UC Irvine
UC Irvine Previously Published Works

Title
Optimizing road capacity and type

Permalink
https://escholarship.org/uc/item/7rp3j4x3

Journal
Economics of Transportation, 3(2)

ISSN
2212-0122

Authors
Small, Kenneth A
Ng, Chen Feng

Publication Date
2014-06-01

DOI
10.1016/j.ecotra.2014.02.001

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Optimizing Road Capacity and Type

Kenneth A. Small
Chen Feng Ng

Feb. 13, 2014

Small: Dept. of Economics, University of California at Irvine, Irvine, CA 92612-5100, USA; ksmall@uci.edu

Ng: Dept. of Economics, California State University at Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840-4607, USA; chen.ng@csulb.edu

Keywords: Capacity; free-flow speed; highway design; optimal highway investment; congestion
JEL codes: L91, R42

Abstract

We extend the traditional road investment model, with its focus on capacity and congestion as measures of capital and its utilization, to include free-flow speed as another dimension of capital. This has practical importance because one can view free-flow speed as a continuous proxy for road type (e.g. freeway, arterial, urban street). We derive conditions for optimal investment in capacity and free-flow speed, and analyze the optimal balance between the two. We then estimate cost functions for capital and user costs and apply the resulting model using parameters representing large US urban areas. We show that providing high free-flow speed may be quite expensive, and there is sometimes a tradeoff between it and capacity. We find suggestive evidence that representative freeways in many large urban areas provide too high a free-flow speed relative to capacity, thus making the case for reexamination of typical design practice.

Acknowledgment

We are grateful for comments by Richard Arnott, Amihai Glazer, Jan Owen Jansson, Marvin Kraus, Odd Larsen, Robin Lindsey, Robert Poole, Kurt Van Dender, anonymous referees, and seminar participants at the University of Tokyo and the 2013 Kuhmo Nectar Conference on Transportation Economics. All responsibility for the content remains with the authors.

Optimizing Road Capacity and Type

Kenneth A. Small
Chen Feng Ng

1. Introduction

The economic analysis of congestion and investment in road capacity is well developed. The research literature contains an abundance of optimality conditions, implications for pricing, and policy implications including such practical matters as second-best pricing, investment under conditions of suboptimal pricing, and financial balance between pricing revenues and investment costs.[1] In such analyses, roads are generally taken to be sufficiently characterized by a single dimension, capacity, with other issues such as safety or aesthetic ride quality dealt with as separate side issues.[2] In part, this emphasis is justified by the apparent dominance of congestion among the costs of urban road trips.[3]

Yet some of the most serious practical issues in road policy involve other aspects of roads such as their safety, environmental impacts, aesthetics, and impacts on neighborhoods and other considerations of urban design. As a result, passionate debates arise about not only the amount of road space to provide, but its type. In particular, the penetration of dense urban development by high-speed and high-capacity expressways has always been controversial.

Transportation economists have had less to say about these latter issues, and a major reason is the single capital dimension in the standard economic models of road investment. Yet it is entirely possible to build very different looking urban road networks of equal capacities, one using high-speed freeways and another using well-engineered arterials. These design tradeoffs require other measures of road capital than capacity.

The goal of this paper is to provide an expanded investment model that lends itself to analyzing such issues, by including free-flow speed as an additional design variable describing

[1] Examples include Mohring and Harwitz (1962), Strotz (1965), Keeler and Small (1977), and Jansson (1984). For reviews see Lindsey and Verhoef (2000) and Small and Verhoef (2007, ch. 5).

[2] In three cases these other road characteristics are explicitly modeled, either as a type of scale economy (Jansson 1984, ch. 10) or as a quality variable (Larsen 1993, Walters 1968). Walters acknowledges that capacity and road quality may vary independently but suggests that "as a very rough approximation it may … be sensible to treat the roads as giving a joint-product with *rigid* proportions." (pp. 35-36, italics in original).

[3] Small and Verhoef (2007), sect. 3.4.6.

road capital. This is of course only a first step toward a more comprehensive goal, in which the planner simultaneously optimizes the many design elements making up road investment (some of which we enumerate in our empirical section), and does so for each road in a network serving diverse trips.[4] While not every issue of interest can be captured with our addition of just one new investment dimension, the advantages of tractability and transparency make this an attractive way to begin bringing the analysis of road types into mainstream transportation economics.

We start by developing the theoretical investment model with a long-run total cost function, consisting of capital costs and user costs, with capacity and free-flow speed as design variables. The first-order conditions of the model lead to the familiar criterion for incremental investment in capacity, and a new criterion for incremental investment in free-flow speed. Combining these criteria gives us an "investment balance condition" that can be used to examine under what conditions a given road is well balanced between these two dimensions: i.e., when does a given road design provide too high or low a free-flow speed relative to its capacity?

To implement the model, we use empirical data to estimate both components of the total cost function. We estimate the capital-cost function using data on construction costs of various road types along with their free-flow speeds and capacities. We estimate the user-cost function from information about speeds and flows of different road types, differentiated by free-flow speed,[5] which we supplement with a queuing analysis to account for situations where input flow exceeds capacity. We then apply the estimation results to examine the investment balance condition for 24 standard road types under hypothetical conditions, and for representative freeways and arterials for a sample of US urban areas under actual conditions.

While our goal here is not primarily policy analysis, the model does permit another look at a question considered by Ng and Small (2012). Given that many high-speed urban expressways operate under severe congestion for several hours each day, is the extra expense of providing such high-speed service under more moderate traffic justified? In the extreme case where all traffic occurred during a peak period impacted by queues behind fixed-capacity bottlenecks, there would be no advantage to high free-flow speed. In more realistic cases, there

[4] A start on such a program is made by Ben-Akiva et al. (1985), who similarly argue for minimizing total costs as a design objective and apply this approach to geometric design of a climbing lane for heavy vehicles.

[5] Such information is compiled in the Highway Capacity Manual (Transportation Research Board 2000) from decades of engineering research.

are tradeoffs involving the duration of peak periods and the relative traffic volumes in peak and off-peak periods. Our earlier paper considers this question by comparing a few specific road types chosen to illustrate the tradeoff between free-flow speed and capacity, or between free-flow speed and construction cost. Here, we develop a more general model of road investment where both capital costs and user costs can vary depending on free-flow speed and capacity, each of which lies along a continuum.

We do find some evidence that typical freeways in large urban areas are over-designed for free-flow speed at the expense of capacity. This arises largely from the finding that the cost elasticity for increasing free-flow speed is, on average, three times that for expanding capacity (roughly 1.2 vs. 0.4); as a result even modest amounts of congestion favor incremental investments in capacity relative to free-flow speed. While the optimal road configuration is very case-specific, we can state a more general policy conclusion: road design needs to allow for variety and flexibility, rather than being constrained to meet a predetermined set of standards such as those for US Interstate Highways. There are probably many situations where urban areas are well served by parkways, high-type arterials, or urban streets with well-engineered intersections as a means of carrying large traffic flows efficiently.

2. Long-run cost functions with two dimensions of infrastructure

Total costs of road travel in our model consist of amortized capital cost and user costs. We adopt simple formulations for each, in order to emphasize what is new in this paper, namely the role of free-flow speed as a design variable. Thus, for example, we ignore road maintenance costs (assuming they would not affect design), accident costs (as there is mixed evidence in the literature regarding the impact of design speed on accident rates), other user costs aside from time (assuming they are proportional to vehicle flow and therefore also do not affect design), and environmental costs (which are best dealt with using other tools).[6] We also ignore capacity

[6] Alam and Kall (2005) calculate that pavement resurfacing costs per lane-mile are higher for freeways than for arterials, not accounting for the fact that roads with higher traffic volumes (like freeways) also tend to be resurfaced more often. Average maintenance costs therefore appear to be correlated with average construction costs, and we believe that including maintenance costs would not change our results significantly. Meanwhile, as discussed in Ng and Small (2012), some of the design features that could result in lower free-flow speeds (like narrower lanes or a lower type of road such as a highway instead of a freeway) do not necessarily lead to higher accident rates, especially if the roads are accompanied by lower speed limits.

fluctuations due to accidents or weather, and the prospect of automated vehicles changing the speed-flow relationships.

Annualized capital cost is composed of initial costs of structures and land, each amortized at a constant rate over its lifetime. These costs depend on road design via the variables measuring capacity and free-flow speed:

$$\rho(V_K, S_F) = \frac{r}{1-e^{-r\Lambda}}\Gamma(V_K, S_F) + rA(V_K, S_F) \tag{1}$$

where V_K and S_F are design capacity and free-flow design speed, respectively, Γ is construction cost, A is right-of-way acquisition cost, r is the interest rate, and Λ is the road life in years, i.e. the time after which the structures and improvements (but not the land) have lost all their value. We assume that Γ and A are increasing in both V_K and S_F. This formulation assumes the annualized cost is constant over the road's lifetime.

