Title

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

Authors
Ng, Chen Feng
Small, Kenneth

Publication Date
2011-05-17

University of California Transportation Center
UCTC-FR-2011-27

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

Chen Feng Ng, California State University, Long Beach, and
Kenneth A. Small, University of California, Irvine
September 2011

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

Chen Feng Ng and Kenneth A. Small

May 17, 2011

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 modern 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 (corresponding author)
Department of Economics
California State University, Long Beach
1250 Bellflower Blvd
Long Beach, CA 90840
cng3@csulb.edu
Tel: 562-985-8513; Fax: 562-985-5804

Kenneth Small
Department of Economics
3151 Social Science Plaza
University of California at Irvine
Irvine, CA 92697
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 (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 spillovers, 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 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 parkway amenities. 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.

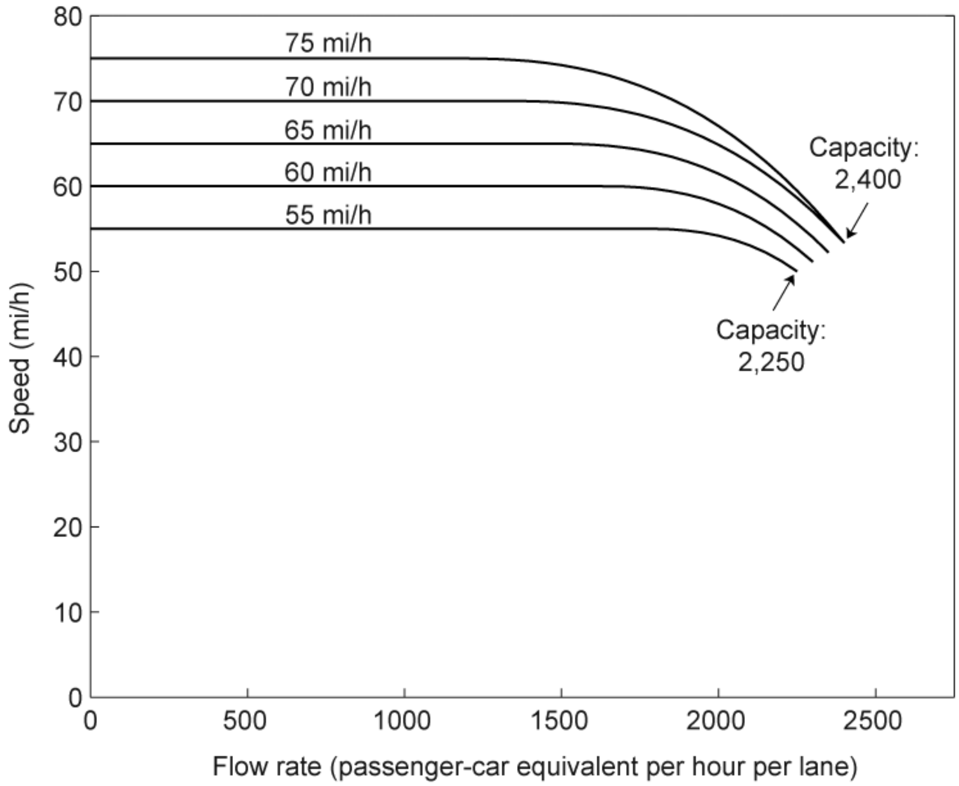
2. Congestion Formation, Capacity, and Travel Time

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 slower speeds 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 and which accounts for the effects of traffic signals. 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 here 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 (HCM) (Transportation Research Board 2000). The free-flow speed is $S^0 \equiv S(0)$. For signalized urban arterials, the HCM 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).This simplification ignores possible effects of spillover queues on other roads, but does not appear to bias our comparisons in any particular direction. For the sake of concreteness, we assume the bottleneck occurs at the entry to the section of road under consideration. We further simplify by assuming the dynamic demand pattern can be approximated by two periods, "peak" and "off-peak", each with constant flow. 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, $D_{max}=P\cdot[(V_p/V_K)-1]$. 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

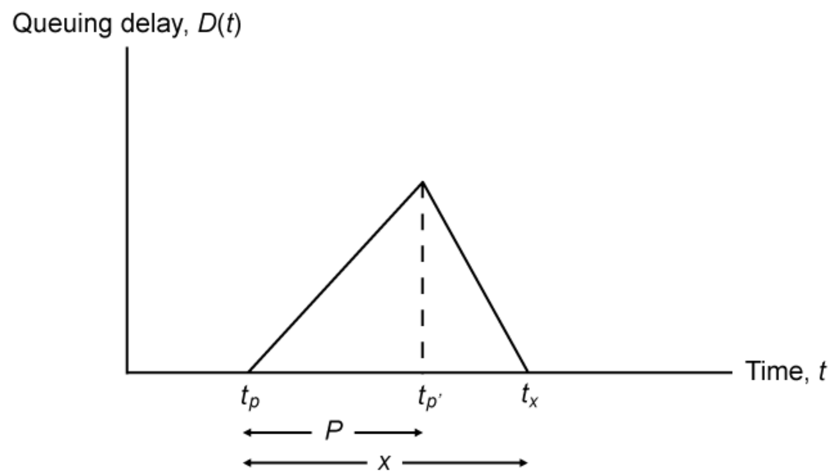
$$\overline{D}_p = \tfrac{1}{2} D_{\max} = \tfrac{1}{2} P \cdot [(V_p / V_K) - 1]. \tag{1}$$

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 additional time lost slowing and waiting for signalized intersections. 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, δ, is:

$$\delta = Z \left\{ \frac{0.5C \cdot (1 - g/C)^2}{1 - [(V/V_K)(g/C)]} + 900T \left[\left(\frac{V}{V_K} - 1 \right) + \sqrt{\left(\frac{V}{V_K} - 1 \right)^2 + \frac{8kV}{TV_K^2}} \right] \right\} \tag{3a}$$

where C is signal cycle length, g is effective green time, V is the volume of traffic going through the intersection, T is the duration of the analysis period (in hours), and k is the incremental delay factor. Time durations C, g, and hence δ are all conventionally measured in seconds. The first and second terms within the curly brackets in Equation (3a) are known as uniform control delay and incremental delay, respectively. The uniform control delay 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. 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). The incremental delay accounts for the saturation level of the road as well as random arrivals, adjusted for the type of signal control through the factor k (described in detail in Appendix A).

