum value,

[3] If the road is not of uniform capacity, the bottleneck would occur at the point of lowest capacity. This would require working out exactly how much flow occurs in each subsection in order to determine the non-queuing congestion delays, with only a small effect on the total travel times.

$D_{max}$=$P\cdot[(V_p/V_K)$-1]. The queue then begins to dissipate, shortening at rate ($V_K$-$V_o$) until it disappears at time $t_x=t_{p'}+(V_p-V_K)P/(V_K-V_o)$.

The resulting queuing delay has the triangular pattern shown in Figure 2. The average queuing delay to anyone entering during the peak period is

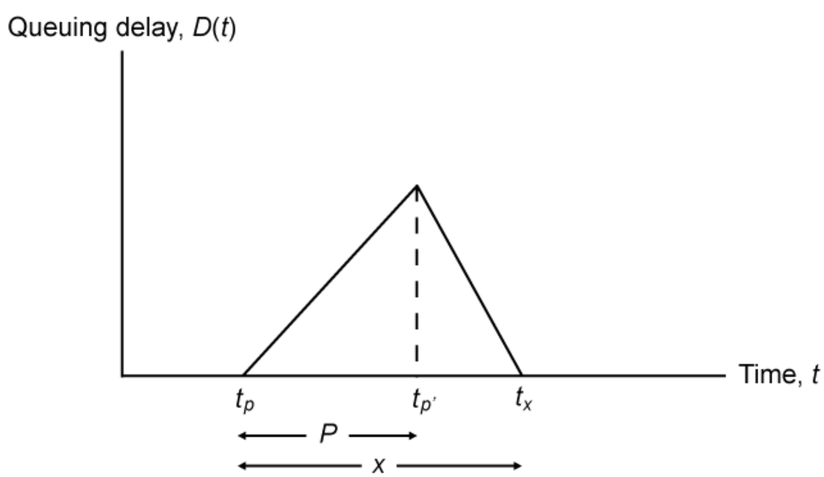
$$\overline{D}_p = \tfrac{1}{2} D_{\max} = \tfrac{1}{2} P \cdot [(V_p / V_K) - 1]\,. \tag{1}$$

The same average queuing delay affects those arriving between times $t_{p'}$ and $t_x$, which when averaged with the other off-peak travelers (who experience no queuing delay) produces the average off-peak travel delay:

$$\overline{D}_o = \frac{\overline{D}_p \cdot (x - P)}{F} = \frac{1}{2} \frac{P^2}{F} \cdot \frac{(V_p - V_K)^2}{V_K \cdot (V_K - V_o)} \tag{2}$$

where $x=t_x-t_p$ is the total duration of queuing (hence $x-P=t_x-t_{p'}$).

**Figure 2. Queuing delay for vehicles arriving at road entrance at time *t***




In addition to queuing delay, urban arterial drivers face "control delay" reflecting time lost slowing and waiting for signalized intersections — aside from the bottleneck queuing just discussed. Let $Z$ denote the number of signalized intersections on the urban street encountered on a trip of length $L$. Based on the HCM's procedure (see Appendix A), the average control delay per vehicle for through movements, $\delta$, is:

$$\delta = Z \cdot \frac{0.5C \cdot (1 - g / C)^2}{1 - \left[(V / V_K)(g / C)\right]} \tag{3a}$$

where $C$ is signal cycle length, $g$ is effective green time, and $V$ is the volume of traffic going through the intersection. Time durations $C$, $g$, and hence $\delta$ are all conventionally measured in seconds. Equation (3a) assumes that vehicles arrive at a signal at a constant rate $V$, which results in queuing if the signal is red, and it incorporates the way this queue dissipates when the signal turns green.[4] That dissipation depends on lane-specific saturation rates, which we are able to relate to the overall capacity $V_K$ of the highway (see Appendix A). Recall that due to our assumption that this same capacity limits upstream flow, $V$ is equal to the queue discharge rate $V_K$ during the time period $t_p$ to $t_x$ (i.e., when queuing occurs at the bottleneck). At all other times, $V$ is equal to the off-peak volume, $V_o$. Thus, for peak travelers and for those off-peak travelers who experience queuing at the bottleneck, the control delay for each vehicle is:

$$\delta_p = Z \cdot 0.5C \cdot (1 - g/C). \tag{3b}$$

For all other off-peak travelers, the control delay is given by (3a) with $V$=$V_o$:

$$\delta_o = Z \cdot \frac{0.5C \cdot (1-g/C)^2}{1 - \left[(V_o / V_K)(g/C)\right]} \tag{3c}$$

We now combine all four sources of delay and add them over vehicles. Peak travelers, of whom there are $V_pP$, experience speed $S(V_K)$ while moving. Adding control delay and queuing delay yields total travel time in hours:

$$TT_p = V_p P \cdot \left[ \frac{L}{S(V_K)} + \frac{\delta_p}{3600} + \overline{D}_p \right]. \tag{4}$$

Among off-peak travelers, a fraction $x$-$P$ travel at speed $S(V_K)$ while moving, whereas a fraction $F$-$x$+$P$ travel at speed $S(V_o)$. Again adding control and queuing delay, their total travel time is:

$$TT_o = V_o \cdot \left[ (x-P)\left( \frac{L}{S(V_K)} + \frac{\delta_p}{3600} \right) + (F - x + P)\left( \frac{L}{S(V_o)} + \frac{\delta_o}{3600} \right) + F\overline{D}_o \right]. \tag{5}$$

Total travel time and average travel time are, respectively:

$$TT_{all} = TT_P + TT_o \tag{6}$$

$$\overline{TT}_{all} = \frac{TT_{all}}{PV_p + FV_o} \tag{7}$$

[4] The delay in equation (3a) is also known as uniform delay, as described in Appendix A. It is based on Webster's delay formulation; see Rouphail *et al* (1996) for a review of this and other methods for estimating control delay.

## 3. Comparisons between Designs with Equal Construction Cost

In this section, we make two comparisons of roads with "regular" and "narrow" designs, one for expressways and one for signalized urban arterials (which we shall refer to interchangeably as urban streets). In each case, we hold constant the total width of the roadway so there is very little cost difference between the two roads in each comparison. We also hold fixed the distance between the two halves of the road, so we need only consider one half, carrying traffic in one direction.

We ignore the difference in cost due to converting part of the paved shoulders in the "regular" design to vehicle-carrying pavements in the "narrow" design; since the largest component of new construction cost is grading and structures, this difference should be minor. We also ignore any differences in maintenance cost that may occur because vehicles on narrow lanes are more likely to veer onto the shoulder or put weight on the edge of the pavement (AASHTO 2004, p. 311).

This comparison enables us to focus on the two primary factors that distinguish these designs from each other: travel time and safety. As we will see, the safety advantages or disadvantages of "narrow" versus "regular" design are not entirely clear, in part because lower speeds help compensate for narrower lanes and less margin for error at the shoulder. We therefore assume in this section that the two designs have identical accident rates, and return to the safety issue in Section 5. This enables us to focus solely on travel time.

The 2000 Highway Capacity Manual (henceforth HCM) provides methodologies for determining road capacities, free-flow speeds, and indeed the entire speed-flow functions for expressways (which are called "freeways" by the HCM) and urban streets with different specifications. As described in detail in Appendix A, we use this information to determine the values $V_K$, $S(V_K)$, and $S(V_o)$ appearing in equations (3)–(5).

Our first example is expressways. Expressway R (the "regular" design) has two 12 ft lanes in one direction, a 6 ft left shoulder, and a 10 ft right shoulder, bringing its total one-directional roadway to 40 feet (see Figure 3a). These are the minimum widths recommended for "urban freeways" by AASHTO (2004) except we have added two feet to the left shoulder. Expressway N (the "narrow" design) has three 10 ft lanes, a 2 ft left shoulder, and an 8 ft right shoulder. As shown in Table 1, this road's narrower lanes and shoulders lead to a lower free-flow speed and thus a lower capacity per lane compared to Expressway R; but its total capacity ($V_K$) is higher since it has more lanes.

For signalized urban arterials, we compare two high-type urban arterial streets, each with the same number of signalized intersections and the same one-directional road width (38 ft). Following the lane and median width recommendations by AAHSTO (2004), the "regular" urban street (Urban Street R) has two 12 ft lanes in one direction for through movement, a 6 ft left shoulder (which provides for a 12 ft median in terms of a two-directional road) and an 8 ft right shoulder (see Figure 3b). At signalized intersections, the entire median (consisting of the left shoulders of *both* directional roadways) is used for a 12 ft exclusive left-turn lane (which therefore occupies the same linear space as the left-turn lane facing it in the opposite direction). The rightmost through lane is a shared right-turn lane, and the right shoulder width remains at 8 ft. We assign this urban street a speed limit of 55 mi/h.

The "narrow" urban street (Urban Street N) has three 10 ft lanes in one direction for through movement, a 2 ft left shoulder, and a 6 ft right shoulder.[5] At signalized intersections, the right shoulder width is reduced to 3 ft; the additional roadway plus the median are used to provide for an exclusive left-turn lane of 10 ft while maintaining three 10 ft through lanes, one of which is a shared right-turn lane.[6] We give it a speed limit of 45 mi/h.

These assumptions enable us to derive free-flow speeds and intersection delays by following procedures in the HCM, as detailed in Appendix A. Table 1 shows selected results. The free-flow time advantage for a trip of $L$=10 miles is 0.77 minutes for the "regular" compared to the "narrow" expressway; and it is 1.17 minutes for the "regular" compared to the "narrow" arterial. Recall that each pairwise comparison is of two roads occupying the same width and hence with nearly identical construction costs.