Total user cost U_t per unit time during a discrete time interval t consists solely of time costs measured at a value of time, α, which for simplicity we take to be constant. User time depends both on free-flow speed and on congestion, the latter via the volume-capacity ratio:

$$U_t(V_t; V_K, S_F) = V_t c_t = V_t \frac{\alpha}{S_t\left(\frac{V_t}{V_K}, S_F\right)} \tag{2}$$

where V_t is traffic volume during time interval t, c_t is average user time cost, and S_t is average speed. The latter is assumed to be increasing in S_F, and to be decreasing in volume-capacity ratio.

The short-run total cost function, including agency costs, is therefore:

$$C(V; V_K, S_F) = \rho(V_K, S_F) + \sum_t q_t U_t(V_t; V_K, S_F) \tag{3}$$

where q_t is the duration of time interval t and V={V_t} is the time pattern of vehicle flows.

The long-run cost function is obtained by choosing the design variables so as to minimize short-run total cost:[7]

$$\tilde{C}(V) = \min_{V_K, S_F} C(V; V_K, S_F)$$
$$= \min_{V_K, S_F} \left[\rho(V_K, S_F) + \alpha \sum_t \frac{q_t V_t}{S_t\left(\frac{V_t}{V_K}, S_F\right)} \right].$$

The conditions for this minimization constitute the investment rules governing capacity and free-flow speed. Assuming interior solutions, they are:

$$\frac{\partial \rho}{\partial V_K} = -\sum_t q_t V_t \frac{\partial c_t}{\partial V_K} \tag{4a}$$

$$\frac{\partial \rho}{\partial S_F} = -\sum_t q_t V_t \frac{\partial c_t}{\partial S_F} \tag{4b}$$

which state that each type of investment should be undertaken to the point where the resulting marginal saving in user cost equals its incremental annualized capital cost.[8] Furthermore, in each

---

[7] The mathematical optimization conditions remain valid even though the design variables are in fact determined as functions of more primitive design options, such as lane width, number of lanes, shoulder widths, etc. Essentially we have reduced the many-dimensional problem to a two-dimensional one, assuming the planner knows how to find the best combination of those primitive design options to achieve a given pair of values for V_K and S_F. This assumes the underlying design options do not create perfect correlation between the two design variables, which they do not as we show in our empirical section.

[8] The second-order optimization conditions for cost minimization require that the Hessian matrix of second derivatives of (3) be positive definite. This means that there are scale diseconomies in the capital cost of providing both V_K and S_F , and/or in the way those two variables affect user costs. Numerically, we find this condition to be met for the road types in this paper when the volume-capacity ratio is moderate or high (in general, when V_t/V_K is about 0.4 and greater for arterials, and 0.7 and greater for highways and freeways). These scale diseconomies are driven by the empirical result that c_t is highly convex in volume-capacity ratio whenever there is any substantial congestion, offsetting the scale economies in the capital cost of providing capacity. This finding appears consistent with the informed speculation by Walters (1968, pp. 36-38) that improvements in road quality are often justified by user-cost savings alone for low-capacity roads, but less likely so for high-capacity roads.

the left- and right-hand sides state, respectively, the incremental costs and incremental benefits of investment.

The first of these investment rules is standard.[9] The second is new to this paper, but obviously follows the same logic. Note that the solutions to these equations will depend strongly on the construction cost function and the level of traffic, leading to very different results in different locations—as is typical of any investment model.

Equations (4a) and (4b) may be simplified by taking advantage of our assumption that user cost is a function of volume and capacity only through their ratio, an assumption which also underlies the analysis of self-financing by Mohring and Harwitz (1962, pp. 84-87).[10] This assumption implies that

$$V_K \frac{\partial c_t}{\partial V_K} = -V_t \frac{\partial c_t}{\partial V_t}$$

from which we can rewrite (4a) and (4b) in elasticity terms as:

$$\rho \varepsilon_{\rho,VK} = \sum_t q_t V_t c_t \left(\varepsilon_{c,V}\right)_t = \sum_t q_t V_t \cdot (mecc)_t \equiv \tilde{R} \tag{5a}$$

$$\rho \varepsilon_{\rho,SF} = \sum_t q_t V_t c_t \cdot \left(\varepsilon_{S,SF}\right)_t \tag{5b}$$

where $(mecc)_t \equiv V_t \bullet (\partial c_t / \partial V_t)$ is the marginal external congestion cost of a trip, $\varepsilon_{\rho,VK}$ and $\varepsilon_{\rho,SF}$ are the elasticities of annualized capital cost with respect to capacity and free-flow speed, respectively, $\varepsilon_{c,V}$ is the elasticity of average travel cost with respect to traffic volume, and $\varepsilon_{S,SF}$ is the elasticity of speed with respect to S_F. (These last two elasticities may vary by time period.) The quantity $\tilde{R}$ is imputed revenues from a hypothetical congestion toll set equal to $mecc_t$ in

---

[9] This investment rule is given in various forms by Mohring and Harwitz (1962, p. 84), Strotz (1965, eq. 1.17), and Keeler and Small (1977), eq (5). See Small and Verhoef (2007, eq. 5.3) for a concise derivation.

[10] This assumption is sometimes described as constant returns to scale in congestion technology: see Small and Verhoef (2007, p. 165).

each period when traffic is fixed at V_t.[11] Therefore (5a) expresses the self-financing theorem, which states that annual revenues from such a toll would equal annualized capital costs times the cost elasticity of capital cost with respect to V_K. Equation (5b) has no comparable interpretation, since there is no efficiency reason to impose a toll for free-flow speed.

The quantities in equations (5a) and (5b) are likely to be quite case-specific, making it difficult to draw general conclusions about the benefits and costs of each type of investment. However, we are more confident in ratios of the two sides of these equations, which reflect the relative costs of the two kinds of investment and the relative cost savings they provide to users. Therefore, we primarily consider what we call "investment balance," defined by dividing (5a) by (5b):

$$\frac{\varepsilon_{\rho,VK}}{\varepsilon_{\rho,SF}} = \frac{\sum_t q_t V_t \cdot (mecc)_t}{\sum_t q_t V_t c_t \cdot \left(\varepsilon_{S,SF}\right)_t} \,. \tag{5c}$$

Consider the implications for any given ratio of elasticities forming the left-hand side of (5c). If congestion is large, so that *mecc* greatly exceeds $c \bullet \varepsilon_{S,SF}$ for a large portion of the time, then investment in capacity will tend to be favored relative to that in free-flow speed since the right-hand side of (5c) will tend to exceed its left-hand side. On the other hand, if peak traffic congestion is not severe and off-peak travel volume is extensive, the ratio on the right-hand side will tend to be small, favoring investment in free-flow speed. In what follows, we refer to the left-hand side (LHS) of equation (5c) as the "ratio of construction cost elasticities," and the right-hand side (RHS) as the "ratio of marginal user costs" (i.e., the ratio of incremental user-cost savings from expanding capacity versus increasing free-flow speed). Our measure of "investment balance" is LHS – RHS; a positive number means that marginal investment in S_F is favored relative to that in V_K.

Intuition is aided by an example. First, suppose travel time is given by the free-flow travel time plus a queuing time applicable only if capacity is exceeded:

[11] As is well known, such a toll can be derived by maximizing the difference between consumers' valuation of their travel (the area under their inverse demand curve) and total costs. See Keeler and Small (1977).

$$\frac{1}{S}=\frac{1}{S_F}+\max\left[\frac{q_t}{2}\cdot\left(\frac{V_t}{V_K}-1\right),0\right]. \tag{6}$$

This piecewise-linear cost function describes the time-averaged user cost for a deterministic bottleneck of constant capacity, assuming there is no queue at the beginning of the time period, because the average queuing delay experienced is proportional to the length of time when bottleneck capacity is exceeded. We then have $mecc=\alpha\bullet[(1/S)-(1/S_F)]$, $\varepsilon_{S,SF} = S/S_F$, and the first-order investment conditions are:

$$\varepsilon_{\rho,VK}=\frac{U^g}{\rho}; \qquad \varepsilon_{\rho,SF}=\frac{U^0}{\rho} \qquad => \qquad \frac{\varepsilon_{\rho,VK}}{\varepsilon_{\rho,SF}}=\frac{U^g}{U^0} \tag{7}$$

where total user cost U over all time periods has been divided into that due to free-flow travel time, $U^0\equiv\alpha\sum_t q_tV_t/S_F$, and that due to congestion, $U^g\equiv U\text{-}U^0$. This example makes clear that a marginal increase in capacity is valuable when user costs of congestion (U^g) are high, whereas an increase in free-flow speed is valuable when user costs of free-flow travel (U^0) are high.[12]

Returning to the more general case, equations (5) can be used to assess current or proposed planning for road capacity and type. A hypothesis motivating this paper is that current planning guidelines for urban areas may place too much emphasis on free-flow speed relative to capacity. This could take the form either of designing a given type of roadway for unnecessarily high speeds, or of choosing a higher type of roadway than necessary. Empirical measurements suggesting that the cost ratio on the right-hand side of (5c) exceeds the elasticity ratio on its left-hand side—i.e., investment balance is negative—would provide evidence for this hypothesis.