Recall that due to our assumption that V_K limits upstream flow (i.e., queuing when $V_p>V_k$ occurs at the entrance to the road, prior to the first signal), V is equal to the queue discharge rate V_K during the time period from t_p to t_x. 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 obtained by setting V=V_K and T=x in (3a):

$$\delta_p = Z\left[0.5C(1-g/C)+900x\sqrt{(8k)/(xV_K)}\right]. \tag{3b}$$

For all other off-peak travelers, the control delay is given by (3a) with V=V_o and T=F-(x-P):

$$\delta_o = Z\left\{\frac{0.5C\cdot(1-g/C)^2}{1-[(V_o/V_K)(g/C)]}+900(F-x+P)\left[\left(\frac{V_o}{V_K}-1\right)+\sqrt{\left(\frac{V_o}{V_K}-1\right)^2+\frac{8kV_o}{(F-x+P)V_K^2}}\right]\right\} \tag{3c}$$

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:

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

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

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

We now need to make explicit assumptions about the number of days per year for which our analysis applies. It is common in large urban areas to have severe congestion on weekends, but for a shorter period than on weekdays. Therefore we assume that the peaking analysis applies to regular work days (255 days per year) plus half the remaining days (55 per year), and we denote the total travel time for all drivers on each of these days as

$$TT_w = TT_o + TT_p \tag{6}$$

It is assumed that the remaining 55 days have traffic volumes typical of the off-peak periods during work days. For simplicity we call them "Sundays" but they actually represent various parts of the 110 annual non-work days. Thus by definition, each driver on a "Sunday" has travel time $L/S(V_o)$, so that total travel time for all drivers on a Sunday is:

$$TT_s = (F+P) \cdot V_o \cdot \frac{L}{S(V_o)} \tag{7}$$

With these assumptions, we can express the average daily traffic (ADT) as

$$ADT = (310/365) \cdot (PV_p + FV_o) + (55/365) \cdot (P+F) \cdot V_o \tag{8}$$

The annual total and average travel time for all vehicle trips are, respectively:

$$TT_{all} = 310 \cdot TT_w + 55 \cdot TT_s \tag{9}$$

$$\overline{TT}_{all} = \frac{TT_{all}}{365 \cdot ADT} \tag{10}$$

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.[3] This comparison enables us to focus on the two primary factors

[3] 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

that distinguish these designs from each other: travel time and safety. We discuss safety in Section 5; in this section, we focus solely on travel time. By choosing designs for which trucks are permitted, we minimize possible differences in cost or safety due to rerouting or separation of truck traffic. In actual implementation, some of the changes analyzed here might be done only during rush hours, thereby lessening the impacts on ease of truck travel.

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 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 specifications are summarized in Figure 3 and Table 1. First, consider expressways. We specify Expressway R (the "regular" design) with 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 (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.

Next, consider signalized urban arterials. Here 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 and an 8 ft right shoulder (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.

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).

Figure 3a. Example expressways (one direction)

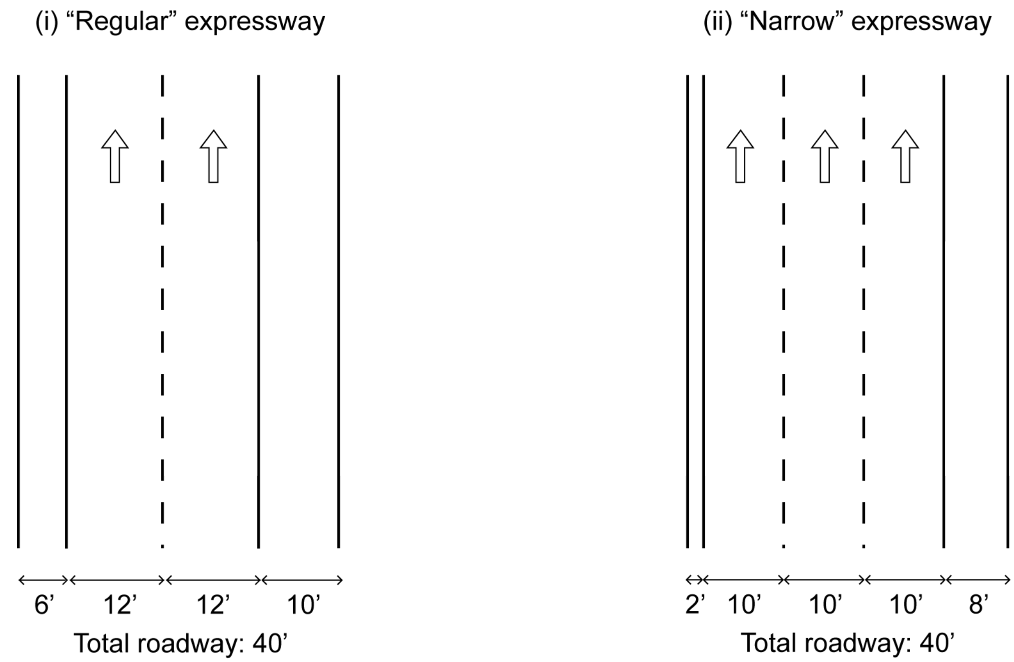


Figure 3b. Example urban streets (one direction)