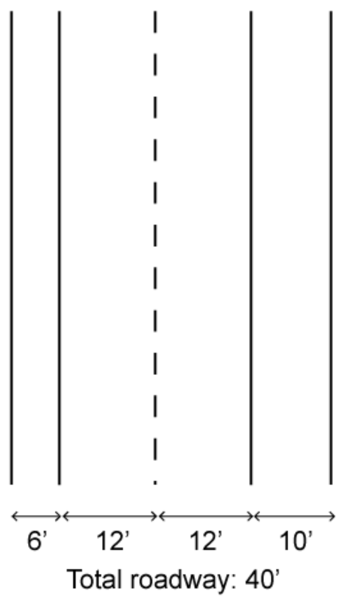
---

[5] A wider right shoulder is given priority over the left shoulder in order to provide adequate space for stopped vehicles and minimize the impact of such incidents on other vehicles.

[6] Note that for the urban streets in Figure 3b, the total two-directional roadway width at the intersection itself is less than the sum of those of the two separate one-directional roadways, because the left turn lanes in both directions share the same linear space. That is, the width of the two directional roadways includes only the width of one, not two, left turn lanes. For the "regular" design this is 2x(12+12+8)+12 = 76 = 2x38, whereas for the "narrow" design it is 2x(10+10+10+3)+10 = 76 = 2x38; hence both are described as having a 38-foot one-directional roadway.

**Figure 3a. Example expressways (one direction)**

(i) “Regular” expressway




(ii) “Narrow” expressway

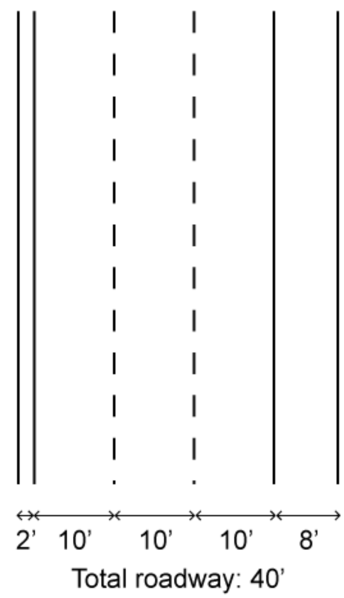
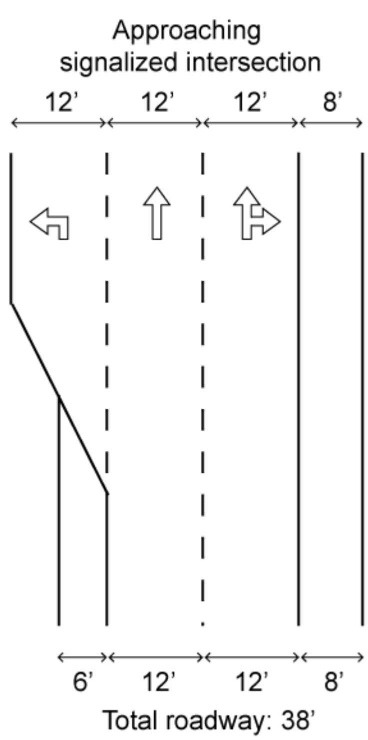



**Figure 3b. Example urban streets (one direction)**

(i) “Regular” urban street




(ii) “Narrow” urban street

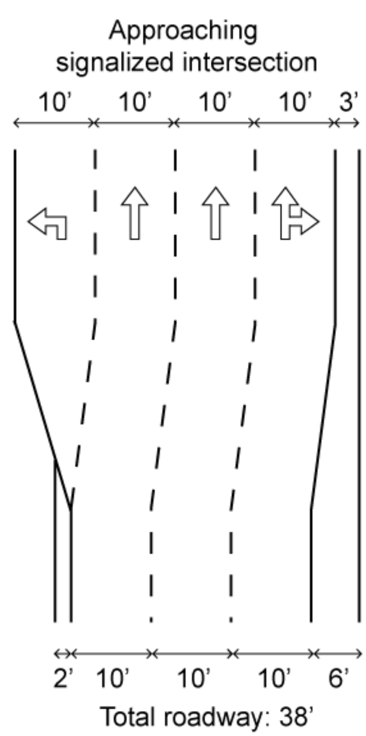


**Table 1. Specifications for examples (all one direction)**

| | Freeway R: regular lanes & shoulders | Freeway N: narrow lanes & shoulders | Urban St R: regular lanes & shoulders | Urban St N: narrow lanes & shoulders |
|---|---|---|---|---|
| ***Parameters*** | | | | |
| Number of lanes | 2 | 3 | 2 | 3 |
| Lane width (ft) | 12 | 10 | 12 | 10 |
| Left shoulder width (ft) | 6 | 2 | 6 | 2 |
| Right shoulder width (ft) | 10 | 8 | 8 | 6 |
| Total roadway (ft) | 40 | 40 | 38 | 38 |
| | | | | |
| Length (mi) | 10 | 10 | 10 | 10 |
| Percentage of heavy vehicles[1] | 0.05 | 0.05 | 0.05 | 0.05 |
| Driver population factor[1] | 1.00 | 1.00 | 1.00 | 1.00 |
| Peak hour factor[1] | 0.92 | 0.92 | 0.92 | 0.92 |
| Interchanges/signals per mile[1] | 0.50 | 0.50 | 0.50 | 0.50 |
| Signal cycle length (s)[1] | - | - | 100 | 100 |
| Effective green time (s)[1] | - | - | 70 | 70 |
| | | | | |
| ***Speed and capacity*** | | | | |
| Free-flow speed (mi/h)[2] | 65.5 | 60.4 | 51.5 | 46.8 |
| Speed at capacity (mi/h)[2] | 52.0 | 51.2 | 51.5 | 46.8 |
| Capacity per lane (veh/h/ln)[3] | 2,113.76 | 2,067.98 | 1,165.33 | 1,087.64 |
| Total capacity, $V_K$ (veh/h)[3] | 4,227.51 | 6,203.94 | 2,490.96 | 3,486.86 |
| | | | | |
| ***Travel time*** | | | | |
| Free-flow travel time (min) | 9.16 | 9.93 | 12.03 | 13.20 |

Notes:
[1] In most cases, the default values recommended by the HCM are used; see Appendix A. The recommended default value for interchanges per mile is used for expressways, and we assume a comparable number for signal density on urban streets. The signal cycle length is based on the HCM's default value for non-CBD areas (see Exhibit 10-16 of the HCM). A relatively high effective green time is chosen.
[2] For expressways, the HCM calculates the average passenger-car speed based on total flow rate using passenger-car equivalents (pces) for heavy vehicles (a car has 1 pce). Our calculations are based on the average speed of passenger cars.
[3] For urban streets, "capacity per lane" is based on the capacity of lanes which allow only through movement. See Appendix A for how total capacity is calculated for both expressways and urban streets.

Using the specifications listed in Table 1 and assuming duration of peak and off-peak periods $P$=4 hours and $F$=12 hours, average travel times (which include queuing delay and control delay, if applicable) can be calculated for a range of traffic volumes. Figures 4a-b show the average travel times for the four different road designs under different values for average daily traffic (ADT) and the ratio of peak volume ($V_p$) to off-peak volume ($V_o$). We see from

Figure 4a that when $V_p/V_o = 1.5$, the "regular" freeway experiences queuing when ADT exceeds 50,724 veh/h, but queuing does not occur on the "narrow" freeway for ADT values up to 60,000 veh/h because the latter has a higher capacity. Once queuing begins, the increase in average travel time is so marked that the average travel time on the "regular" freeway begins to exceed that of the "narrow" freeway when ADT is just a little higher than the value at which queuing begins. Some of this travel-time increase can be attributed to the lower speed when the lanes become more crowded, but most of it is due to queuing delay. In the case of the signalized urban arterials, the "regular" and "narrow" urban streets experience queuing when ADT exceeds 29,891 veh/h and 41,842 veh/h, respectively. The average travel time on the "regular" urban street starts to exceed that of the "narrow" urban street when ADT is greater than 30,536 veh/h.[7]

Figure 4b shows the average travel times for the four highway types when $V_p/V_o = 4$. With much higher traffic volumes during the peak hour compared to the previous scenario, queuing now begins at lower values of ADT. Once queuing begins, average travel time on the "narrow" design increases at a lower rate compared to the "regular" design because the former has more capacity and thus discharges vehicles from the queue at a higher rate.

Thus even though the "regular" roads have slightly shorter average travel times (compared to the "narrow" roads) when traffic volumes are low, this advantage is quickly erased when they experience queuing — all the more so when $V_p/V_o$ is large, since then more vehicles experience queuing, the duration of the queue is longer, and fewer vehicles reap the advantages of higher free-flow speed.

We can also calculate the values of ADT and $V_p/V_o$ for which the difference in average travel time between the "regular" and "narrow" designs is zero. Figures 5a-b show this (and other) contour lines for freeways and urban streets, respectively — plotted so that a positive number favors the "narrow" design. In both figures, the "narrow" design has shorter average travel times compared to the "regular" design in the region to the right of the "0" contour line. For the example freeways, the lowest value for the difference in average travel time (i.e. the largest possible advantage for the regular design) is -0.77 minutes, which occurs under free-flow conditions for both freeways. For the urban streets, the lowest difference is -1.17 minutes.

[7] In Figures 4a and b, travel times for the "regular" and "narrow" urban streets are shown for ADT values up to 39,855 veh/h and 55,789 veh/h respectively because beyond that, the duration of the queue ($x$) exceeds the analysis period ($F + P$).