Equations (5) apply to a first-best optimum. Each is valid conditional on V, but if V is sensitive to investment and one or more optimality conditions does not hold, then the other does not necessarily describe a second-best situation. In Section 5, we will apply our theory to

---

[12] Another example is when time spent in congestion is modeled, as is common, as a power function of the volume-capacity ratio with power b. Then $mecc=\alpha b\bullet[(1/S)-(1/S_F)]$ and $\varepsilon_{S,SF}=1$; the optimization conditions are $\varepsilon_{\rho,VK}=bU^g/\rho$ and $\varepsilon_{\rho,SF}=U/\rho$. In this case cost added by congestion is affected by S_F, which is why the numerator of the second equation includes total user cost U and not just the uncongested portion U^0 as it did in the other example.

existing cities, which are subject to strong underpricing of congestion, i.e. (5a) does not hold; so we will need to modify (5c) accordingly if we wish to make second-best policy statements.

3. Empirical estimation of cost functions

3.1 Data for costs, free-flow speeds, and capacities

We wish to estimate construction costs as a function of capacity and free-flow speed, while holding constant other factors such as terrain, climate, and input prices. Since we are more interested in the relative costs of different types of roads than their absolute costs, we are not too concerned about whether we have representative values for those other factors, but do want detailed differences among road types. Such data are provided by the Specifications and Estimates Office of the Florida Department of Transportation (FDOT). These data contain estimated quantities and prices of inputs needed for various types of roads in urban areas, while holding other factors constant.

The basic data, shown in Table 1, tell us about the tradeoffs among alternative road designs. For example, as we shall see shortly, a 4-lane divided urban street has the same free-flow speed as an undivided 5-lane urban street with a center turn lane, but the former costs more and has higher capacity. Meanwhile a 4-lane Interstate offers greater free-flow speed but lower capacity than a 6-lane multilane highway, with only a small cost difference.

Table 1. FDOT cost estimates for urban areas (in 2011 prices)

Description	No. lanes	Bike lane (width)	Median (width)	Shoulders (inside & outside)	Cost per mile (mill. $)
Undivided arterial	2	4 ft	---	---	4.179
Undivided arterial with center lane	3	4 ft	---	---	4.769
Undivided arterial	4	4 ft	---	---	5.132
Undivided arterial with center lane	5	4 ft	---	---	5.814
Divided arterial	4	4 ft	22 ft	---	7.123
Divided arterial	6	4 ft	22 ft	---	7.986
Divided Interstate, closed median with barrier wall	4	---	22 ft	10 ft	8.875
Divided Interstate, closed median with barrier wall	6	---	22 ft	10 ft	9.858

Source: Statewide cost estimates published in January 2012 by the Specifications and Estimates Office of the Florida Department of Transportation (http://www.dot.state.fl.us/planning/policy/costs/). The cost estimates are calculated based on the quantities of different materials needed for these roads and are updated every year using statewide average price estimates. See Appendix B for more detail.

These cost estimates are even more useful because they contain detailed information on individual components such as embankment, pavement, pipe culverts, lighting, etc. This additional information enables us to double our sample size by estimating, for each road type, the cost of an otherwise identical road but with 11-foot lanes instead of the default lane width of 12 feet. This is done by reducing the relevant costs (embankment, stabilization and pavement costs) proportionately, while keeping other costs (such as the costs of pipe culverts, curbs and gutters, pavement markings, lighting and signage) constant. Since 11-foot lanes are recognized in the Highway Capacity Manual (HCM) (Transportation Research Board 2000), we will be able to measure the deterioration of service quality and capacity that accompanies the lower costs and, as we shall see, these two dimensions are not degraded proportionally.

In order to calculate free-flow speeds and capacities for each road type, we use the 2000 Highway Capacity Manual, supplemented where necessary by the FDOT road descriptions and HCM default values; see Appendix A for other assumptions and the equations.[13] The HCM has separate procedures for freeways, urban streets, and “highways” (which have design standards

[13] Although there is a newer edition of the HCM (the 2010 version), we use the 2000 version so that the results in this paper are consistent with those presented in Ng and Small (2012).

between those of freeways and urban streets).[14] We are therefore able to further expand our data set by assuming that FDOT's "arterial" can be either an urban street with traffic signals or a highway (except we assume only an urban street can have a center lane). We assume that highways have grade-separated intersections at all major crossings and there are no signals but like urban streets, there are some at-grade access points (e.g., driveways). It is further assumed that urban streets have one signal per mile while highways and freeways have an interchange with an urban street every two miles. We use the cost estimates for traffic signals and interchanges included in the FDOT dataset and add them to the costs shown in Table 1 (see Appendix B for more detail).

Urban streets require several further assumptions. We assume they have limited parking and little pedestrian activity. We assign speed limits of 45 mi/h and 40 mi/h for the roads with 12-foot lanes and 11-foot lanes, respectively (since free-flow speed depends on, though is not equal to, the speed limit). We also must make assumptions about the number of turn lanes and signal phasing for left-turn lanes (see Appendix A).[15] For each assumed turn-lane and signal configuration, we calculate the saturation flow rate, i.e., the highest flow rate that can pass through a signalized intersection while the light is green, and from that we calculate capacity following the HCM.

The assumptions just described lead to 24 road types, each with its unique cost, capacity, and free-flow speed. From these 24 observations, summarized in Table 2, we fit function $\Gamma(V_K,S_F)$ describing initial construction cost.

[14] In deference to this distinction, we use "road" as a general term encompassing all three types, so as to avoid the ambiguity of the term "highway" that exists in the HCM (even in its title) between the general or specific meaning of "highway."

[15] Signal phasing means the types of turns permitted on successive parts of a complete cycle for a traffic signal. The two categories of phasing of primary concern to us are permitted versus protected left turns: "permitted" means left turns are allowed whenever the light is green and there is a gap in oncoming traffic, whereas "protected" means left turns are allowed only with a green arrow during which oncoming traffic is stopped with a red signal.

Table 2. Road types and construction cost per mile

No. of lanes (two-directional)	Road type	Lane width (feet)	Unimpeded speed (mi/h)	Free-flow speed (mi/h)	Two-directional capacity (veh/h)	Road cost per mile (thousands of $)	Signal/interchange cost (thousands of $)	Total cost per mile (thousands of $)
2 lanes, undivided	Urban street	12	42.1	35.8	1,277.6	4,179	155	4,334
		11	40.2	34.4	1,245.1	4,033	155	4,188
	Two-lane highway	12	52.5	52.5	3,112.4	4,179	5,581	9,760
		11	47.1	47.1	3,112.4	4,033	5,385	9,417
3 lanes, ctr turn lane	Urban street	12	42.1	35.8	1,637.0	4,769	155	4,924
		11	40.2	34.4	1,582.4	4,581	155	4,736
4 lanes, undivided	Urban street	12	43.1	36.5	1,930.2	5,132	195	5,328
		11	41.2	35.1	1,891.9	4,909	195	5,104
	Multilane highway	12	51.8	51.8	7,306.1	5,132	6,853	11,985
		11	49.9	49.9	7,169.7	4,909	6,555	11,463
5 lanes, ctr turn lane	Urban street	12	43.1	36.5	3,273.1	5,814	195	6,009
		11	41.2	35.1	3,164.0	5,537	195	5,732
4 lanes, divided	Urban street	12	43.1	36.5	3,745.7	7,123	195	7,318
		11	41.2	35.1	3,620.9	6,854	195	7,050
	Multilane highway	12	53.4	53.4	7,421.0	7,123	9,511	16,634
		11	51.5	51.5	7,284.6	6,854	9,152	16,007
	Freeway	12	65.5	65.5	8,455.0	8,875	11,850	20,725
		11	63.6	63.6	8,386.8	8,353	11,153	19,506
6 lanes, divided	Urban street	12	43.5	36.8	5,618.6	7,986	236	8,222
		11	41.6	35.4	5,431.3	7,639	236	7,876
	Multilane highway	12	53.4	53.4	11,131.6	7,986	10,664	18,651
		11	51.5	51.5	10,926.9	7,639	10,201	17,840
	Freeway	12	67.0	67.0	12,763.3	9,858	13,163	23,020
		11	65.1	65.1	12,661.0	9,215	12,304	21,519

Note: We use "free-flow speed" to designate the speed at very low traffic levels, as does Schrank et al. (2012b). The HCM defines it the same way for freeways and highways. But for urban streets, the HCM defines free-flow speed to exclude the effects of "control delay", which is the delay caused at intersections by stopping and/or waiting behind other stopped vehicles while they start up and proceed through the intersection; here we call this the "unimpeded speed." Formulas for calculating both unimpeded speed and control delay are provided by Zegeer et al. (2008) and the HCM (see Appendix A), and used here to compute "free-flow speed" as well as, in the next section, speed as a function of traffic volume. Road capacities and speeds are calculated using the HCM default assumption that the percentage of heavy vehicles is 5%.