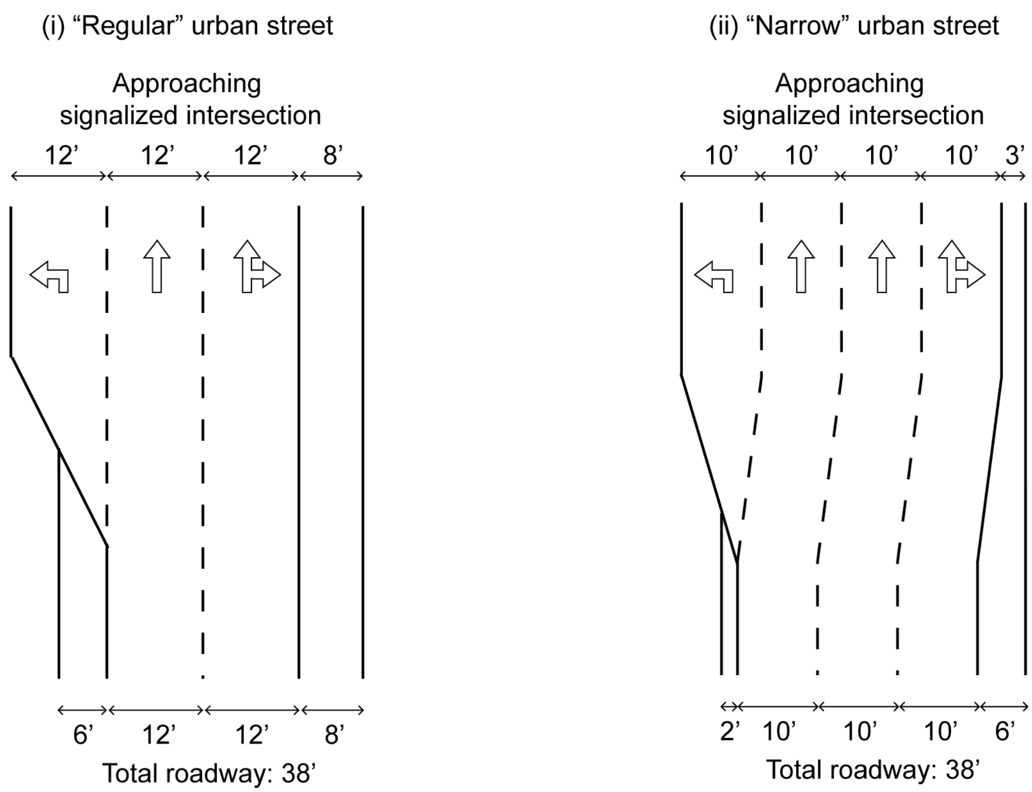


Table 1. Specifications for examples (all one direction)

	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[a]	0.05	0.05	0.05	0.05
Driver population factor[a]	1.00	1.00	1.00	1.00
Peak hour factor[a]	0.92	0.92	0.92	0.92
Interchanges/signals per mile[a]	0.50	0.50	0.50	0.50
Signal cycle length (s)[a]	-	-	100	100
Effective green time (s)[a]	-	-	70	70
Speed and capacity				
Free-flow speed (mi/h)[b]	65.5	60.4	51.5	46.8
Speed at capacity (mi/h)[b]	52.3	51.2	51.5	46.8
Capacity per lane (veh/h/ln)[c]	2,113.76	2,067.98	1,165.33	1,087.64
Total capacity, V_K (veh/h)[c]	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:
[a] 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.
[b] 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.
[c] 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.

The "narrow" urban street (Urban Street N), by contrast, has three 10 ft lanes in one direction for through movement, a 2 ft left shoulder, and a 6 ft right shoulder—the latter serving to minimize the impact of improperly stopped vehicles on traffic flow. 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.[4] 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.

Using the specifications listed in Table 1, we can calculate average travel times, including queuing delay and control delay, if applicable, for a range of traffic volumes and time-of-day distributions. We consider the 16-hour daily period (6 a.m. – 10 p.m.), during which about 95 percent of all trips take place (Hu and Reuscher 2004, Table 28), and assume each one-directional roadway has a peak period lasting 4 hours: hence P=4 and F=12. We report results for two fairly extreme assumptions, corresponding to relatively little to a great deal of peaking: V_p/V_o = 1.25 and V_p/V_o = 2.[5]

Figures 4a-b show the resulting average travel times for the four different road designs under different values for average daily traffic (see equation 8), and the ratio of peak volume (V_p) to off-peak volume (V_o). (Times are shown for ADT up to the value that would leave a queue remaining at the end of the off-peak period.) We see from Figure 4a that when V_p/V_o = 1.25, the "regular" freeway experiences queuing when ADT exceeds 56,984, but queuing does not occur on the "narrow" freeway for ADT values up to 65,000 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

[4] 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.

[5] By way of comparison, one may consider that 37 percent of all national person trips occur within one of the two peak periods reported by Hu and Reuscher (namely 6–9 a.m. and 1–4 p.m.), or 5.9 percent per hour, much of which we can presume is concentrated in one direction; whereas 59 percent occur during the off-peak hours (between 6 a.m. and 10 p.m. but outside those peak periods), or 6.2 percent per hour, presumably spread rather evenly across both directions on most roads. This would suggest a peaking ratio V_p/V_o between one and two as typical for an average across all areas. A much higher value can be expected for some commuting roads that carry heavy traffic from mostly residential outlying areas into central business districts.

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 33,576 and 47,001, respectively. The average travel time on the “regular” urban street starts to exceed that of the “narrow” urban street when ADT is greater than 33,144.

Figure 4b shows the average travel times for the four highway types when $V_p/V_o = 2$. 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.