**Figure 4a. Average travel times for $V_p/V_o = 1.5$**

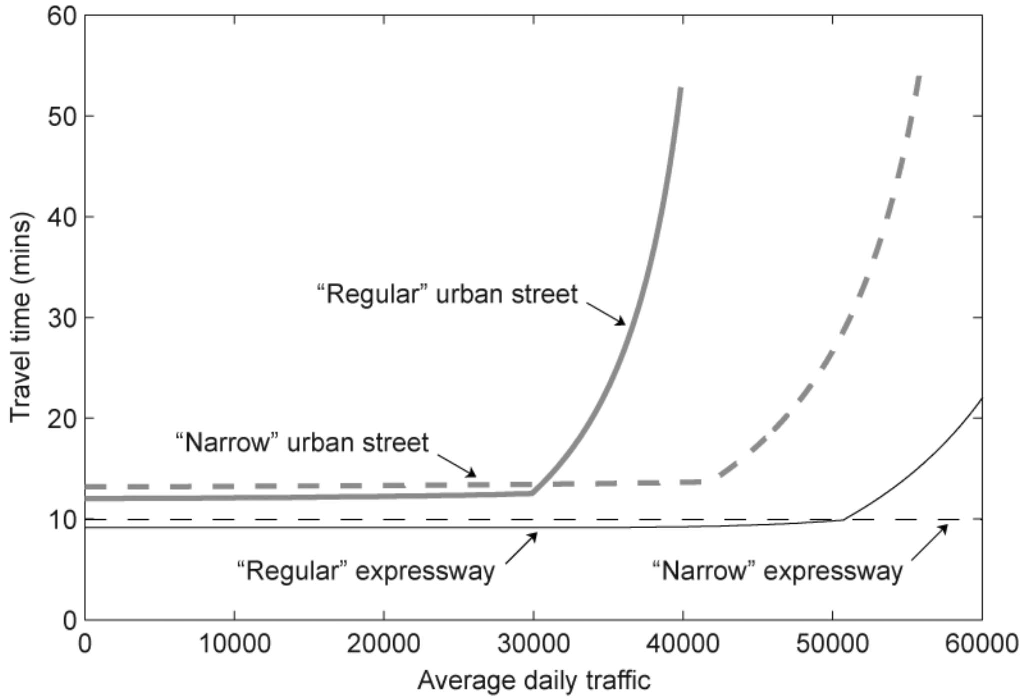



**Figure 4b. Average travel times for $V_p/V_o = 4$**

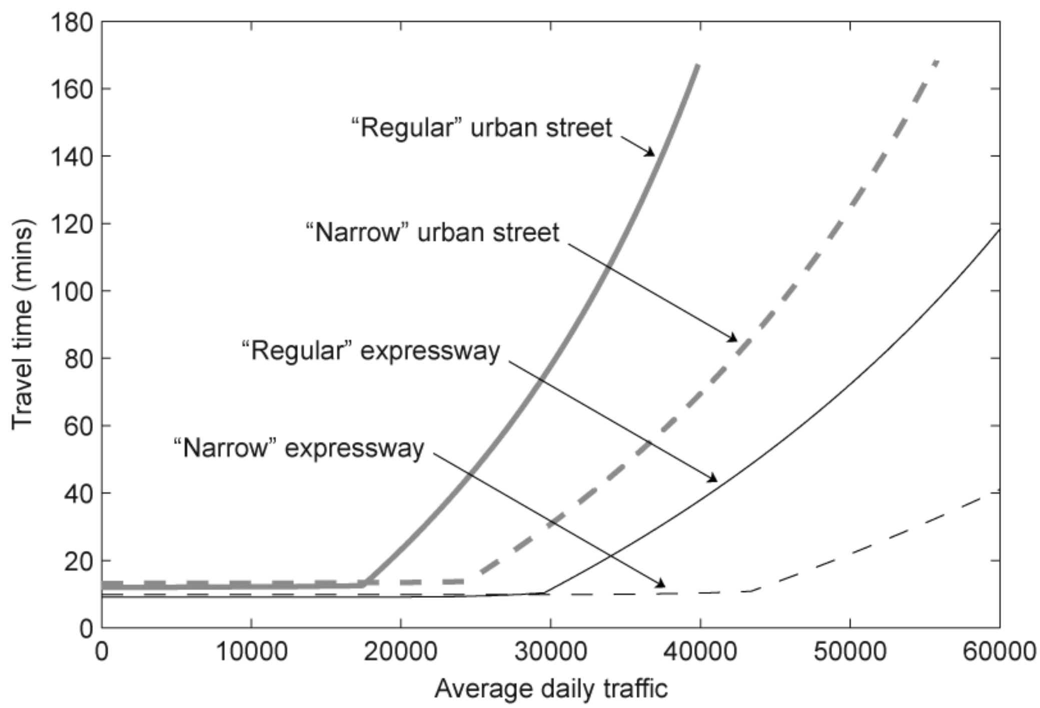


**Figure 5a. Contour map of the difference in average travel time (“regular” expressway minus “narrow” expressway)**

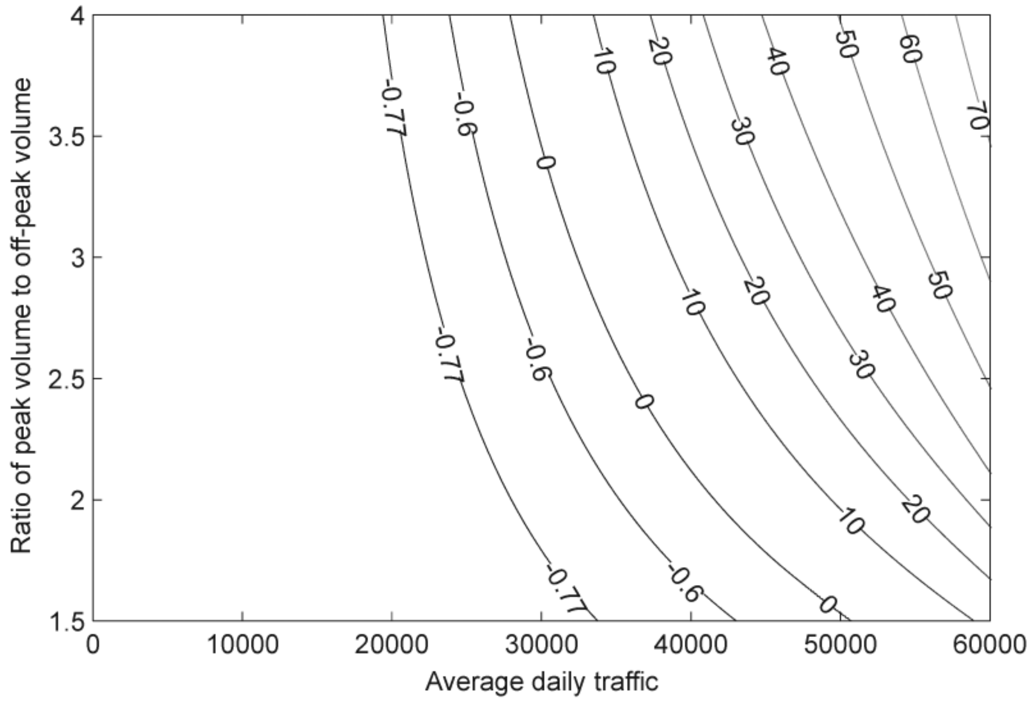
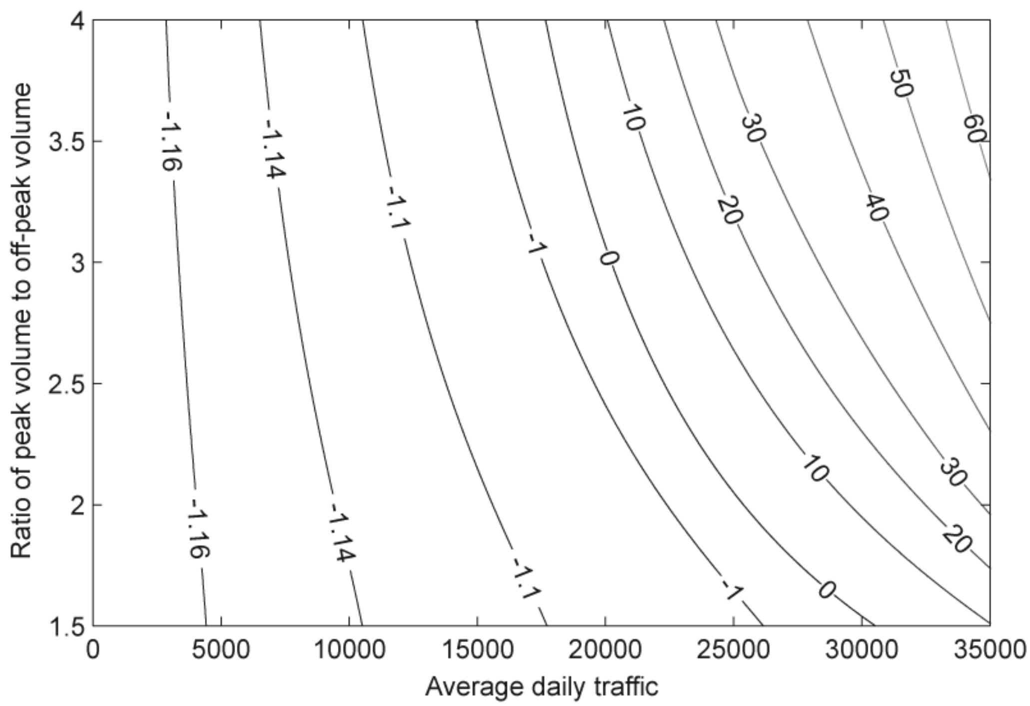



**Figure 5b. Contour map of the difference in average travel time (“regular” urban street minus “narrow” urban street)**



We observe that the "narrow" design is strongly favored under all conditions in which there is appreciable queuing. Most strikingly, the advantage of the "narrow" design increases extremely rapidly with traffic. By contrast, the advantage of the "regular" design for light traffic volumes is very modest and increases very slowly as traffic decreases. This is because the "narrow" design's advantage depends on queuing, whereas the "regular" design's advantage depends on the difference in free-flow speeds, which is quite small. While the specific numbers depend on our particular examples, these broad features result from well-established properties of highway design, and so are quite general.