These estimates imply construction costs per lane-mile, for 12-foot lanes, of roughly $3.8–5.2 million for freeways and $1.3–2.5 million for urban streets, with multilane highways in between. As a comparison, Schrank et al. (2012a) estimate that new construction can cost between $5-20 million per lane-mile for freeways, and around $1.5 million for "major surface streets," although their numbers likely include land acquisition costs.

3.2 Estimation of capital cost function

We use a Cobb-Douglas function to estimate the relationship between construction cost per mile (denoted by Γ, measured in thousands of dollars), free-flow speed (S_F), and capacity (V_K), with the right-hand-side variables as ratios to their sample means:

$$\ln \Gamma = \beta_0 + \beta_1 \ln(S_F / \overline{S_F}) + \beta_2 \ln(V_K / \overline{V_K}) + \xi \tag{8}$$

where ξ is the error term. The sample means for free-flow speed and capacity are 45.80 mi/h and 5,589 veh/h, respectively. We also estimated a translog cost function with second-order terms as seen in Table 3.

The regression results, using ordinary least squares on 24 observations, are shown in Table 3. Since none of the second-order terms are statistically significant (at a ten-percent level), we prefer the Cobb-Douglas specification (Model 1). Using that specification, the implied elasticities of construction cost with respect to free-flow-speed and capacity are

$$\varepsilon_{\Gamma,SF} = \beta_1; \quad \varepsilon_{\Gamma,VK} = \beta_2 .$$

As indicated by the first two coefficients of Model 1, these elasticities are 1.22 and 0.41, respectively. Thus increasing capacity—for example, by building more lanes of a given road type—is subject to strong scale economies, a finding consistent with evidence in Meyer et al. (1965) and Kraus (1981).[16] What is new here, and potentially important, is the finding of scale *dis*economies with respect to free-flow speed. Our estimate suggests that increasing free-flow speed is quite expensive, even holding capacity constant.

[16] Kraus finds scale economies are substantially reduced, though not eliminated, by considering the effects produced by the high cost of enlarging intersections as an entire network of roads is expanded. Such costs are not considered here, at least not explicitly.

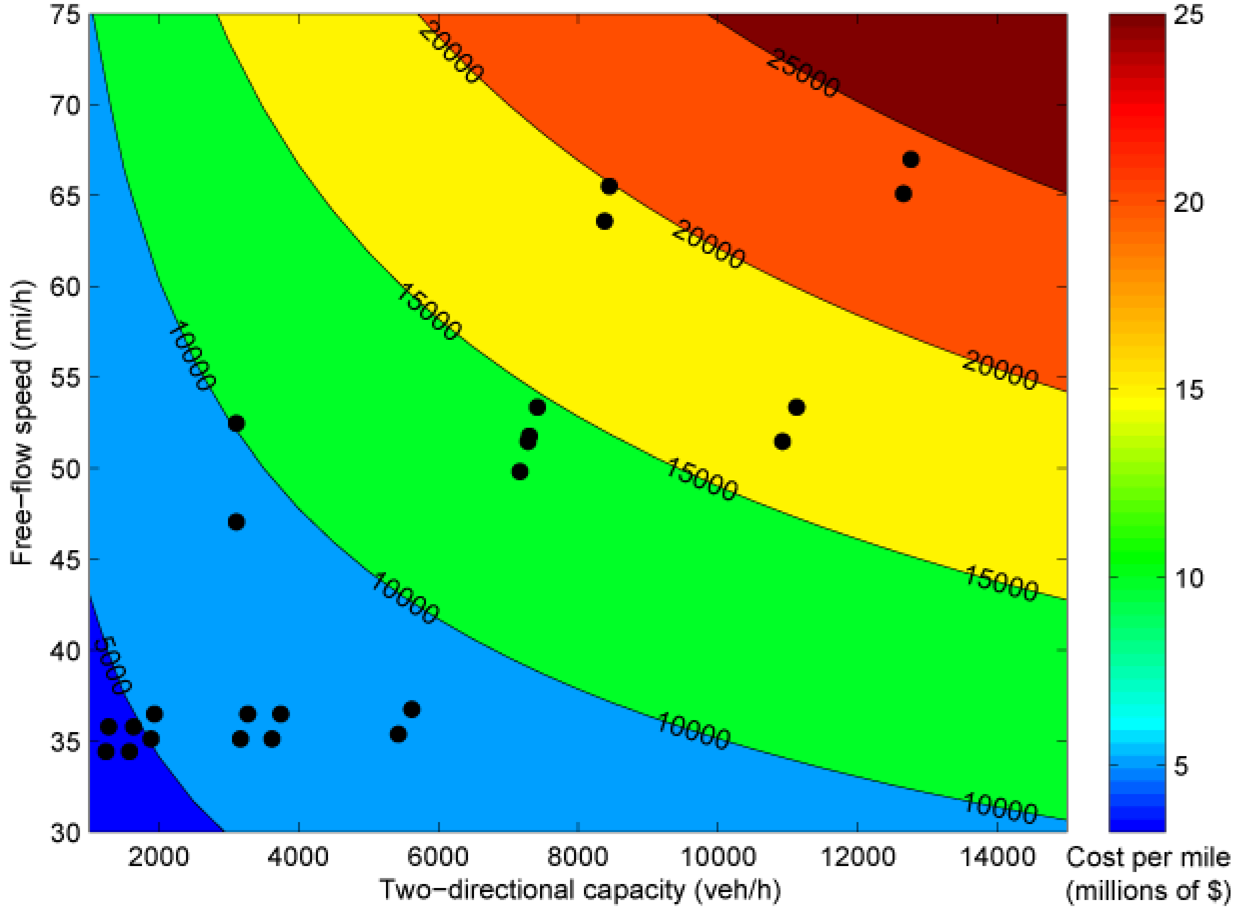
Table 3. Construction cost regression results with ln Γ as the dependent variable

Variables	Model 1: Cobb-Douglas	Model 2: Translog
$\ln(S_F / \overline{S_F})$	1.2170***	1.1961***
	(0.117)	(0.148)
$\ln(V_K / \overline{V_K})$	0.4090***	0.4298***
	(0.038)	(0.066)
$0.5[\ln(S_F / \overline{S_F})]^2$		-0.1003
		(1.742)
$0.5[\ln(V_K / \overline{V_K})]^2$		0.0943
		(0.211)
$[\ln(S_F / \overline{S_F})][\ln(V_K / \overline{V_K})]$		-0.1763
		(0.504)
Constant	9.2950***	9.3003***
	(0.018)	(0.037)
Observations	24	24
R-squared	0.976	0.982

Note: Standard errors in parentheses.
***, ** and * indicate statistical significance at the 1, 5 and 10 percent levels, respectively.

The regression results can be used to predict construction costs for a range of free-flow speeds and capacities. Figure 1 shows these predicted costs as well as a scatter plot of the actual 24 data points. It provides an illustration of how construction costs increase as both free-flow speed and capacity increase.

Figure 1. Contour plot of predicted costs using Cobb-Douglas coefficient estimates and scatterplot (in black) of observed data points



To estimate the annualized capital cost of building a road, we combine the construction costs (Γ) from equation (8) with some assumptions on right-of-way acquisition cost (*A*), the interest rate (*r*), and the road life in years (Λ), in order to calculate equation (1). Based on Ng and Small (2012), variable *A* typically ranges from about 3 to 6 percent of total capital cost for urban areas with a population of 0.2 to 1 million people, and is about 18.3 percent for urban areas with one million people or more.[17] Denoting these percentages as *x* (expressed as a decimal), we can express the right-of-way acquisition cost as a fraction of construction cost: $A = \Gamma \cdot [x/(1-x)]$. The annualized capital cost per mile from equation (1) can therefore be rewritten as:

[17] These statements from Ng and Small (2012) are in turn based on cited figures from Alam and Ye (2003) and Alam and Kall (2005).

$$\rho(V_K, S_F) = \left(\frac{r}{1-e^{-r\Lambda}} + \frac{rx}{1-x}\right)\Gamma(V_K, S_F)\,. \tag{9}$$

Given exogenous values of r, Λ and x, the factor in parentheses on the right-hand side of (9) is a constant, which we denote as κ. Taking the natural logarithm of equation (9) and substituting in equation (8) (without the error term) yields:

$$\ln \rho(V_K, S_F) = \ln \kappa + \beta_0 + \beta_1 \ln(S_F / \overline{S_F}) + \beta_2 \ln(V_K / \overline{V_K}) \tag{10}$$

Therefore the capital cost elasticities are the same as those from the construction cost function.