Figure 4a. Average travel times for $V_p/V_o = 1.25$

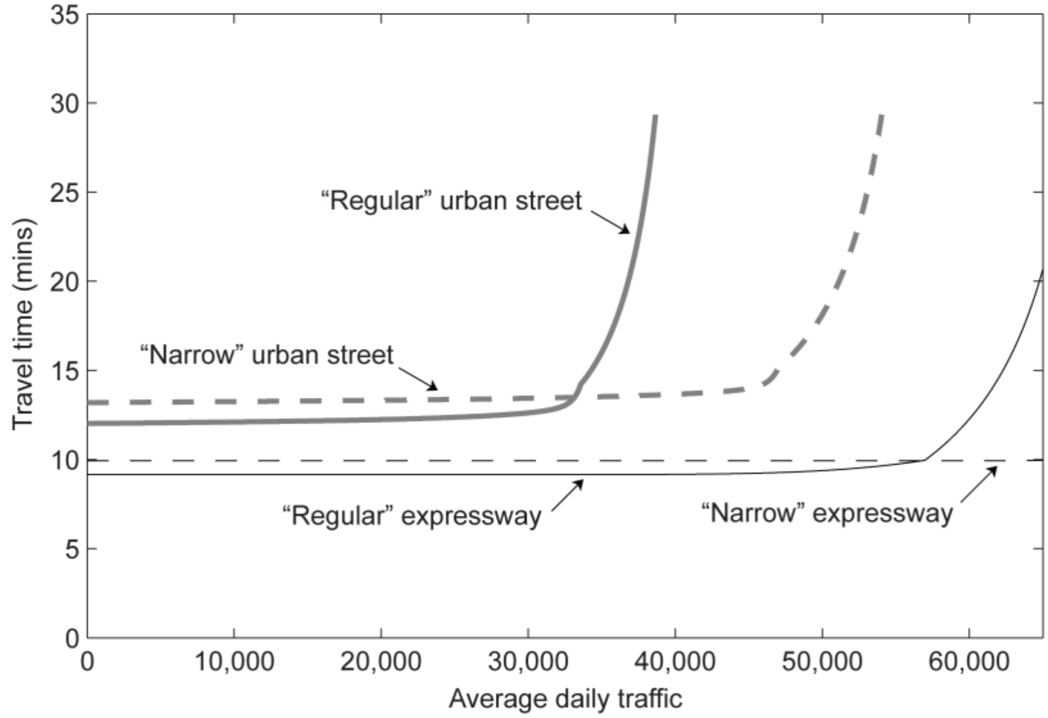


Figure 4b. Average travel times for $V_p/V_o = 2$

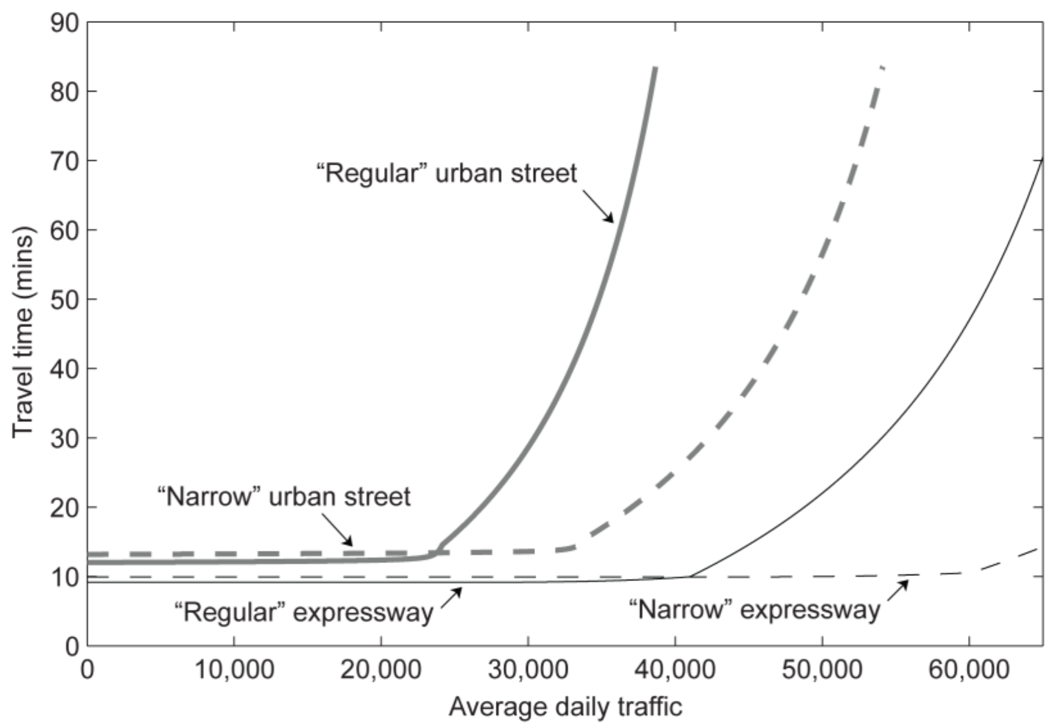


Figure 5a. Contour map of the difference in average travel time ("regular" expressway minus "narrow" expressway)

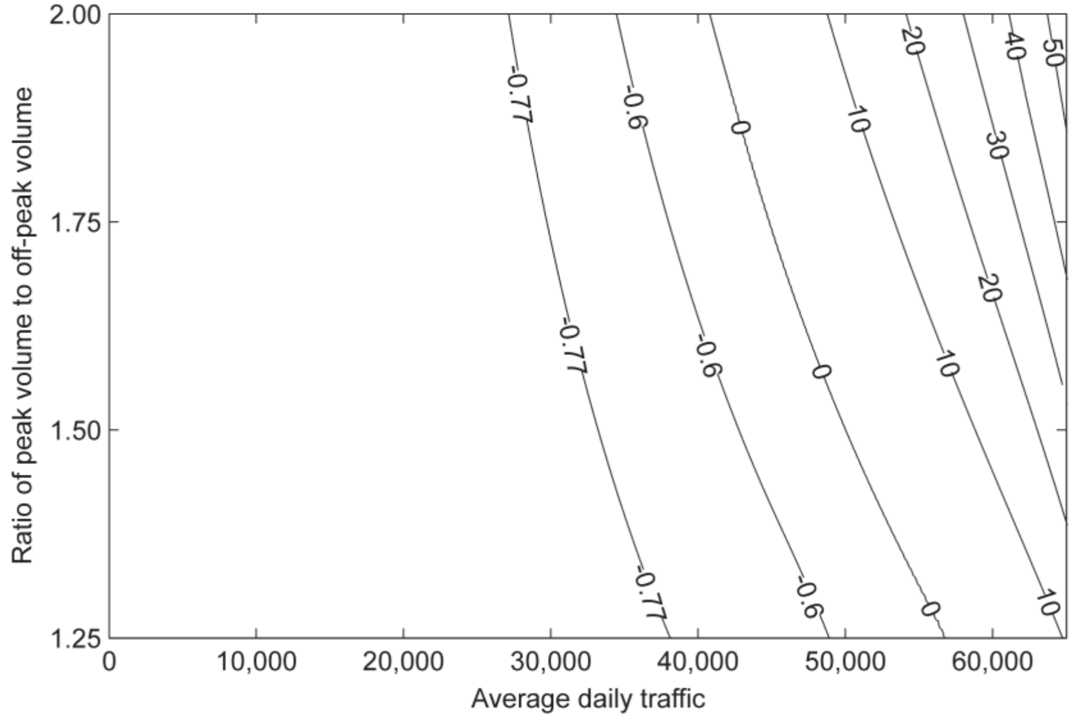
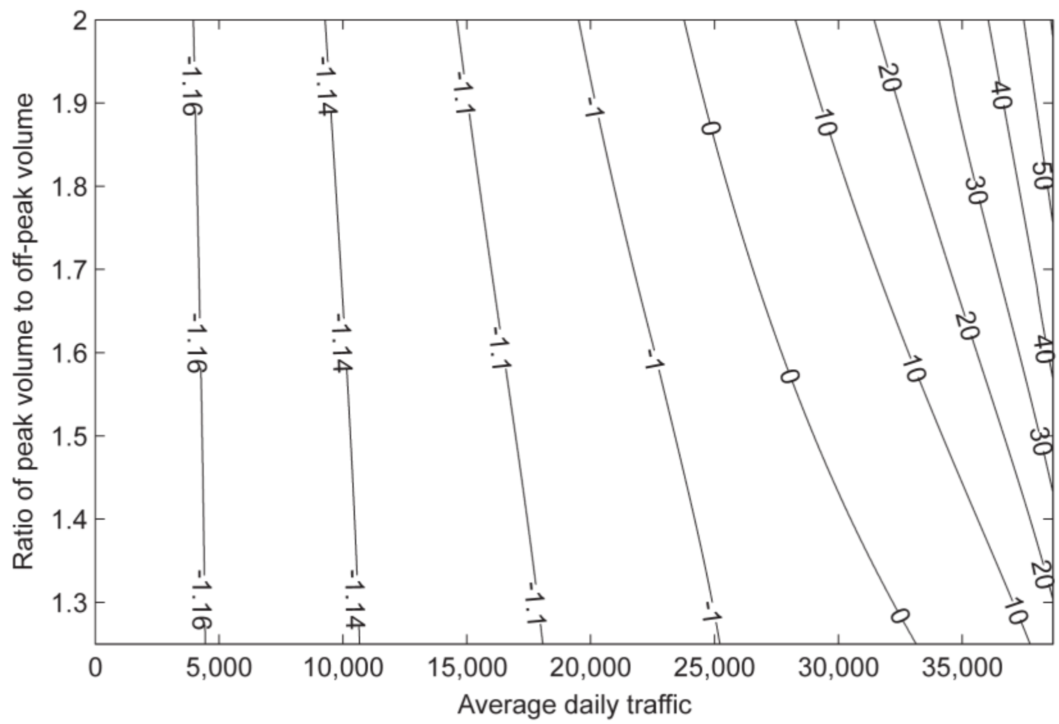