## 4. Comparison between Expressways and Arterials with Equal Capacities

As seen in the previous section, signalized urban arterials tend have to lower free-flow speeds and capacities than expressways. However, arterials also tend to have lower construction costs. Thus it is useful to consider when it can be more cost-effective to build a lower-speed, less expensive arterial instead of an expressway.

In order to make realistic comparisons, we now consider a very high type of arterial, considerably higher than those of Section 3: namely, one that is divided, is uninterrupted by traffic signals, and has driveway or side-street access no more than once every two miles. This type of road is a type of "multilane highway" and differs from an expressway by allowing some access by other than entrance and exit ramps; but it has grade-separated intersections for all major crossings.[8]

We wish to examine total costs, including construction and travel-time costs, of a network of expressways versus a network of unsignalized arterials, each with the same capacity. We therefore compare the "regular" expressway from the previous section with an unsignalized arterial with similar characteristics.[9] Using the procedures outlined in Chapters 12 and 21 of the HCM, this unsignalized arterial has a capacity of 3,945 veh/h — only seven percent less than that

---

[8] According to the HCM (ch. 12-13), the most notable difference between a freeway and a multilane highway is that the former is characterized by full control of access. A "multilane highway" can have at-grade intersections and traffic signals with average spacing more than two miles.

[9] We also considered a cost comparison between the "regular" expressway and the "regular" signalized urban street in the previous section. This comparison always favored the expressway because the urban street's relatively low capacity (due to the presence of signalized intersections) means that 1.7 urban streets would have to be built in order to achieve the same capacity as the expressway. However, an expressway's construction cost is only 1.3-1.5 times higher than that of an arterial, so the expressway is favored by both construction cost and travel-time cost.

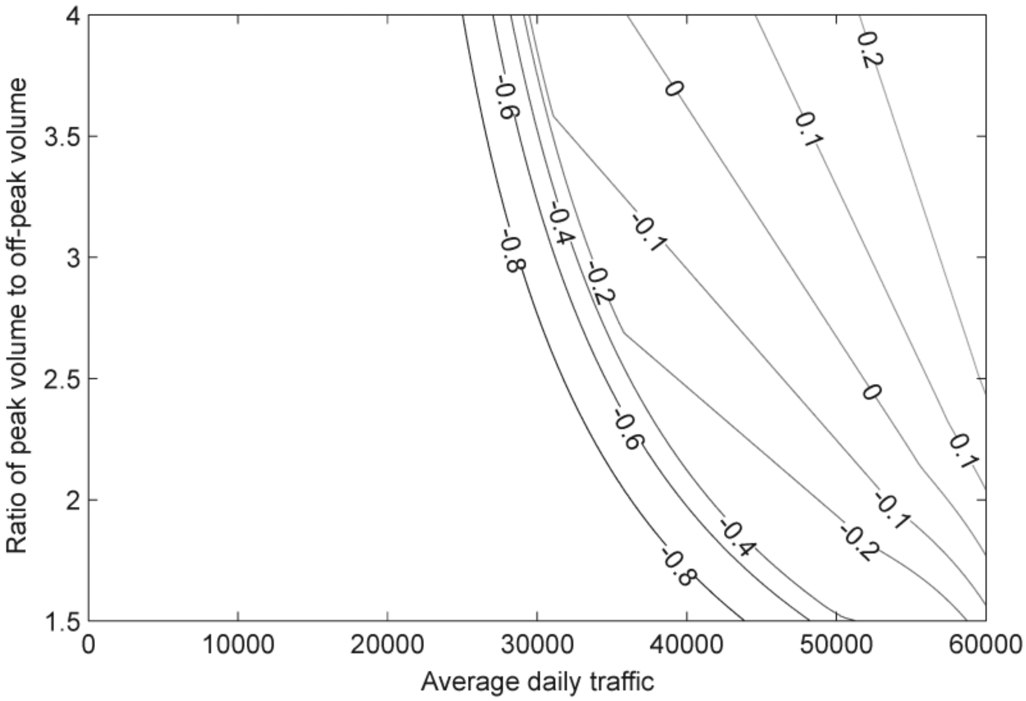
of "Freeway R" of Table 1. Its free-flow speed is 59.9 mi/h while average speed at capacity is 54.9 mi/h. This value for the arterial's speed at capacity is slightly *higher* than that of the expressway, due to the fact that speed falls less steeply with flow for unsignalized arterials than for expressways (see Exhibits 21-3 and 23-2 of the HCM). While this situation might seem artificial, it reflects what may well be a real advantage of the unsignalized arterial: more rapid accelerations and decelerations on the expressway may introduce turbulence into its flow when near capacity.

To equalize capacities, then, requires that the arterial network have 1.07 times the number of lane-miles as the expressway network — implicitly assuming the network is large enough to ignore indivisibilities. We therefore consider again a road section of $L$=10 miles, but multiply the arterial construction cost by 1.07. We assume that roads in each network provide access to the same origin and destination points, and that traffic volumes are distributed proportionally throughout a given network so that average travel times for a ten-mile trip are the same everywhere.

The free-flow travel time on the expressway is 9.16 minutes, compared to 10.02 minutes for the unsignalized arterial network (difference: 0.86 minutes). At positive traffic volumes, the expressway's speed advantage erodes and eventually turns negative due to its steeper speed-flow curve. Once capacity is reached (at the same travel volume, due to our equalizing capacities), queuing sets in, with queuing delay identical for the two road networks.

Figure 6 shows the resulting contour lines of the difference in travel times (average travel time on the expressway minus average travel time on the arterial network) for a range of $V_p/V_o$ and ADT values. The kinks seen in the -0.1 and -0.2 contour lines indicate the ADT at which queuing begins for the corresponding $V_p/V_o$. For relatively low ADT (in the region to the left of the "0" contour line), the expressway has shorter average travel times, but as just described it loses this advantage eventually as ADT increases. Similarly, for a given ADT, the expressway's time advantage erodes and eventually is lost as $V_p/V_o$ increases because then fewer travelers benefit from its higher speed under light traffic conditions.

**Figure 6. Contour map of the difference in average travel time (expressway network minus unsignalized arterial network)**




Thus, in the region to the right of the "0" contour line, building a network of unsignalized arterials is more cost-effective than building an expressway network of the same capacity, since the arterial has lower construction costs in addition to shorter travel times. However, when the expressway has shorter average travel times, we need to weigh the annual travel time savings against amortized construction costs. We examine this for the case where peak volume, $V_p$, is equal to $1.05 \times V_K$, so that each road network experiences a small amount of queuing, We again consider two values of $V_p/V_o$: 1.5 and 4. Under these conditions, average travel time on the expressway is 0.37 minutes shorter than that on the arterial network for $V_p/V_o = 1.5$, and 0.05 minutes shorter for $V_p/V_o = 4$.

To calculate travel time savings, we multiply the difference in average travel time by the ADT, by 250 days per year, and by an assumed value of travel time. We take the latter to be \$10.04 in our base case, with plus or minus 30% as low and high cases.[10] The resulting aggregate travel-time cost savings of the expressway versus the arterial network are shown in Table 2.

[10] As suggested by Small and Verhoef (2007, sect. 2.6.5), the value of time is estimated as 50% of the wage rate. According to the US BLS (2007, Table 37), the mean hourly wage for civilian workers in metropolitan areas was \$20.08 in 2006.

**Table 2. Travel-time cost savings for $V_p = 1.05V_K$ (in thousands of dollars/year)**

| | $V_p/V_o = 1.5$ | $V_p/V_o = 4$ |
|---|---|---|
| Base value of time: $10.04 | 825 | 63 |
| Low value of time: $7.03 | 577 | 44 |
| High value of time: $13.05 | 1,072 | 82 |

We now turn to construction costs. Table 3 presents estimates of construction costs for expressways and other principal arterials compiled by Alam and Kall (2005) for the US Federal Highway Administration's Highway Economic Requirements System (HERS) model, based on samples of actual projects. These figures show that construction cost per additional lane on new alignment is typically 23-31 percent lower for an arterial than for an expressway, with the larger differences applying to larger urban areas. Here we restrict our consideration to urban areas of more than 200,000 people.

**Table 3. Construction costs per lane on new alignment in urban areas (thousands of 2007 $ per mile)[1]**

| Urban area population (1000s) | Expressway | | Other Principal Arterial | | High-Type Arterial | |
|---|---|---|---|---|---|---|
| | Total | % ROW | Total | % ROW | Total | % ROW |
| 200-1,000 | 15,392 | 3.0% | 10,686 | 5.7% | 13,039 | 4.1% |
| >1,000 | 19,260 | 18.3% | 14,717 | 18.3% | 16,988 | 18.3% |

Source: See text for last two columns. Other columns computed as follows. Roadway costs are from Alam and Kall (2005, Table 9). Non-roadway costs other than right of way (*i.e.*, engineering, environmental impact and mitigation, intelligent transportation systems, urban traffic management, and bridges) are from multipliers for road way costs, in Alam and Kall (2005, Table 13). Right of way (ROW) costs are from Alam and Ye (2003, Table C-10) in the case of urban areas of 200-1,000 thousand, and from the multiplier 0.39 as recommended by Alam and Kall (2005, Table 13) in the case of urban areas of more than 1 million. All costs have been updated from 2002 to 2007 price levels using the US price deflator index of new one-family houses under construction (US Census Bureau 2008), which rose by 30.6% over those five years.