4. Speeds and travel times

To determine travel times on the road types described in the previous section, we consider four factors: (1) free-flow-speed; (2) slower speeds, based on the HCM speed-flow curves, when traffic flow increases but is still below capacity; (3) control delay due to traffic signals, applicable only to urban streets; and (4) congestion delay from queuing when demand exceeds capacity. The first three components are based on the HCM procedures described in Appendix A. The last arises from the considerable evidence that much of the delay on roads is due to local queuing behind bottlenecks arising from capacity variation, entry ramps, or exit ramps (Daganzo et al. 1999, Small and Chu 2003).

The fourth component of travel time, congestion delay, is based on the bottleneck queuing model, which with some minor modifications is the same as that in Ng and Small (2012) as well as in the first example in Section 2. We assume that the bottleneck occurs at the entry to the road, and there are two time periods for one-directional traffic: a "peak" period of duration P (in hours) with constant demand V_p, and an "off-peak" period of duration F with constant demand V_o. A queue (assumed to have zero physical length) builds up if demand exceeds capacity V_K, and the average queuing delay experienced by a peak traveler is $(P/2)\bullet[(V_p/V_K) - 1]$. The model of Ng and Small assumes that the queue gradually discharges when demand falls below capacity, and so if $V_o<V_K<V_p$, off-peak travelers typically experience some queuing delay. However, this would be inconsistent with the assumptions of the theoretical model in Section 2

where it is assumed that travelers in one time period do not affect the travel times of travelers in other time periods (i.e., user cost, c_t, depends only on traffic conditions in time period t and not on those in any other time period). Therefore, when calculating travel times in this section we simplify by ignoring the queuing delay experienced by some off-peak travelers; thus off-peak travel times are underestimated when peak volumes exceed capacity.[18]

We assume that the road is 10 miles in length, which is close to the average vehicle trip length of 9.72 miles reported in the National Household Travel Survey (Santos et al. 2011, Table 3). The durations of the time periods are assumed to be P = 4 hours and F = 12 hours, respectively. (Under our assumptions the value of F does not affect travel time, but it is used later when calculating aggregate travel times for all travelers.)

Average travel times incorporating all four components just described are calculated for each of the 24 road types listed in Table 2 at volume-capacity ratios ranging from 0 to 1.5 (at 0.01 increments). This results in a panel dataset with 3,624 observations of average travel time in minutes, $avgtt_{ij}$, where i indexes road type and j indexes the volume-capacity ratio. We shall refer to these data as the HCM data.

However, these calculations depend explicitly on the road type. Noting that the speed function in equation (2) can be expressed in terms of travel time (T) for a road of length L, $T_t \equiv L/S_t$, we seek a functional form for travel time that depends only on free-flow speed (S_F) and volume-capacity ratio ($v \equiv V/V_K$) in order to apply the theory developed in Section 2. The function should also capture the small increase in travel times when v is low, and the much steeper rise when v exceeds 1 due to queuing (as seen with the HCM data in Figures 2-3). We tried linear and nonlinear estimating equations, and the most realistic fit is obtained using a variation of the function proposed by Akçelik (1991) for the purpose of representing both normal flow (volume less than capacity) and queued flow in a single function, as described by Small and Verhoef (2007, eq. 3.11). The original Akçelik travel time function is:

$$T = T_F + 0.25P \cdot \left[(v-1) + \sqrt{(v-1)^2 + \frac{8J_a v}{V_K P}} \right] \tag{11}$$

[18] We also ignore the possibility of hypercongested steady states, which may occur in urban street networks (Small and Chu 2003). This simplification will also tend to cause us to underestimate the value of congestion relief.

where $T_F \equiv L/S_F$ is free-flow travel time and J_a is a constant taking on different values depending on the type of road, ranging from 0.1 for freeways to 1.6 for high-friction secondary arterials. The term under the square root provides for a modest increase in travel time with v when $v<1$, and for an increase approaching that from deterministic queuing behind a bottleneck when incoming flow is significantly greater than capacity.[19]

To fit with our theoretical model, however, the function cannot depend on road type except through S_F, nor can it depend on capacity except through the ratio $v \equiv V/V_K$. We therefore estimate a variant, motivated by two facts: (i) in Akçelik's derivation, the first term depends on the length of the road L but the second does not since it represents queuing delay at a single choke point; and (ii) empirically, S_F is positively correlated with (J_a/V_K). The modified Akçelik function to be estimated is:

$$T = \frac{L}{S_F} + \gamma_1 P\left[(v-1) + \sqrt{(v-1)^2 + (\gamma_2 / P)\exp(\gamma_3 \cdot S_F) \cdot v}\right] + \mu \tag{12}$$

where γ_1, γ_2 and γ_3 are the parameters to be estimated and μ is the error term. We estimate the equation holding constant P=4 hours and L=10 miles, which are the parameters we use to compute the HCM travel times that are the observations in the estimation. Each observation consists of one of our 24 road types and one of 151 values of v distributed evenly between zero and 1.5.

Our estimates, using nonlinear least squares, are given in Table 4. We note that our estimate of γ_1 is close to the value of 0.25 derived by Akçelik on theoretical grounds, as shown in equation (11).

[19] When the "delay parameter" J_a is zero, this equation simplifies to $T=T_F$ for $v \leq 1$ and $T=T_F+(1/2)P \bullet [v-1]$ for $v>1$.

Table 4. Estimates of modified Akçelik function

Parameter	Estimate	Standard error
γ_1	0.2929	0.0010
γ_2	126.3	38.0
γ_3	-0.1726	0.0085

Note: Based on 3,624 observations. R-squared = 0.9866.

Figures 2 through 4 compare the predicted travel times from equation (12) with those from which it was fitted (what we call "the HCM procedure," which means the HCM supplemented by our queuing model). The figures do this for a variety of road types with 12-foot lanes. For convenience, travel times are given in minutes. Figures 2 and 3 graph these travel times as a function of volume-capacity ratio v, whereas Figure 4 graphs them as a function of free-flow speed S_F.