Figure 5b. Contour map of the difference in average travel time ("regular" urban street minus "narrow" urban street)



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 meets the HCM definition 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. Examples include the Arroyo-Seco Parkway (SR-110) in the Los Angeles region, Storrow Drive in Boston, and most of Lake Shore Drive in Chicago.[6] Since not all intersections are grade-separated, drivers on minor roads may have to take a longer route to one of the grade-separated crossings; this additional delay could be accounted for in our model with some additional assumptions, but we think it would not make enough difference to be worth the extra complication. In the case of Storrow Drive and Lake Shore Drive, cross traffic is less of an issue due to the fact that they both run along a river or lakefront, so there are few streets that cross all the way through.

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.

[6] Lake Shore Drive includes six signalized intersections (only five going southbound) within its 15-mile length, for an average spacing of over two miles; but all the signals are within a central section about 2.4 miles in length. This highway opened in 1937 (Chicago Area Transportation Study 1998).

We therefore compare the "regular" expressway from the previous section with an unsignalized arterial with similar characteristics.[7] 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 of "Freeway R" of Table 1. 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. (We also assume travel distances are the same for both networks, thereby ignoring the arterial network's advantage in having more total road-miles, thereby providing more access points and therefore reducing some trip distances.)

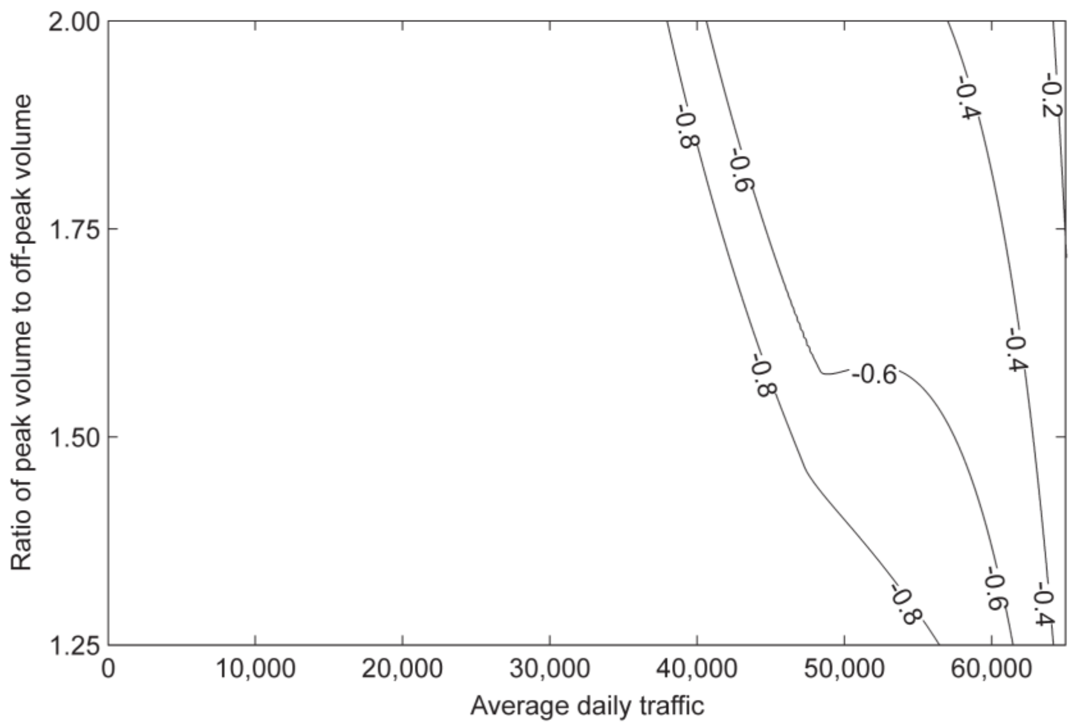
The HCM free-flow speeds are is 65.5 and 59.9 mi/h for the expressway and arterial, respectively (exclusive of any signal delays). There is an anomaly, however, in comparing their speeds at capacity: the HCM formulas imply a slightly *lower* value for the expressway than for the arterial (52.3 versus 54.9 mi/h), 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 could possibly be explained by a disruptive effect of more rapid accelerations and decelerations on the expressway, we take the more conservative stance that it is an anomaly resulting from different chapters of the HCM being developed by different research groups. Thus, we equalize the speeds at capacity by increasing the rate at which arterial speed falls with traffic volume (see Appendix A).

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 gradually erodes. 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.

[7] 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.

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. We can see that the expressway always has shorter travel times up to an ADT of 65,000 (beyond this, the queuing period is longer than the 12 hour off-peak period). The kink seen in the -0.6 contour line indicates the ADT at which queuing begins for the corresponding value of V_p/V_o.

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



Even though the expressway has shorter average travel times, we need to weigh the annual travel time savings against amortized construction costs since the arterial has lower 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: 1.25 and 2. Under these conditions, average travel time on the expressway is 0.68 minutes shorter than that on the arterial network for $V_p/V_o = 1.25$, and 0.54 minutes shorter for $V_p/V_o = 2$. We value these time savings at \$10.50 per vehicle in our base case, with plus or minus 30% as low and high cases.[8] The resulting aggregate travel-time cost savings of the expressway versus the arterial network are shown in Table 2.