However, the cost differences from the HERS model are likely to overstate those applying to our comparison for two reasons. First, we are considering a higher type of arterial than the average in the sample. Second, these figures are based on averages for traffic conditions prevailing on actual roads built, and so may overstate somewhat the differences that would occur for a given (fixed) set of traffic conditions. Therefore we assume that the applicable costs for our high-type arterial are midway between those for expressway and other principal arterials shown in Table 3. The resulting costs are shown in the last column of the table.

Using the costs in Table 3, the amortized construction cost per lane-mile is calculated as $r \cdot C_{ROW} + [r/(1 - e^{-r\lambda})] \cdot C_{other}$, where $C_{ROW}$ and $C_{other}$ are right-of-way and other construction costs, $r$ is the interest rate (assumed to be 7%), and $\lambda$ is the effective lifetime of the road (assumed to be 20 years).[11] This amortized cost is then multiplied by the number of lanes (2), the length of the road (10 miles), and the relative number of roads (i.e., 1.00 expressway, 1.07 arterials) to obtain the total amortized construction cost of the compared road sections. Table 4 shows the results for six different cases governing the construction-cost differential between arterials and expressways. The "base" cost differential reflects costs given in Table 3 (for the larger two sizes of metropolitan areas). The "higher" and "lower" differentials are 1.5 and 0.5 times the base differential.

**Table 4. Difference in amortized construction costs for equal-capacity expressway and arterial networks (in thousands of 2007 dollars per year)**

| | Urban area population (1000s) | |
|---|---|---|
| Cost differential | 200-1,000 | >1,000 |
| Base | 2,725 | 1,920 |
| Low (x0.5) | 1,363 | 960 |
| High (x1.5) | 4,088 | 2,880 |

We can immediately see that unsignalized arterials are more cost-effective than expressways under most scenarios here, because the difference in travel-time cost is relatively small while the difference in construction cost is much higher. That is, nearly all the numbers in Table 3 are smaller than any of the numbers in Table 4. There is only one exception: if we use the high value of time, along with the low peak-to-off-peak ratio, then the travel-time saving of the expressway is quite large, $1072 thousand per year; whereas if the construction-cost differential is low and the area has population above 1 million, then the expressway's cost disadvantage is smaller than this, only $960 thousand. In all other cases shown, the higher cost of

[11] The US Office of Management and Budget (US OMB 1992) recommends 7 percent as the real interest rate for cost-benefit analysis of transportation and other projects.

building the expressway more than offsets its travel-time advantage. Selected cases are shown in Table 5.

**Table 5. Construction and travel-time cost differential (expressway minus arterial), for $V_p$=1.05$V_K$, urban population > 1 million, value of time = \$10.04/h (\$1000s/year)**

| Construction | $V_p/V_o$ | |
|---|---|---|
| Cost Differential | 1.5 | 4.0 |
| Low | 135 | 897 |
| Base | 1,095 | 1,857 |
| High | 2,055 | 2,817 |

## 5. Safety

Conventional wisdom is that many of the smaller-footprint design features considered here – narrow lanes, narrow shoulders, sharper curves, and so forth – would increase traffic accidents. This belief underlies many of the recommended standards in AASHTO (2005). We consider in this section whether increased accident costs would be likely to alter the results found so far. Because accident costs are strongly dominated by injuries and fatalities (Small and Verhoef 2007, p. 100-103), we consider evidence on them rather than on all accidents. We also consider only urban arterials or freeways of four or more lanes.

The conventional wisdom relies on several posited uses of extra lane or shoulder width: more room to accommodate temporary inattention, more room to maneuver in case of a near-accident, ability to make emergency stops off the main roadway. But there are compensating behaviors that tend to offset these advantages: higher speeds, possibly closer vehicle spacing, and a tendency to use wide shoulders for discretionary stops — which are dangerous because many accidents are associated with a vehicle stopped on the shoulder.[12] The speed increase is especially well documented, with some evidence that it occurs unconsciously (Lewis-Evans and Charlton 2006).

[12] See for example Hauer (2000a, p. 1.1) and Hauer (2000b, p. 2.1). The latter source states that 10 percent of fatal freeway accidents are due to vehicles stopped in the shoulder, and that among such stopped vehicles discretionary stops outnumber emergency stops by a factor of 7 for cars and 5 for trucks.

Statistical studies of the safety effects of wider or straighter roads do not lead to any consistent conclusion. This is partly because there are many unmeasured road attributes, such as age and hazardous terrain, that are closely correlated with design features and so may confound an attempt to isolate the effects of those design features. We review a few of the more enlightening studies here.

Hadi *et al.* (1995, Table 2) find a statistically significant effect of lane width for just two of the five types of urban multilane roads studied, with narrow lanes increasing fatal and injury accidents. The same is true of paved shoulder width. In each case, the statistical models for the other three road types omit the variable in question, either because of statistical insignificance or unexpected sign: thus, we cannot tell whether there may be cases in which narrow lanes or shoulders tend to *reduce* accidents.

Two studies have compared freeway segments before and after conversion from standard to narrow lanes, sometimes with shoulder narrowing as well. Curren (1995, pp. 35-41) finds a substantial and statistically significant increase in accident rates in three corridors studied; but a decrease, albeit not statistically significant, in the other two corridors studied. The increases were observed in those corridors in which both lanes and shoulders were quite narrow after conversion. Bauer *et al.* (2004) find an average 10–11 percent increase in accident frequency in before-and-after studies of urban freeway conversions from four to five lanes. A limitation of this approach is that other aspects of road geometry are likely to have been changed at the same time.

Potts *et al*. (2007) find generally inconsistent or statistically insignificant effects of lane width in cross-sectional comparisons across urban and suburban arterial roads, both in Minnesota (Minneapolis-St. Paul area) and Michigan (northern Detroit metropolitan area). Relative to a 12-foot lane, they find a small increase in injury accidents for 11-foot lanes, and a bigger increase for 10-foot lanes, in one out of four cases of multilane arterials in Minnesota; but inconsistent effects, including several where narrower lanes *reduced* injury crashes, in Michigan (Tables 4 and 5). They conclude:

> There was no indication that the use of … 10- or 11-ft lanes, rather than … 12-ft lanes, … led to increases in accident frequency [with the caveat that] one of the states analyzed [Minnesota] showed an increase in crash rates for four-lane undivided arterials with lane widths of … 10 ft or less. (p. 81)

The full study from which the results of Potts *et al.* are drawn also investigates shoulder width, finding that wider shoulders do substantially reduce one important category of accidents: multi-vehicle collisions not associated with intersections or driveways.[13]

Noland and Oh (2004) take a quite different approach, analyzing the aggregate number of accidents in each of 102 counties in Illinois in year 2000, as a function of average road characteristics in that county. They basically find no effects, although this may be because the data are so highly aggregated.[14] They offer as an explanation a version of Peltzman's offsetting behavior hypothesis (Peltzman 1975), by which safety improvements are partially or fully offset by more aggressive driving. Noland and Oh also point out that most studies finding a negative relationship between traffic accidents and conventional design elements (such as lane width) include few if any controls for demographics. We should mention that analyzing offsetting behavior is complex, because the various factors affecting drivers' speed and aggressiveness include some that affect how onerous or how pleasurable driving is (Steimetz 2008).

Milton and Mannering (1998) examine accident data from Washington State and find some evidence that narrow lanes and shoulders increase accident frequencies at least in some cases. These comparisons hold constant the posted speed limit but not the frequency of access points, making it difficult to know whether the finding would still hold in comparing two roads with different speed limits but identical numbers of access points. Similarly, Kweon and Kockelman (2005, Table 3) find that narrower shoulders reduce fatal and injury crash rates, by about 2.6 percent for every foot.

To summarize, both theoretical and empirical evidence linking road design to safety are ambiguous, although on balance they contain some indications that greater lane width and shoulder width may increase safety. Thus, we think it is an open question whether the "narrow" road designs considered here would in fact reduce safety, but it is certainly a potential concern. Probably it would depend on factors that vary from case to case, especially the speeds chosen by drivers.

---

[13] Harwood *et al.* (2007), ch. 5, especially Table 34.

[14] One indication is that even population, the only scale variable included, has a very small and statistically insignificant effect (elasticity about 0.01–0.06), despite the logical expectation that number of accidents would depend in some way on number of vehicles being driven.

This suggests a strategy of accompanying such roads with lower speed limits and/or other measures to discourage speeding. Of course, that raises the question of why not adopt such measures in any case. It appears that governments live in an uneasy balance between taking measures known to reduce injuries and fatalities, yet avoiding measures viewed by drivers as too intrusive. Whatever the effects of road design taken by itself, the way road design might interact with these other measures probably has a greater impact on safety. Our reading of the guidelines on highway design suggest that speed-reducing measures are more likely to be accepted when the road design is modest, because drivers then intuitively understand the rationale for them. Therefore it seems quite likely that a strategy can be developed that includes lower-footprint road designs as well as equal or better safety.