Figure 2. Travel times for selected streets and highways

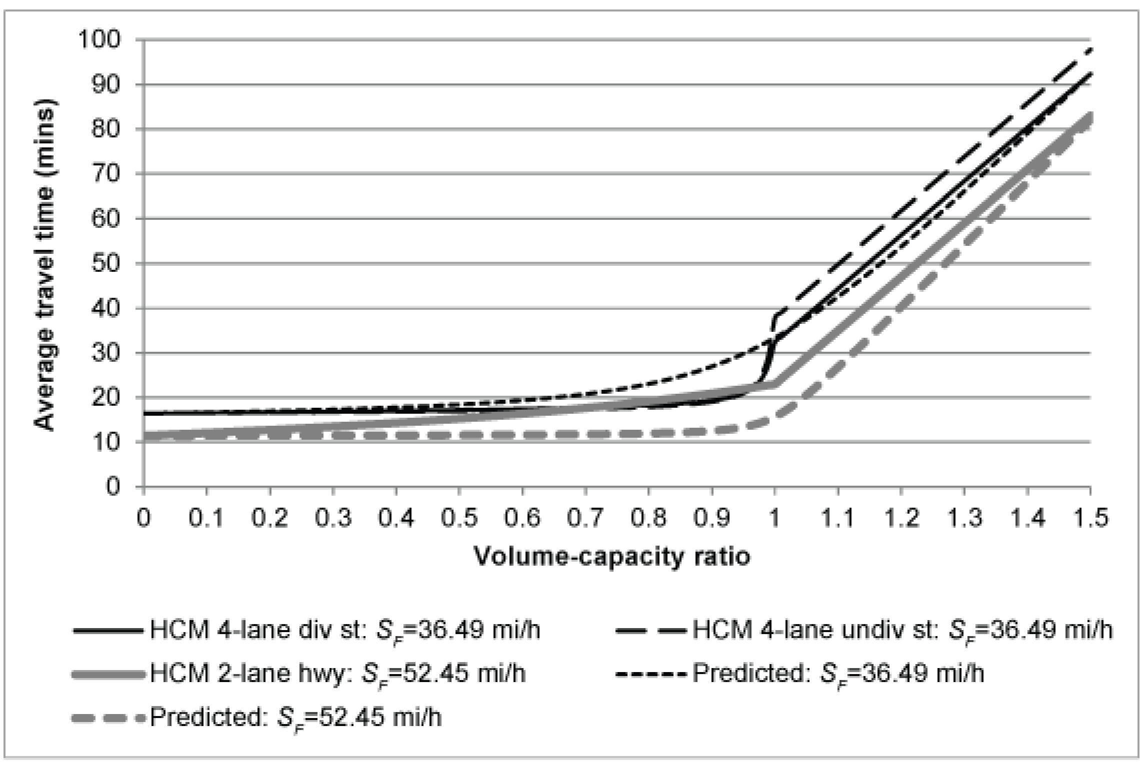


Figure 3. Travel times for a four-lane divided highway and freeway

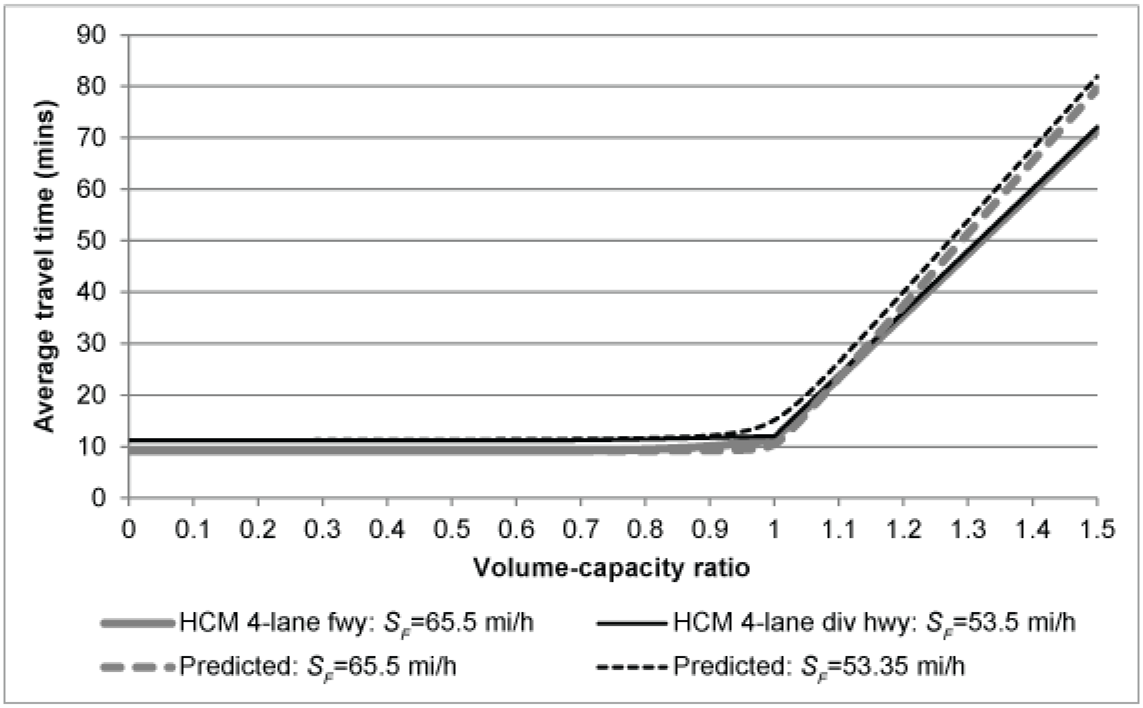
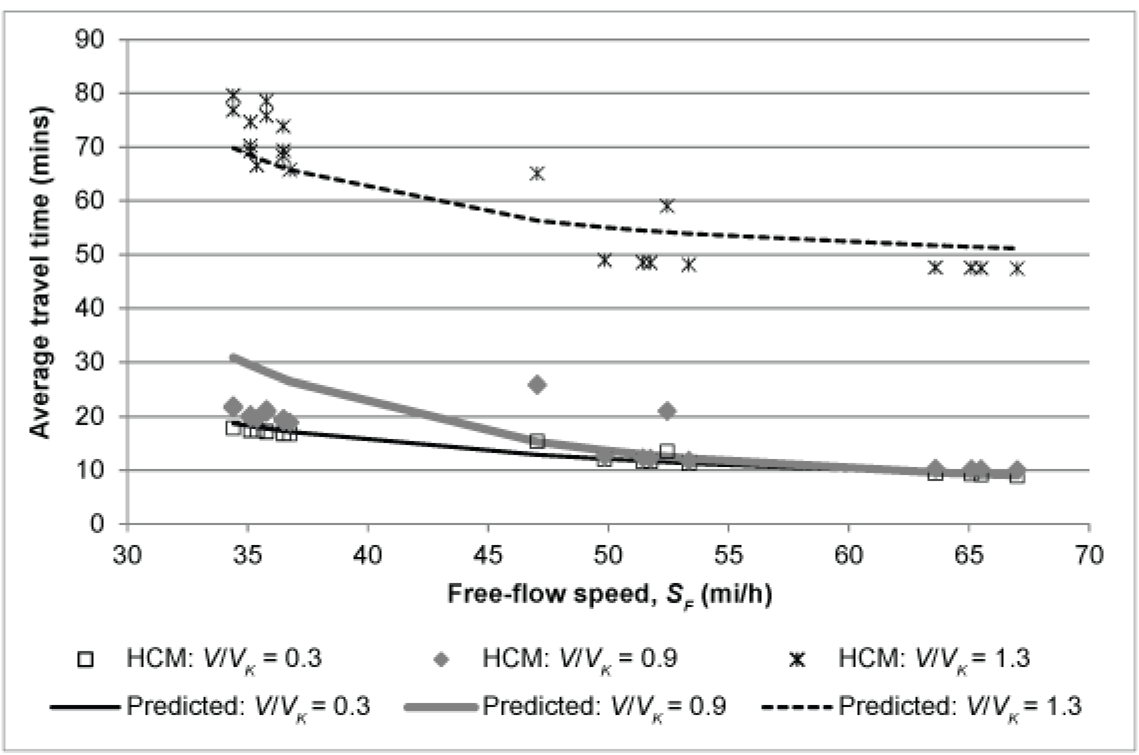


Figure 4. Travel times as a function of free-flow speed, for selected values of volume-capacity ratio



In general, the modified Akçelik function reproduces the shapes of the relationships quite well, while eliminating the kinks at v=1 that are an unrealistic artifact of the use of different procedures for v<1 and v>1. Especially helpful is that the modified function eliminates the unrealistic non-convexity at v=1 that occurs in our HCM procedure for urban streets, seen in Figure 2. The modified Akçelik function also captures the feature, arising directly from the HCM, that the travel time function is very flat almost up to $v = 1$ for higher road types. However, it underestimates travel times for two-lane highways because it interprets their relatively high free-flow speed as indicating a high road type, whereas actually traffic slows noticeably on two-lane highways even for moderate traffic levels. When queuing occurs (e.g., at v =1.3 as seen in Figure 4), predicted travel times are slightly underestimated for urban streets and two-lane highways, and overestimated for multilane highways and freeways.

Figures 2 through 4 show that our modified Akçelik function is convex in both traffic level (v) and free-flow speed (S_F) (see earlier footnote for how this relates to the second-order conditions for cost minimization).

The derivatives of the modified Akçelik function lead to the following values needed to calculate equation (5c):

$$(L/\alpha)c\varepsilon_{S,SF} \equiv T\varepsilon_{S,SF} = \frac{L}{S_F} - \frac{\gamma_1\gamma_2\gamma_3 v S_F \exp(\gamma_3 S_F)}{2z} \tag{13}$$

$$(L/\alpha)mecc \equiv v\frac{\partial T}{\partial v} = \gamma_1 P v\left(1 + \frac{v-1}{z} + \frac{\gamma_2 \exp(\gamma_3 S_F)}{2Pz}\right) \tag{14}$$

where

$$z = \left[(v-1)^2 + \frac{\gamma_2 v}{P}\exp(\gamma_3 S_F)\right]^{1/2}.$$

The asymptotic slope of (14) is proportional to P, just as for a simple bottleneck.[20]

[20] As $v\rightarrow\infty$, the second term in parentheses in (14) approaches 1 while the third term disappears, so that $\partial T/\partial V\rightarrow 2\gamma_1 P/V_K$. If γ_1 were equal to 0.25 as in the original Akçelik formula, this would be exactly the asymptotic

5. Numerical results for investment balance

We now apply the model to some examples of roads to see under what conditions these roads embody the optimal balance between S_F and V_K indicated by equation (5c). Specifically, we first calculate the left-hand side (LHS) of (5c), which we call the ratio of the construction cost elasticities, using the results from Section 3.2. Since this ratio does not depend on free-flow speed or capacity, it is equal to a constant (0.4090/1.2170 = 0.34). Next, for each road (characterized by free-flow speed and capacity), we substitute equations (13) and (14) (incorporating the estimated parameters from Table 4) into the right-hand side (RHS) of equation (5c), which we call the ratio of marginal user costs. We calculate this ratio for each road for given peak and off-peak traffic volumes, with assumptions on peak and off-peak durations, number of days of travel, etc, as listed below. We then evaluate LHS – RHS (the investment balance condition); as a reminder, a positive number favors marginal investment in S_F relative to that in V_K.

In Section 5.1, we evaluate investment balance for the hypothetical roads seen in the previous sections under a variety of traffic levels, in order to explore the range of conditions when each type of road is appropriate. In Section 5.2, we apply this methodology to empirical data for representative roads in various U.S. cities to see if conditions in these cities warrant improvement in capacity at the expense of free-flow speed (or vice versa). In Section 5.3, we go further and examine the absolute criteria in equations (5a-b), i.e., we calculate the benefit-cost ratios for investing in capacity or free-flow speed, for the same sample of cities.