Table 2. Travel-time cost savings of the expressway for $V_p = 1.05V_K$ (in thousands of 2009 dollars/year)

	$V_p/V_o = 1.25$	$V_p/V_o = 2$
Base value of time: \$10.50	2,601	1,494
Low value of time: \$7.35	1,821	1,046
High value of time: \$13.64	3,379	1,941

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 smaller differences applying to larger urban areas. We restrict our consideration to urban areas of more than 200,000 people.

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; since the expressways are likely handling more traffic than the arterials, their costs are likely to reflect some design features motivated by this higher traffic; thus the HERS cost differences may overstate the differences that would occur for a given (fixed) set of traffic conditions such as we consider here. 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.

[8] According to Small and Verhoef (2007, sect. 2.6.5), the value of time for work trips is typically estimated as 50% of the wage rate, which would be about \$10.50 per hour for 2009 (BLS 2010, Table 1, reporting mean hourly wage for civilian workers). We assume these value of time studies apply to vehicles, not people, although authors are often ambiguous. According to Hu and Reuscher (2004, Table 16), average vehicle occupancy is 1.63 for all trip purposes averaged. Values of time are higher for work trips than for others, but occupancies are lower; we assume these two factors balance out between peak and off-peak travel so assign them both the same value of time.

Table 3. Construction costs per lane on new alignment in urban areas (thousands of 2009 dollars per mile)[1]

Urban area population (1000s)	Expressway		Other Principal Arterial		High-Type Arterial	
	Total	% ROW	Total	% ROW	Total	% ROW
200-1,000	14,507	3.0%	10,072	5.7%	12,289	4.1%
>1,000	18,152	18.3%	13,871	18.3%	16,012	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 adjusted from 2002 dollars to the average price level in the years 2007-2009 using the Federal Highway Administration's Composite Bid Price Index (FHWA 2003) and National Highway Construction Cost Index (FHWA 2010). The average price level was used as construction costs were very volatile during those years.

.

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%, as recommended by the Office of Management and Budget [1992] for cost-benefit analyses of transportation and other projects), and λ is the effective lifetime of the road (assumed to be 25 years). 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.

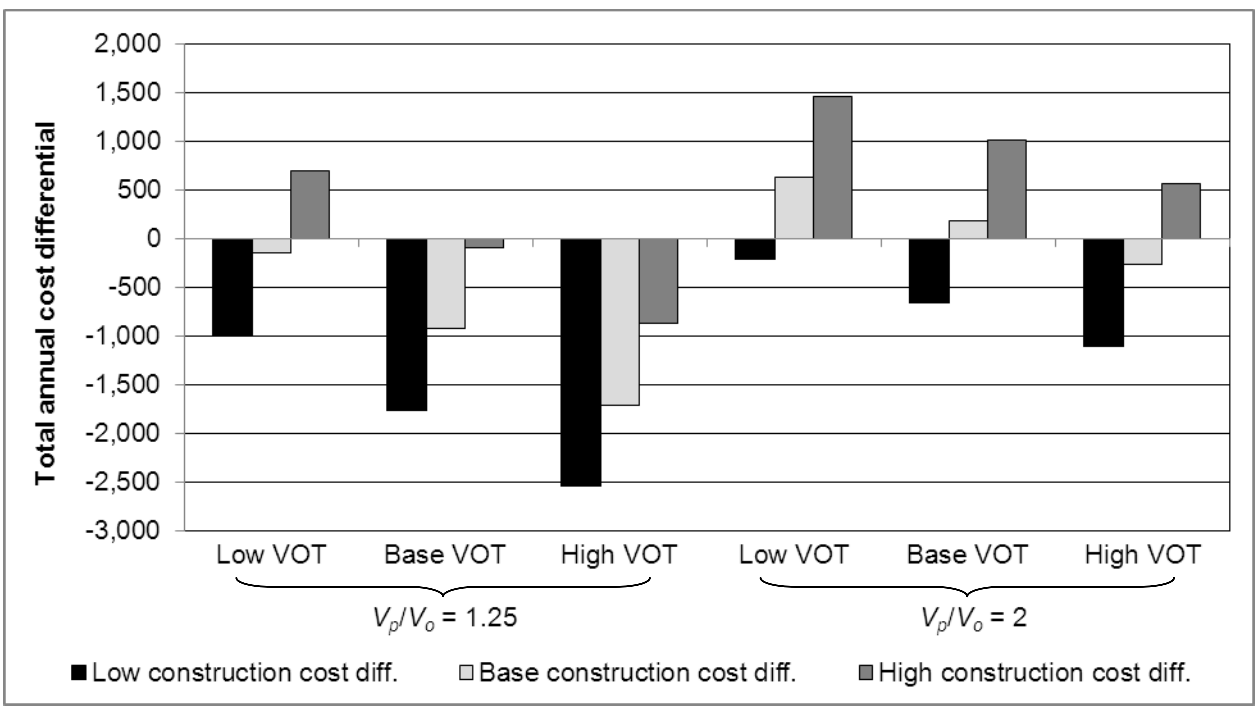
Table 4. Difference in amortized construction costs for equal-capacity expressway and arterial networks (in thousands of 2009 dollars/year)

	Urban area population (1000s)	
Cost differential	200-1,000	>1,000
Base	2,330	1,673
Low (x0.5)	1,165	837
High (x1.5)	3,495	2,510

We can compare the difference in travel-time cost in Table 2 to the difference in construction cost in Table 4 to see which road is favored under different scenarios. Figure 7 illustrates the total annual cost differential (expressway minus arterial) for urban areas of more than 1 million people, where negative numbers indicate that the expressway has lower total annual costs because its travel time savings outweigh its higher construction costs. The expressway's advantage can be clearly seen when values of time are assumed to be high and construction cost differentials are low. However, there are several scenarios where the arterial is favored, especially when the ratio of peak traffic to off-peak traffic is high.

We also ran sensitivity analyses with two different interest rates for amortized construction costs: a lower rate of 5% and a higher rate of 9%. Since higher interest rates widen the gap in construction costs, the arterial is favored under more situations when interest rates are higher.

Figure 7. Total annual cost differential (expressway minus arterial network): V_p=1.05V_K, urban population > 1 million (in thousands of 2009 dollars/year)



Notes: Travel time savings are calculated with the base value of time (VOT) as $10.50/hour, with plus or minus 30% as the high and low values. The base construction cost differential is $1,674 thousand, with plus or minus 50% as the high and low values. Negative numbers indicate that the expressway has lower total annual costs.

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.