Several innovations in Europe offer hope that roads designed for lower speeds could be accompanied by measures to ensure that lower speeds in fact prevail. Variable speed limits have been used for many years in Germany and the Netherlands, and recently in Copenhagen, primarily to smooth traffic during the onset of congestion — but also with a strong reduction in injury accidents in one German implementation.[15] Another approach, used occasionally in the Netherlands and Denmark, is to is to enforce speeds by tracking vehicle licenses between control points separated by a known distance. The most draconian approach is the installation of on-vehicle systems that limit speed, with varying degrees of driver option to ignore or disable them. On-road demonstration studies of such systems have been carried out in Sweden, the Netherlands, Spain, and the UK (Liu and Tate 2004). The most effective speed limiters use "intelligent speed adaptation" to vary the allowed speed with traffic conditions. Studies using laboratory driving simulators and traffic simulation models have found that speed limiters are likely to reduce average speed, speed variation, and lane-changing movements during uncongested times — all of which are known to reduce accident rates — with little or no detrimental effects on traffic flow during congested periods.[16]

Another potentially significant safety consideration is use of automobile-only roads. As noted in the Introduction, such roads offer considerable savings in construction cost and use a lower environmental footprint. Might they also alleviate negative safety impacts of more environmentally friendly designs? Here again the evidence is mixed, but suggests that there

[15] See Mirshahi *et al.* (2007), pp. 17-18, 24, 28-29.

[16] See Compte (2000), Liu and Tate (2004), and Toledo, Albert, and Hakkert (2007).

probably would be such safety advantages. Lord, Middleton, and Whitacre (2005) find that on the New Jersey Turnpike, which has two car-only and mixed-traffic roadways that are approximately equivalent in other respects, accident rates are higher in the mixed-traffic roadway and trucks are disproportionately involved in accidents there. Fridstrøm (1999) finds that overall injury accident rates are nearly four times as responsive to amount of truck travel than to amount of car travel. Countering this is a finding by Hiselius (2004) that truck travel *reduces* the number of total accidents (injuries are not broken out separately); but this study is restricted to rural roads (in Sweden). Overall, then, it seems likely that the kind of strategy investigated in this paper — especially reducing lane widths — could be enhanced by combining it with more use of automobile-only roads (perhaps supplemented by some truck-only roads). To investigate this more fully would require additional analysis of traffic flow, construction and maintenance cost, and accident rates in mixed traffic compared to homogeneous traffic.

Despite the empirical uncertainty about safety effects of road designs, it is useful to assess the possible magnitudes of costs of increased accidents that could occur. Consider, then, the finding of Bauer *et al.* (2004) that narrower lanes might increase accident rates by around 10 percent, and suppose this applies equally to fatality and injury accidents. This is almost identical to the reductions found by Kweon and Kockelman (2005) for the four-foot shoulder reduction that we consider in Section 3. Small and Verhoef (2007, pp. 100-103) estimate average social costs of accidents for a US urban commuting trip at $0.14 per vehicle-mile. Thus a 10 percent increase in accident rates would be an increase of $0.014 per vehicle-mile. Over a ten-mile road section, this increased cost of $0.14 per vehicle would offset a travel-time savings of 0.84 minutes if value of time is $10 per hour (approximately the value used in the base scenario in the previous section). This would reverse the advantage of the "narrow" design in only a very small slice of the parameter space illustrated in Figure 5. Thus, it seems unlikely that safety considerations would reverse any of the conclusions of our analysis in Section 3.

What about our comparison of expressways versus arterials in Section 4, which holds lane width constant? Expressways are generally recognized to be safer than arterials, despite the higher speeds and more difficult lane maneuvers often encountered on expressways. In 2006, the fatality rate per 100 million vehicle-miles traveled on expressways was 0.62, compared to 1.13 for other principal arterials (US FHWA, 2006a). Kweon and Kockelman (2005) examine the effects of functional category, speed limit, curvature, and other variables on accident rates in

Washington State. They find that expressways of Interstate standards are associated with a 46.1% and a 14.7% decrease in fatal and injury crash rates, respectively, compared to other limited-access principal arterials.[17] These effects could be modified if we accounted for the higher speed limit of an expressway; we do not make such an adjustment because although speed limit is controlled for in their model, it produces non-monotonic and somewhat unreliable results according to their discussion — as indeed it does in several other studies, probably due to confounding variables.[18] However, we do adjust for the likelihood that our high-type arterial is safer than an average arterial; consistent with our treatment of the cost differential, we assume its safety disadvantage relative to an expressway is half that of an average arterial.

Using these results and Small and Verhoef's (2007, Table 3.4) estimates that the social costs per mile of fatal and injury accidents are \$0.077 and \$0.057, respectively, the social costs for a ten-mile expressway section are lower by \$0.18 and \$0.04 per trip due to lower fatal and injury accidents, respectively, compared to an arterial. Multiplying these costs by ADT and 250 days per year, the difference in the annual social costs of both types of accidents for the example networks in Section 4 is about \$2.9 million and \$1.7 million for $V_p/V_o = 1.5$ and $V_p/V_o = 4$, respectively. Comparing these differences with the values in Table 5, we find that the arterial network has lower costs only in two of the six parameter values shown there: namely the base and high cost differentials when $V_p/V_o = 4$. The numbers used here are imprecise, especially since Kweon and Kockelman point out that there is not much variation in their data on fatal accidents; but these numbers suggest that the relative safety features of expressways versus arterials may be quite important in a cost-benefit analysis.

We conclude that accident costs potentially, although not certainly, offset the potential advantage of an arterial compared to an expressway, given the current determinants of safety on roads. Thus we believe that any move to replace expressways with arterials in metropolitan

---

[17] These numbers are computed as the differences between the coefficients on "indicator for interstate highway" and "indicator for principal arterial" for the columns "fatal crash" and "injury crash," respectively, in Kweon and Kockelman (2005, Table 3). Note the "indicator for limited access" applies to both roads so is not part of the comparison.

[18] The effect of speed limit in Kweon and Kockelman's multi-equation model is complex and largely absent for fatality crashes; they consider it unreliable due to limited fatality observations and to correlation with unobserved design variables. In an auxiliary model, they find that observed *increases* in speed limits over the time period of their data decreases accidents up to a change to 60 mi/hr, then increases them. Other studies obtaining inconsistent or counter-intuitive results for speed include

planning would need to be accompanied by a thorough analysis of accidents and would best be part of a comprehensive approach to using speed control or other measures to address safety.

## 6. Conclusion

It seems that the intuitive arguments made in the introduction hold up under quantitative analysis. For both freeways and signalized urban arterials, squeezing more lanes into a fixed roadway width has huge payoffs when highway capacity is exceeded by even a small amount during peak periods, whereas the payoff from the higher off-peak speeds offered by wider lanes and shoulders is very modest. The advantage of the road with narrower lanes is accentuated when the ratio of peak to off-peak traffic is large. Similarly, for a wide range of parameters governing construction costs and values of time, the savings in travel-time costs offered by a network of expressways, compared to an equal-capacity network of high-type unsignalized arterials, is considerably smaller than the amortized value of the extra construction costs incurred. However, if accident costs are included, arterials may no longer have a cost advantage in some scenarios since expressways are associated with substantially lower accident rates, especially for fatal crashes.

Of course, the pairwise comparisons presented here do not come close to depicting the full range of relevant alternatives for road design. And because so many properties of highways are site-specific, results comparable to ours cannot be assumed to apply to any particular case without more detailed calculations. Nevertheless, we think these results provide guidance as to what types of designs deserve close analysis in specific cases, and they may provide guidance for overall policy in terms of the type of road network to be planned for. We suspect that in many cases such a network will have fewer expressways built to interstate standards, and more lower-speed expressways and high-type arterials, than are now common in the US.

Current trends present a mixed picture as to how the relative advantages of different highway designs are likely to change over time. Intractable congestion and general growth of travel, along with limited capital budgets, seem to dictate increasing traffic but probably some peak spreading, thus moving highway parameters toward the lower right in Figures 5 and 6, with uncertain implications for the comparison. If congestion pricing became widespread, that would curtail traffic while tending to make it more evenly distributed, thus moving parameters toward the lower left and making current practice relatively more attractive.

Aside from the advantages quantified here, it seems likely that the more modest highway designs suggested by these comparisons will also have more pleasing environmental and aesthetic impacts. Highways with slower free-flow speeds can fit better into existing geographical landforms and urban landscapes, permitting more curvature and grades and so requiring less earth-moving and smaller structures such as bridges and retaining walls. Tire noise and nitrogen oxides emissions are likely to be lower. Neighborhood disruption due to land condemnation and construction should be less. These advantages depend on reductions in speeds commensurate with the highway design, implying an important interaction between policies toward highway design and those toward speed control.

## Appendix A. Speeds and capacities from the HCM

This appendix discusses the HCM's methodology for calculating speeds and capacities for expressways, which the HCM calls "freeways" (based on HCM ch. 13, 23), and urban streets (based on HCM, ch. 10, 15, 16). The procedure for unsignalized urban arterials ("multilane highways" in the HCM's terminology) is generally quite similar to that of expressways, with slightly different parameter values in the speed/capacity equations (HCM ch. 12, 21).

### A.1 Expressways/Freeways

Capacity varies by free-flow speed, and so the first step is to estimate free-flow speed. The equation below is used to estimate free-flow speed (*FFS*) of a basic freeway segment (see equation 23-1 in HCM):

$$FFS = BFFS - f_{LW} - f_{LC} - f_N - f_{ID} \tag{A.1}$$

where *BFFS* is the base free-flow speed (70 mi/h for urban freeways as stated in Exhibit 13-5 of the HCM), $f_{LW}$ is the adjustment for lane width, $f_{LC}$ is the adjustment for right-shoulder lateral clearance, $f_N$ is the adjustment for number of lanes, and $f_{ID}$ is the adjustment for interchange density. The tables for these adjustment factors can be found in Exhibits 23-4 to 23-7 in the HCM, and are described below in relation to our example freeways in Section 3 of this paper.