5.1 Sampling the universe of urban road conditions

We first consider the investment balance condition for the specific road types we have been analyzing, shown in Table 2. We do so under the assumption of two time periods, each of uniform flow, and with traffic volume exogenous in each time period. We consider peak volume-

slope of the average wait through a bottleneck of capacity V_K over period P when that capacity is exceeded, as in equation (6). This is why our predicted travel-time curves rise nearly linearly with traffic at high traffic levels in Figures 2 and 3; their slopes are slightly higher than for the "HCM procedure" because our estimate of γ_1 slightly exceeds 0.25.

capacity ratios ranging from 0.1 to 1.25, holding constant the peak and off-peak durations (P=4 hours and F=12 hours, respectively), and with a constant ratio of peak to off-peak volume (V_p/V_o=1.25),[21] and other assumptions taken from Ng and Small (2012).[22]

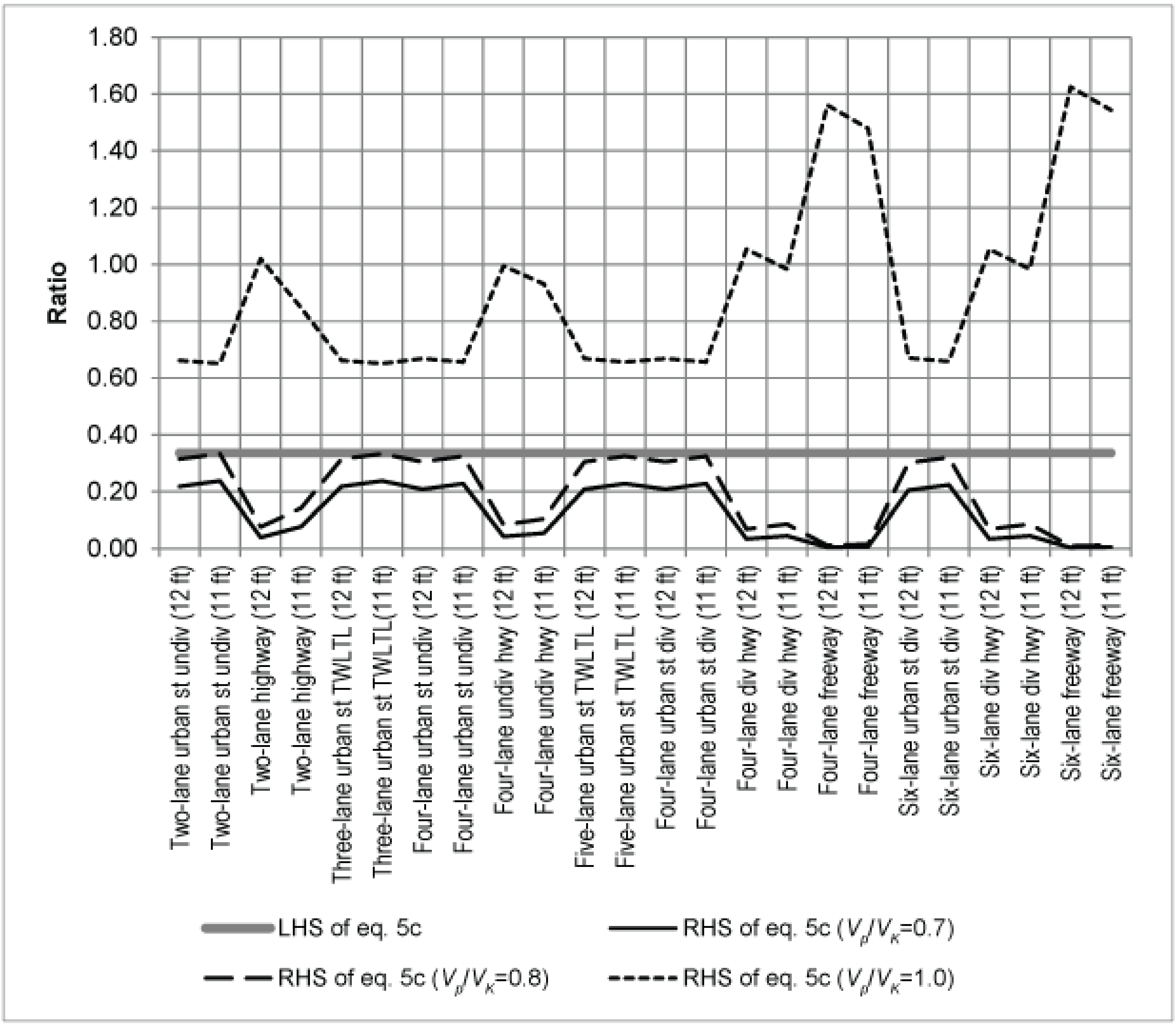
Some results are shown in Figure 5 (Appendix C has further details). The thick line shows the left-hand side of equation (5c) (the ratio of construction cost elasticities); whereas the thin and the dashed lines show the right-hand side (the ratio of marginal user costs) for three values of peak volume-capacity ratio (V_p/V_K). Incremental investment in S_F is more favorable than investment in V_K when the ratio of construction cost elasticities exceeds the ratio of marginal user costs, i.e., when the thick line lies above the thin or dashed line. We can see that when V_p/V_K is 0.8 or less, investing in S_F is more beneficial for all types of roads. But under highly congested conditions, as when V_p/V_K = 1, investment in S_F is never favored: rather, it is always better at the design stage to sacrifice some free-flow speed in order to increase capacity.

The case where V_p/V_K = 0.8 is illuminating. With this level of peak traffic, the ratio of construction cost elasticities is almost equal to the ratio of marginal user costs for arterials, indicating that arterials are close to meeting the investment balance condition. By contrast, all the highways and expressways of four lanes or more offer inefficiently high free-flow speeds relative to their capacity. A corollary is that if peak traffic congestion is at this level and if capacity is being optimized as called for by (4a), then (4b) indicates that most highways and expressways exhibit over-investment in free-flow speed under the design standards embedded in the Florida cost data.

[21] For example, the time-of-day profiles shown by Schrank et al. (2012b, Exhibit A-4) for "severe congestion" portray approximately 29 percent of weekday traffic occurring during the four highest hours for those highway links demonstrating afternoon peaking. If all the remaining traffic occurs during 12 off-peak hours, the peak-to-offpeak volumes are (0.29/4) / [(1-0.29)/12] = 1.22. Interdependence of demand across time periods would probably increase the benefits of investment in free-flow speed relative to capacity, since higher free-flow speed would reduce the external cost of congestion by attracting some peak traffic to off-peak, while the latter would increase it by attracting new traffic to the peak period (the well-known induced demand effect).

[22] These are: Peak period (in a given direction) occurs 310 days per year; off-peak period occurs for 12 hours/day on those same 310 days, and also occurs for 16 hours/day on the other 55 days.

Figure 5. The investment balance condition (5c) for 24 road types



Note: Investment in S_F is favored relative to that in V_K when the LHS (ratio of construction cost elasticities: thick line) exceeds the RHS (ratio of marginal user costs: thin and dashed lines).

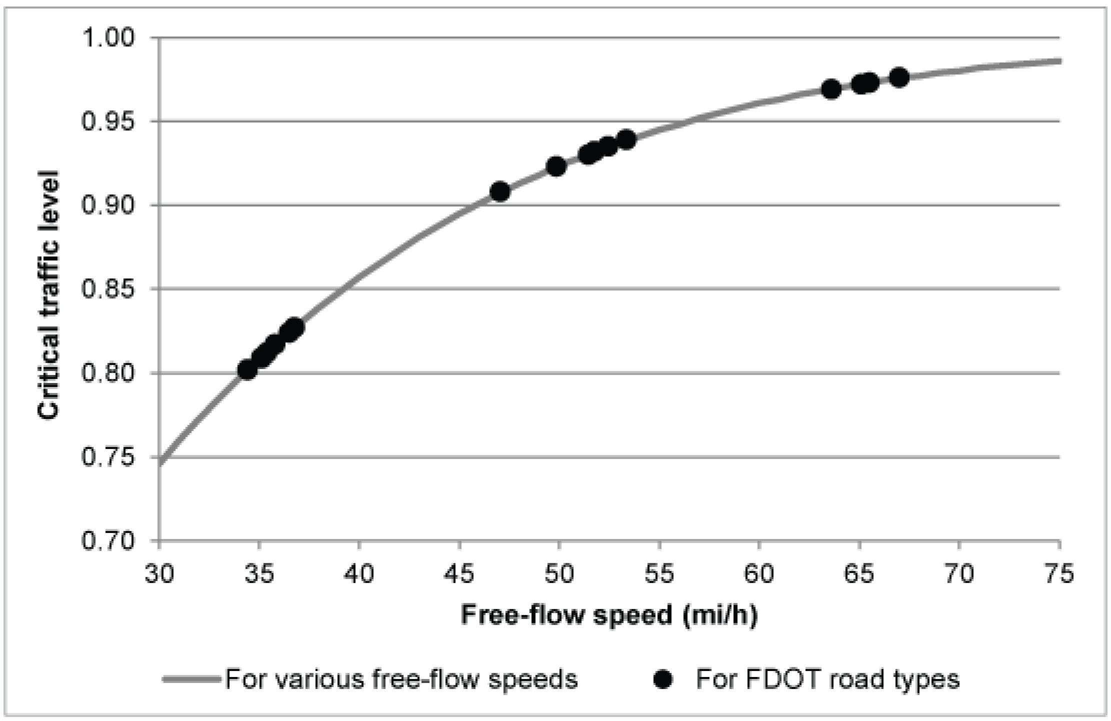
While these results are computed for a particular ratio of peak to off-peak traffic volume (V_p/V_o=1.25), they are quite insensitive to that ratio.[23]