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 (Lewis-Evans and Charlton 2006), 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 (Hauer 2000a,b). These are examples of the well-known Peltzman hypothesis (Peltzman 1975), by which safety improvements are partially or fully offset by more aggressive driving.

Statistical studies of the safety effects of such design features are ambiguous.[9] 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 their effects. Another problem is that these comparisons sometimes hold constant the posted speed limit but not the frequency of access points, just the opposite of the conceptual comparison relevant to this paper.

Thus, we think it is an open question whether the "narrow" road designs considered here would in fact reduce safety. It is certainly a potential concern, but one requiring a more complete analysis than we can provide here. Probably the result would depend on factors that vary from

[9] Hadi *et al.* (1995, Table 2) find a statistically significant decrease in injury accident rates with lane and shoulder width for just two of the five types of urban multilane roads studied. Curren (1995, pp. 35-41) finds a substantial and statistically significant increase in accident rates in three corridors where roads were reconfigured with narrower lanes; but a decrease, albeit not statistically significant, in the other two corridors studied. 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. Potts *et al*. (2007) find generally inconsistent or statistically insignificant effects of lane width in cross-sectional comparisons across urban and suburban arterial roads in the Minneapolis-St. Paul and northern Detroit metropolitan areas; in some cases narrower lanes *reduced* injury crashes. Harwood *et al.* (2007, ch. 5) find that wider shoulders do substantially reduce multi-vehicle collisions not associated with intersections or driveways. Noland and Oh (2004) find no effect of average road characteristics on county-loevel accident rates in Illinois. Kweon and Kockelman (2005, Table 3) find that narrower shoulders reduce fatal and injury crash rates, by about 2.6 percent for every foot. Milton and Mannering (1998) find some evidence that narrow lanes and shoulders increase accident frequencies in Washington State.

case to case, especially the speeds chosen by drivers. This suggests a strategy of accompanying such roads with lower speed limits and/or other measures to discourage speeding.

Would speed-control measures be accepted by drivers? Our reading of the guidelines on highway design suggests 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. This is also a conclusion of Ivan *et al.* (2009), who consider in great detail how design features affect chosen speeds.

There is some supporting evidence from Europe for the possibility of implementing effective speed control. Variable speed limits have been used for many years in Germany and the Netherlands, and recently in Denmark, primarily to smooth traffic during the onset of congestion. Such a policy was accompanied by a strong reduction in injury accidents in one German implementation (Mirshahi *et al.* (2007). 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 on-vehicle systems have been carried out in Sweden, the Netherlands, Spain, and the UK (Liu and Tate 2004). Laboratory studies and traffic simulation models show that such measures reduce average speed, speed variation, and lane-changing movements — all of which are known to reduce accident rates — with little or no detrimental effects on traffic flow during congested periods.[10]

If smaller-footprint roads were also automobile-only, a possibility mentioned earlier though not analyzed here, they probably would have further safety advantages. For example, Lord, Middleton, and Whitacre (2005) show that on the New Jersey Turnpike, which has two parallel roadways of which one is car-only, accident rates are higher in the mixed-traffic roadway and trucks are disproportionately involved. 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. On the other hand, Hiselius (2004) finds that trucks *reduce* the number of total accidents on rural roads in Sweden, so the issue is not settled.

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

To complete our brief review of safety, we present two illustrative calculations in order to assess how safety costs might alter to the comparisons of Sections 3 and 4. For this exercise, we assume there are no complementary speed-control policies.

First, we examine the implications for our two alternate expressway designs considered in Section 3. We consider the finding of Bauer *et al.* (2004) that narrower lanes might increase accident rates by around 10 percent — which is almost identical to the reductions computed from Kweon and Kockelman (2005) for the four-foot shoulder reduction that we consider. To this, we apply the estimate by Small and Verhoef (2007, p. 102) of average social costs of accidents for a US urban commuting trip: namely, \$0.16 per vehicle-mile.[11] Thus a 10 percent increase in accident rates over a ten-mile section would imply an increased accident cost of \$0.16 per vehicle per trip, equivalent (at \$10.50/hour value of time) to 0.91 minutes. 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.

Next, we examine implications for our comparison of expressways versus arterials (Section 4). Expressways are safer than arterials, despite their higher speeds. In 2006, the fatality rate per 100 million vehicle-miles traveled on expressways was 0.62, compared to 1.13 for other principal arterials (FHWA, 2006a). Kweon and Kockelman (2005) 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.[12] Using the same accident-cost estimates just mentioned, but restricted to costs due to fatalities and injuries,[13] we can calculate that the social costs for a ten-mile expressway section would then be lower by \$0.2 and \$0.05 per trip due to lower fatal and injury accidents, respectively, compared to an arterial.[14]

---

[11] Updated to 2009 dollars from of \$0.14 in 2005 dollars, based on Small and Verhoef's method of using the average between the growth factors of hourly earnings and of the consumer price index for all urban consumers, all items (US CEA, 2011, Tables B-47 and B-60).

[12] We have computed these numbers 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.

[13] Small and Verhoef (2007, Table 3.4) estimate the average social cost of fatalities and injuries for an urban commuting trip at \$0.087 and \$0.064 in 2009 dollars, respectively, per vehicle-mile (updated from \$0.077 and \$0.057 in 2005 dollars using the same method described in footnote 13).

[14] In this calculation, we adjust for the likelihood that our high-type arterial is safer than an average arterial.: consistent with our treatment of the cost differential in Section 4, we assume its safety disadvantage relative to an

These figures imply that for the example networks in Section 4, accident costs could be higher for the low-footprint design by about $5.4 million and $3.9 million for the lower and higher congestion cases, respectively. In Figure 7, we saw that the arterial's cost advantage was never greater than about $1.5 million under the different scenarios, and so we conclude that accident costs could offset the advantage of an arterial compared to an expressway. Thus we believe that any move to replace expressways with arterials in metropolitan 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 and Discussion

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. Meanwhile, the savings in travel-time costs offered by a network of expressways, compared to an equal-capacity network of high-type unsignalized arterials, is smaller than the amortized value of the extra construction costs incurred under certain conditions (generally, when the value of time is low, the construction cost differential is high, and when the ratio of peak volume to off-peak volume is high). However, if accident costs are included, arterials may no longer have a cost advantage 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. Furthermore, we did not account for fuel consumption and emissions, which depend on speed, grade, etc., for effects on truck routing, or for induced travel, which is important when comparing roads of different capacities; instead, we chose comparisons that minimize the likely impact of these factors, for example by choosing networks of equal total