The lane width adjustment, $f_{LW}$, is 0 when lane width is 12 ft and 6.6 when lane width is 10 ft. The width of the left shoulder has no impact on free-flow speed, while there is no reduction in free-flow speed when the right shoulder is wider than 6 ft. Thus, the right shoulder lateral clearance adjustment, $f_{LC}$, is 0 for the freeways in our example. The "regular" freeway in our example has 2 lanes and the "narrow" freeway has 3 lanes, and so the adjustment for number of lanes ($f_N$) are 4.5 and 3.0 respectively. Assuming that there are 0.5 interchanges per mile as recommended by the HCM gives us $f_{ID} = 0$.

The HCM states that base capacity is "2,400, 2,350, 2,300, and 2,250 pc/h/ln at free-flow speeds of 70 and greater, 65, 60, and 55 mi/h, respectively" (p. 23-5). A simple formula can be derived from this information, as shown in Appendix N of the Highway Performance Monitoring System (HPMS) Field Manual (Federal Highway Administration, 2002):

$$BaseCap = \begin{cases} 1{,}700 + 10FFS & \text{for } FFS < 70 \\ 2{,}400 & \text{for } FFS \geq 70 \end{cases} \tag{A.2}$$

where *BaseCap* is in passenger-cars per hour per lane.

The HCM also gives a formula (equation 23-2) for converting hourly volume $V$, which is typically in vehicles per hour, to the equivalent passenger-car flow rate $v_p$, which is in passenger-car equivalents per hour per lane (pce/h/ln) and is used later on to estimate speed:

$$v_p = V / (PHF \times N \times f_{HV} \times f_p) \tag{A.3}$$

where $PHF$ is the peak-hour factor (which represents variation in traffic flow within an hour), $N$ is the number of lanes in one direction, $f_{HV}$ is the adjustment for heavy vehicles, and $f_p$ is the adjustment for driver population (which indicates whether drivers consist of commuters who are familiar with the road or recreational drivers). For the general case where there are no data available, the HCM in Exhibit 13-5 recommends $PHF = 0.92$ for urban roads and $f_p = 1.00$ (i.e., familiar drivers).

The HCM also recommends that in the general case, the percentage of heavy vehicles on the road can be assumed to be 5% in urban settings (Exhibit 13-5). We assume that heavy vehicles on the road consist only of trucks and buses (no recreational vehicles), and that the expressway is on level terrain. This gives us $f_{HV} = 0.98$ (based on equation 23-3 of the HCM).

Using equation A.3, we can convert *BaseCap* (which is in passenger-car equivalents per hour per lane) to capacity in terms of vehicles per hour for all lanes, as shown in Appendix N of the HPMS Field Manual. The HPMS Field Manual calls this *PeakCap*, and we refer to it as $V_K$ in the model:

$$PeakCap = BaseCap \times PHF \times N \times f_{HV} \times f_p \tag{A.4}$$

The HCM also has speed-flow diagrams which depict average passenger-car speed $S$ (mi/h) as a function of the flow rate $v_p$ (pce/h/ln). We consider free-flow speeds between 55 and 70 mi/h, in which case the following formulas apply (see Exhibit 23-3 in the HCM):

For $v_p \leq (3{,}400 - 30FFS)$:

$$S = FFS \tag{A.5a}$$

For $(3{,}400 - 30FFS) < v_p \leq (1{,}700 + 10FFS)$:

$$S = FFS - \left[\frac{1}{9}(7FFS - 340)\left(\frac{v_p + 30FFS - 3{,}400}{40FFS - 1{,}700}\right)^{2.6}\right] \tag{A.5b}$$

## A.2 Urban arterials

The HCM groups urban arterial streets into several design categories. We focus on high-speed principal arterials (design category 1), which have speed limits of 45-55 mi/h and a default free-flow speed of 50 mi/h (Transportation Research Board 2000, Exhibits 10-4 and 10-5). The HCM provides little guidance for estimating free-flow speeds when field measurements are not available, and so we use the procedure recommended by Zegeer *et al* (2008, pp. 66-73). We take the case with no curbs or driveway access points along the road. Free-flow speed is then equal to a "speed constant" which in turn depends on the speed limit of the road. We assume the speed limits on the "regular" and "narrow" arterials are 55 mi/h and 45 mi/h respectively, which gives us free-flow speeds of 51.5 mi/h and 46.8 mi/h.

A vehicle's travel time on an urban street (ignoring queuing due to volumes exceeding capacity, computed separately in the text) consists of running time plus "control delay" at a signalized intersection. Based on Exhibit 15-3 of the HCM, running time for an urban arterial longer than one mile is calculated as simply the length divided by the free-flow speed; that is, the speed-flow curve is flat.

Control delay is the additional delay caused at intersections by stopping and/or waiting behind other stopped vehicles while they start up and proceed through the intersection. The HCM considers separately each "lane group" consisting of through lanes, exclusive left-turn lanes, or shared turn/through lanes. It also states that "[t]he control delay for the through movement is the appropriate delay to use in an urban street evaluation" (p. 15-4). With this, we will focus on only two lane groups, through lanes and shared right-turn/through lanes, because a given trip would make at most one left turn and we are not concerned with the time that requires.

The formula for calculating control delay for each lane group (equation 16-9 in the HCM) is the sum of three components: (1) uniform control delay, which assumes uniform arrivals; (2) incremental delay, which takes into account random arrivals and oversaturated conditions (volume exceeding capacity); and (3) initial queue delay, which considers the additional time required to clear an existing initial queue left over from the previous green period. As described in Section 2, our model assumes that queuing due to oversaturation occurs at the entry to the road, and that the queue discharges at a rate equal to the capacity of the road. Thus the traffic volume arriving at the intersection is never greater than the intersection's capacity, so only the uniform control delay is applicable. (Capacity is calculated based on an intersection's through movement capacity, as detailed below.)

The control delay is then calculated for each lane group using equations 16-9 and 16-11 of the HCM:

$$d = \frac{0.5C(1-g/C)^2}{1-\left[\min(1,X)(g/C)\right]} \cdot PF \tag{A.6}$$

where $C$ is the cycle length, $g$ is effective green time, $X$ is the volume-to-capacity ratio of that lane group, and $PF$ is a "progression adjustment factor" which accounts for the effects of synchronization (or lack of it) between adjacent signals. Using the defaults recommended by the HCM for signals spaced 3,200 or more feet apart (denoted as Arrival Type 3, see p. 10-23 of the HCM), we have $PF = 1$. We assume that the through lane group and the shared right-turn/through lane group have identical values of $g/C$ and that traffic distributes across lanes so that they have identical values of $X$.[19] Therefore both lane groups have the same delay, given by (A.6). The control delay in equation (3), then, is just $d$ multiplied by the number of signals. Because we assume that all the lanes carrying through traffic equalize their volume-capacity ratios, we can substitute our overall volume-capacity ratio $V/V_K$ for $X$, with capacity $V_K$ defined appropriately as we now describe.

The arterial's capacity is based on the saturation flow rates, $s_i$, of the two lane groups, along with the fraction of time the signal is green and the proportion of traffic at each intersection that is making left turns. (This latter proportion gets to use the left-turn lane, assumed to have ample capacity, so can be added to the capacity of the other two lane groups.) Saturation flow means the highest flow rate that can pass through the intersection while the light is green. Based on equation 16-6 of the HCM and using $i$ to index lane groups, the capacity of each lane group (denoted in the HCM by $c_i$) is:

$$c_i = s_i \cdot (g_i / C) \tag{A.7}$$

where the effective green ratio $g_i/C$ is here taken to be identical for both the through group and shared right-turn/through group, hence $g/C$.

Adding the fraction $\tau_{LT}$ of traffic volume that is making left turns, the total capacity of the road — $V_K$ in our model — is:

[19] It is also assumed that vehicles are not allowed to turn right during red signal phases.

$$V_K = (1-\tau_{LT})^{-1}(c_T + c_{RT}) \tag{A.8}$$

where $c_T$ and $c_{RT}$ are the capacities of the through and right-turn lane groups. We assume that 7.5% of the total traffic volume will be vehicles turning left, and similarly for vehicles turning right, so $\tau_{LT} = \tau_{RT} = 0.075$.[20]

The saturation flow rates needed for (A.7) are given by equation 16-4 of the HCM, which includes various adjustment factors. Many of these are equal to one because we use the corresponding HCM recommended default values (see Chapters 10 and 16). Specifically, we assume that the road is located in a non-CBD area and on level terrain, no parking is allowed, there are no buses that stop within the intersection area, and no adjustments are necessary for pedestrians or bicycles. Since we are interested in estimating capacity, we assume that there is uniform use of the available lanes (i.e., there is no adjustment for lane utilization), as recommended by the HCM (p. 10-26). Also, there are no left-turn adjustments since we assume a separate, exclusive left-turn lane and treat any delay in making left turns (which happens at most once in any trip) as part of the access time to a final destination rather than delay time on the road in question. We also follow the HPMS Field Manual's lead and multiply the HCM's original equation for saturation flow by the peak hour factor (*PHF*) rather than adjusting volumes by that factor (see p. N-19 of the HPMS Field Manual).