---

expressway is half that of an average arterial. The figures given in the text are therefore calculated as $0.077x(0.461/2)x10 = $0.18, and $0.057x(0.147/2)x10 = $0.04.

capacity when comparing freeways and arterials. 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 (FHWA, 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 $v_p \le (3{,}400 - 30FFS)$:

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

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

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

In Section 4, we modified the HCM multilane highway speed-flow function for the high-type arterial since the HCM function results in the high-type arterial having a higher speed at capacity than the expressway. For free-flow speeds between 55 and 60 mi/h as in our case, the HCM speed-flow function as seen in Exhibit 21-3 of the HCM is:

$$S_a = FFS - \left[\left(\frac{3}{10} FFS - 10 \right) \left(\frac{v_p - 1{,}400}{28FFS - 880} \right)^{1.31} \right] \quad \text{(A.6)}$$

The high-type arterial's speed at capacity (which we shall denote as S_a^K) can be calculated from the equation above by setting v_p equal to the volume at capacity. Denoting the expressway's speed at capacity as S_e^K, we modify the HCM speed-flow function so that $S_a^K = S_e^K$, essentially by increasing the rate at which speed falls with traffic volume. With this, speeds on the high-type arterial are now obtained using the following formula (with S_a calculated from equation A.6 above):

$$\widetilde{S}_a = \left(\frac{S_a - S_a^K}{FFS - S_a^K} \right) (FFS - S_e^K) + S_e^K \quad \text{(A.7)}$$

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. Because the initial queue limits entry flow to the road's capacity, the initial queue occurs only once at the entry to the road (prior to the first signal) since the traffic volume arriving at each intersection is never greater than the intersection's capacity. As a result, the control delay in this paper consists only of uniform control delay and incremental delay.

The control delay is then calculated for each lane group using equations 16-9, 16-11 and 16-12 of the HCM:

$$d = \frac{0.5C(1-g/C)^2}{1-\left[\min(1,X)(g/C)\right]} \cdot PF + 900T\left[(X-1)+\sqrt{(X-1)^2+\frac{8kIX}{cT}}\right] \tag{A.8}$$

where C is the cycle length, g is effective green time, X is the volume-to-capacity ratio of that lane group, PF is the progression adjustment factor, T is the duration of the analysis period (in hours), k is the incremental delay factor, and I is the upstream filtering factor (equal to 1 since the upstream signal is more than a mile away). The first term in equation A.8 is the uniform control delay while the second term is incremental delay. Note that if we consider the incremental delay for oversaturated conditions, i.e., $X \geq 1$, and drop the part containing k, the incremental delay (converted into hours) turns out to be identical to our paper's average delay for peak travelers in the bottleneck queuing model when oversaturation occurs (equation 1).

The progression adjustment factor, PF, 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 $PF = 1$. k is a calibration factor that depends on whether the signal is actuated or pretimed; it is assumed in this paper that the signals are actuated with snappy intersection operation (unit extension of 2 seconds). With this, k is given by the formula $k = 0.92(X - 0.5) + 0.04$, where $0.04 \le k \le 0.5$.

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.[15] Therefore both lane groups have the same delay, given by (A.8). 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:

$$c_i = s_i \cdot (g_i / C) \tag{A.9}$$

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 τ_{LT} of traffic volume that is making left turns, the total capacity of the road — V_K in our model — is:

$$V_K = (1 - \tau_{LT})^{-1}(c_T + c_{RT}) \tag{A.10}$$

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 $\tau_{LT} = \tau_{RT} = 0.075$.[16]

---

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

[16] 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.

The saturation flow rates needed for (A.9) 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 1-0.15P_{RT}, where P_{RT} is the 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 τ_{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.[17] 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.

[17] 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: $\tau_{RT}[ns_T + s_T(1 -$

In addition to delays at signalized intersections, we must consider queuing at the entrance to the road, where there is no intersection. As described in Section 2, our model assumes that such queuing occurs according to a simple bottleneck model, and that the queue discharges at a rate equal to the capacity of the road. (Capacity is calculated based on an intersection's through movement capacity, as detailed above.) Applying the HCM control delay equation to the road entry, the uniform control delay (the first term in equation A.8) is zero because the road entry has no signal, which gives us $g/C = 1$. The incremental delay due to randomness (the term with k in equation A.8) is negligible because of the large traffic volume, as also shown by Newell (1971, p. 125). The remaining components in equation A.8 plus the HCM initial queue delay (both converted to hours) give precisely the same result as the simple bottleneck model described in the text. That is, the HCM incremental delay excluding the term containing k is the average delay for peak travelers in our bottleneck model when there is oversaturation in the peak period (equation 1), while the HCM initial queue delay is identical to the average delay for non-peak travelers in our bottleneck model when oversaturation occurs (equation 2).

$0.15P_{RT})]/[s_T(1 - 0.15P_{RT})]$. 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 τ_{RT}.

References

AASHTO (2004) *A Policy on Geometric Design of Highways and Streets*, 5th 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.

BLS (2010) *National Compensation Survey: Occupational Earnings in the United States, 2006*, Washington D.C.: Bureau of Labor Statistics. http://www.bls.gov/ncs/ncswage2009.htm

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.

Chicago Area Transportation Study (1998) *1995 Travel Atlas for the Northeastern Illinois Expressway System*, Chicago: Chicago Area Transportation Study.

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.

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

FHWA (2003) *Price Trends for Federal-Aid Highway Construction*, Washington, D.C.: Federal Highway Administration. http://www.fhwa.dot.gov/programadmin/pt2003q1.pdf

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

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/

FHWA (2011), *National Highway Construction Cost Index*, Washing D.C.: Federal Highway Administration. http://www.fhwa.dot.gov/policyinformation/nhcci.cfm

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

Hu, Pat S., and Timothy R. Reuscher (2004) *Summary of Travel Trends: 2001 National Household Travel Survey*. Washington, D.C.: US Federal Highway Administration, December. http://nhts.ornl.gov/2001/pub/STT.pdf

Ivan, John N., Norman W. Garrick, and Gilbert Hanson (2009) *Designing Roads that Guide Drivers to Choose Safer Speeds*, Report No. JHR 09-321. Rocky Hill, Connecticut: Connecticut Department of Transportation, November.

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/

Newell, Gordon F. (1971) *Applications of Queuing Theory*, London: Chapman and Hall.

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.

Office of Management and Budget (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.

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.

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 CEA (2011) "Annual report of the Council of Economic Advisers," in *Economic Report of the President*, Washington, DC: US Government Printing Office. http://www.gpoaccess.gov/eop/2011/pdf/ERP-2011.pdf

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.