With these assumptions, the saturation flow rate for a lane group is:

$$s = s_0 N f_w f_{HV} f_{RT} PHF$$

where $s_0$ is the base saturation flow rate per lane (pce/h/ln), $N$ is the number of lanes in the lane group, $f_w$ is the adjustment factor for lane width, $f_{HV}$ is the adjustment factor for heavy vehicles, and $f_{RT}$ is the adjustment factor (applying only to the right-turn group and accounting for vehicles having to reduce speed to make the turn). The HCM recommends $s_0$ = 1,900 pce/h/ln. The lane width adjustment, $f_{LW}$, is 1 when lane width is 12 ft and 0.93 when lane width is 10 ft. Assuming the percentage of heavy vehicles in the traffic stream is 5% as in the case of expressways, $f_{HV}$ = 0.95 (different from the heavy vehicle adjustment factor for expressways). For the shared right-turn/through lane group, the right-turn adjustment factor $f_{RT}$ is 1-0.15$P_{RT}$, where $P_{RT}$ is the

[20] A typical urban trip length is 10 miles (Pisarski, 2006). Using Lake Shore Drive in Chicago as an example road, there are 1.5 exits per mile, or 7.5 exits passed by a typical trip if half of it takes place on the arterial. We assume therefore that a fraction 1/7.5 = 0.133 of trips exit at each exit. We raise this to 0.15 to account for the likelihood that the critical bottleneck intersections are those with the most turning traffic. If left and right turns are evenly balanced, that gives 7.5% right turns and 7.5% left turns at the intersection whose capacity is being computed.

percentage of right-turning vehicles in that lane group. As mentioned earlier, the proportion of total traffic volume that turns right at each intersection is $\tau_{RT} = 0.075$. This gives us $P_{RT}$ equal to 0.152 and 0.230, and correspondingly, $f_{RT}$ equal to 0.978 and 0.966, for the two-lane and three-lane arterials in our example.[21] For the through lane group, $f_{RT} = 1$ by definition. As in the case of expressways, the peak hour factor, *PHF*, is assumed to be 0.92.

[21] Let the saturation flow rate of the through lane and the number of through lanes of a particular arterial be denoted by $s_T$ and $n$, respectively. The saturation flow rate for the right-turn lane group is $s_{RT} = s_T(1 - 0.15P_{RT})$. At capacity, $P_{RT}$ is the number of vehicles turning right divided by the capacity of the right-turning lane group: $\tau_{RT}[ns_T + s_T(1 - 0.15P_{RT})]/[s_T(1 - 0.15P_{RT})]$. After simplifying, we obtain a quadratic equation in terms of $P_{RT}$, which can be solved for a value of $P_{RT}$ between 0 and 1 for given $n$ and $\tau_{RT}$.

## References


AASHTO (2004) *A Policy on Geometric Design of Highways and Streets*, $5^{th}$ edition, Washington, D.C.: American Association of State Highway and Transportation Officials.

AASHTO (2005) *A Policy on Design Standards – Interstate System*, 5th Edition, Washington, D.C.: American Association of State Highway and Transportation Officials.

Alam, Mohammed and David Kall (2005) *Improvement Cost Data: Final Draft Report*, prepared for the Office of Policy, Federal Highway Administration.

Alam, Mohammed and Qing Ye (2003) *Highway Economic Requirements System Improvement Cost and Pavement Life: Final Report*, prepared for the Office of Policy, Federal Highway Administration.

Bauer, Karin M., Douglas W. Harwood, Warren E. Hughes, and Karen R. Richard (2004) "Safety Effects of Narrow Lanes and Shoulder-Use Lanes to Increase Capacity of Urban Freeways," *Transportation Research Record: Journal of the Transportation Research Board*, 1897: 71-80.

Cassidy, Michael J. and Robert L. Bertini (1999) "Some traffic features at freeway bottlenecks," *Transportation Research Part B*, 33: 25-42.

Cassidy, Michael J. and Jittichai Rudjanakanoknad (2005) "Increasing the capacity of an isolated merge by metering its on-ramp," *Transportation Research Part B*, 39: 896-913.

Compte, Samantha L. (2000) "New Systems: New Behaviour?" *Transportation Research Part F*, 3: 95-111.

Curren, J.E. (1995), *Use of Shoulders and Narrow Lanes to Increase Freeway Capacity*, National Cooperative Highway Research Program Report 369. Washington, D.C.: Transportation Research Board.

Downs, Anthony (1962) "The law of peak-hour expressway congestion," *Traffic Quarterly*, 6: 393-409.

Downs, Anthony (2004) *Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion*, Washington, DC: Brookings Institution Press.

Federal Highway Administration (2002) *Highway Performance Monitoring System Field Manual*. http://www.fhwa.dot.gov/ohim/hpmsmanl/hpms.htm

Fridstrøm, Lasse (1999) "An Econometric Model of Car Ownership, Road Use, Accidents, and their Severity," TØI Report 457/1999, Oslo: Institute of Transport Economics. http://www.toi.no/getfile.php/Publikasjoner/T%D8I%20rapporter/1999/457-1999/457-1999.pdf

Hadi, Mohammed A., Jacob Aruldhas, Lee-Fang Chow, and Joseph A. Wattleworth (1995) "Estimating Safety Effects of Cross-Section Design for Various Highway Types Using Negative Binomial Regression," *Transportation Research Record: Journal of the Transportation Research Board*, 1500: 169-177.

Harwood, Douglas W., Karin M. Bauer, Karen R. Richard, David K. Gilmore, Jerry L. Graham, Ingrid B. Potts, Darren J. Torbic, and Ezra Hauer (2007) *Methodology to Predict the Safety Performance of Urban and Suburban Arterials*, National Cooperative Highway Research Program Web-Only Document 129: Phases I and II. Washington, D.C.: Transportation Research Board. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w129p1&2.pdf

Hauer, Ezra (2000a) "Lane Width and Safety," unpublished draft. Toronto: RoadSafetyResearch .com., March. http://ca.geocities.com/hauer@rogers.com/Pubs/Lanewidth.pdf

Hauer, Ezra (2000b) "Shoulder Width, Shoulder Paving and Safety," unpublished draft. Toronto: RoadSafetyResearch.com, March. http://ca.geocities.com/hauer@rogers.com/Pubs/Shoulderwidth.pdf

Kweon, Young-Jun and Kara M. Kockleman (2005) "Safety Effects of Speed Limit Changes: Use of Panel Models, Including Speed, Use, and Design Variables," *Transportation Research Record: Journal of the Transportation Research Board*, 1908: 148-158.

Lewis-Evans, Ben, and Samuel G. Charlton (2006) "Explicit and Implicit Processes in Behavioural Adaptation to Road Width," *Accident Analysis and Prevention*, 38: 610-617.

Liu, Ronghu, and James Tate (2004) "Network Effects of Intelligent Speed Adaptation Systems," *Transportation*, 31: 297-325.

Milton, John, and Fred Mannering (1998) "The Relationship among Highway Geometrics, Traffic-Related Elements and Motor-Vehicle Accident Frequencies," *Transportation*, 25: 395-413.

Mirshahi, Mohammad, Jon Obenberger, Charles A. Fuhs, Charles E. Howard, Raymond A. Krammes, Beverly T. Kuhn, Robin M. Mayhew, Margaret A. Moore, Khani Sahebjam, Craig J. Stone, and Jessie L. Yung (2007) *Active Traffic Management: The Next Step in Congestion Management*, Report No. FHWA-PL-07-012. Alexandria, Virginia: American Trade Initiatives. http://international.fhwa.dot.gov/pubs/pl07012/

Noland, Robert B., and Lyoong Oh (2004) "The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data," *Accident Analysis and Prevention*, 36: 525-532.

Peltzman, Sam (1975) "The effects of automobile safety regulation," *Journal of Political Economy* 83: 677-725.

Pigou, Arthur C. (1920) *The Economics of Welfare,* London: Macmillan.

Pisarski, Alan E. (2006) *Commuting in America III: The Third National Report on Commuting Patterns and Trends*, ed. Nathan Gartner, Carroll J. Messer, and Ajay K. Rathi, National Cooperative Highway Research Program Report 550 and Transit Cooperative Research Program Report 110, Washington, D.C.: Transportation Research Board.

Potts, Ingrid B., Douglas H. Harwood, and Karen R. Richard (2007) "Relationship of lane width to safety for urban and suburban arterials," *Transportation Research Record: Journal of the Transportation Research Board*, 2023: 63-82.

Rouphail, Nagui, Andrzej Tarko, and Jing Li (1996) "Traffic Flow at Signalized Intersections," in *Traffic Flow Theory: A State of the Art Report – Revised Monograph on Traffic Flow Theory*, Turner-Fairbank Highway Research Center. http://www.tfhrc.gov/its/tft/tft.htm

Small, Kenneth A. and Erik T. Verhoef (2007) *The Economics of Urban Transportation*, London and New York: Routledge.

Steimetz, Seiji S.C. (2008) "Defensive Driving and the External Costs of Accidents and Travel Delays," *Transportation Research Part B*, forthcoming.

Toledo, Tomer, Gila Albert, and Shalom Hakkert (2007) "Impact of Active Speed Limiters on Traffic Flow and Safety: Simulation-Based Evaluation," *Transportation Research Record*, 2019: 169-180.

Transportation Research Board (2000) *Highway Capacity Manual 2000*. Washington, D.C.: Transportation Research Board.

US BLS (2007) *National Compensation Survey: Occupational Earnings in the United States, 2006*, Washington D.C.: US Bureau of Labor Statistics. http://www.bls.gov/ncs/ncswage2006.htm

US Census Bureau (2008) *Construction Price Indexes: Price Deflator Index of New One-Family Houses Under Construction*, Washington, D.C.: US Government Printing Office. http://www.census.gov/const/price_deflator.pdf

US FHWA (2006a) *Highway Statistics*, Washington, D.C.: Federal Highway Administration. http://www.fhwa.dot.gov/policy/ohim/hs06/index.htm

US FHWA (2006b) *Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance*, Washington, D.C.: Federal Highway Administration. http://www.fhwa.dot.gov/policy/2006cpr/

US OMB (1992) *Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs*, circular no. A-94, revised, Section 8. Washington, DC: US Office of Management and Budget.

Zegeer, John D., Mark Vandehey, Miranda Blogg, Khang Nguyen, and Michael Ereti (2008) *Default Values for Highway Capacity and Level of Service Analyses*, National Cooperative Highway Research Program Report 599, Washington D.C.: Transportation Research Board.