Figure 6 broadens the computations to a wide range of free-flow speeds and displays the "critical traffic level," defined as the maximum value of V_p/V_K for which the ratio of construction

[23] This is because, as V_p/V_o increases, both the marginal external congestion cost and the average user cost of peak travelers rise relative to those of off-peak travelers; but since one is in the numerator and the other in the denominator of the ratio of marginal user costs, that ratio, which is the right-hand side of (5c), remains relatively constant. The left-hand side of the equation does not depend on traffic volumes at all; thus, the relationship between the two sides of the equation is relatively unaffected.

cost elasticities exceeds the ratio of marginal user costs (a situation favoring investment in free-flow speed relative to that in capacity). In other words, for any given road type, investment balance is realized when peak traffic congestion is described by the critical traffic level; if congestion is less the road is too slow at low flows, whereas if congestion is greater the road is over-invested in free-flow speed. Note that the critical traffic levels do not depend on capacity, because neither the construction cost elasticities nor the ratio of marginal user costs depend on capacity.

Figure 6. Critical traffic levels for various free-flow speeds and scatter plot (in black) of FDOT road types



Note: The critical traffic level is the maximum V_p/V_K for which incremental investment in S_F is more favorable than investment in V_K, according to equation (5c). It is calculated for 1 mi/h increments of free-flow speed.

The critical traffic levels increase with free-flow speed; this is due to the fact that as the volume-capacity ratio increases, the travel time function (seen in Figures 2-3) rises more quickly for low free-flow speed roads such as urban streets than for high free-flow speed roads such as multilane highways and freeways. As a result, for roads with low free-flow speed, marginal investment in capacity rather than in free-flow speed is beneficial at a lower volume-capacity ratio. Note that the critical traffic level is never greater than 1.0 because even a small amount of queuing causes the marginal external congestion cost to rise significantly, making the case for capacity investment very strong.

For the road types in our sample, shown as black dots in the figure, the critical traffic levels range from around 0.8 for urban streets, and from 0.9 to almost 1.0 for all other road types. Corresponding average peak speeds for these critical traffic levels, shown in Appendix C, range from 23 to 26 mi/h for urban streets, 39 to 48 mi/h for highways, and from 61 to 65 mi/h for freeways. It is apparent that whenever there is substantial peak congestion, a reconfiguration of these roads to extract more capacity at the expense of free-flow speed would be beneficial if it could be done at the design stage.

5.2 Investment balance for typical urban roads in the United States

We now examine the investment balance condition for some road conditions observed in US urban areas in 2011, using data from Schrank et al. (2012b) in their *Urban Mobility Report*. The report, which is published annually and provides an overview of congestion trends in the U.S., has detailed information on speeds, vehicle-miles traveled, and congestion delay for 498 urban areas. We use the average free-flow speed and average peak speed for "freeways" and "arterials" for "very large" and "large" urban areas.[24] We choose these areas because they are typically where congestion is worst and therefore we are most likely to see effects of underinvestment in capacity. Furthermore, it is in these cities where broad intrusive expressways have been most controversial, while at the same time these cities offer occasional examples of alternative high-capacity but slower highway designs such as Lakeshore Drive in Chicago and Storrow Drive in Boston.

We now must acknowledge that real roads underprice congestion, and this affects the first-order conditions for investment because a byproduct of investment may be to alter the marginal external cost of congestion. We focus on two primary channels for such alteration. Both

[24] These areas (of which there are 47 total) are defined as having population more than 3 million and 1–3 million, respectively. The data are from Schrank et al. (2012b), Appendix A, Exhibit A-8.

are due to investment in capacity – therefore leading to a modification of equation (5a) – and both are subjects of past research, described by Small and Verhoef (2007, Sect. 5.1.4).[25]

First, investment in capacity will lower the user cost of travel during congested periods, thereby attracting more traffic to the extent that there is a non-zero elasticity of demand for such travel. This is the problem of "induced demand," which if ignored results in overbuilding because of assuming more congestion relief than is actually achieved. Small and Verhoef (2007, eq. 5.8) show that the resulting second-best investment criterion for capacity replaces the factor q_tV_t in our equation (5a) by (q_tV_t-λ_t), where λ_t is a Lagrangian multiplier for the pricing constraint. The second-best solution for λ_t is:

$$\lambda_t = q_t V_t \frac{(mecc)_t - \tau_t}{(mecc)_t + \left(p_t / |\varepsilon_t|\right)}$$

where τ_t is the toll, p_t is the generalized price of travel (average user cost plus toll) and ε_t is the elasticity of demand with respect to generalized price. Taking the toll as zero, this implies that the marginal benefit of capacity expansion, i.e. the right-hand side of (4a), must be multiplied by an adjustment factor for induced demand:

$$R_t^1 = 1 - \frac{1}{1 + \left(p_t / mecc_t\right) \cdot \left(1 / |\varepsilon_t|\right)} \quad .$$

Based on standard conditions and the literature on demand elasticities, we take this factor to be 0.75 and 1.0 for the peak and off-peak periods, respectively.[26]

---

[25] Wilson (1983) provides an important contribution and a good summary of literature up to that point. It is also possible that the first-order condition for investment in free-flow speed, equation (5b), would need modification in a second-best world. We suspect that investment in free-flow speed would decrease the marginal external cost of congestion, because higher free-flow speed benefits off-peak travelers more than peak travelers; therefore such investment would shift some people from peak to off-peak. However, we think this would be a much smaller effect since it operates indirectly, in contrast to the direct effects of capacity investment.

[26] For the demand elasticity, we consider overall elasticity of travel, not just travel during a given time period, because cross-demand between periods is incorporated into our second adjustment factor. Estimates of the long-run elasticity of motor vehicle travel with respect to fuel price in the US suggest historical values on the order of -0.15 to -0.30, and current values considerably lower; we follow NHTSA (2012, p. 6) which uses -0.10 as the best estimate for analyzing future impacts from current or pending regulations. We take fuel cost to be 13.3% of the generalized price of travel, based on the calculation of 11% by Small and Verhoef (2007, Table 3.3) updated to 25% higher fuel

Second, higher capacity will attract traffic from off-peak to peak periods. Small and Verhoef (2007, Sect. 5.1.4) discuss implications for investment analysis. Using the model of Henderson (1992), in which demand patterns adjust continuously to maintain equilibrium among commuters, the adjustment factor would be about R^2_{peak}=0.95.[27] This is close to one, suggesting that our model ignoring cross-elasticities is a good approximation. Again, we may take the off-peak adjustment factor to be 1.0. In what follows, the $(mecc)_t$ calculations in equation (5a) are multiplied by the two adjustment factors, R^1_t and R^2_t.

To compute the investment balance condition, we also need to know road capacity and peak volume-capacity ratio. We combine data on road mileage from the Federal Highway Administration's *Highway Statistics* (2013) with lane-miles data from Schrank et al. (2012b) to obtain the average number of lanes for freeways and arterials in each urban area and use this to estimate capacity, assuming that arterials are equivalent to urban streets with signals (see Appendix C for details). Knowing both free-flow speed and peak speed, we can solve (12) iteratively to determine the peak volume-capacity ratio v_p; we then assume v_p/v_o=1.25, as before, to get the off-peak volume-capacity ratio v_o. Thus, for each urban area we have a representative "average" road (either a freeway or arterial) with unique free-flow speed, capacity, and peak and off-peak volume-capacity ratios; we use this information to calculate the two sides of the investment balance condition (equation [5c]). Note that because our calculations are highly non-linear and the data from Schrank et al. are averaged over the entire urban area, the investment balance for a representative road does not apply to all roads in the urban area.

---

prices prevailing now. Similarly, based again on the same table, we take travel-time cost to be 3.2 times today's fuel cost, or $3.2 \cdot 0.133 p_t = 0.43 p_t$. That implies $|\varepsilon_t|$=0.1/0.133=0.75. For a reasonably high level of peak congestion in which *mecc* equals the average travel-time cost, this would imply that $p_t/mecc_t \approx 2.3$ for the peak period. The above equation then implies a peak-period adjustment factor of R^1_{peak}=1-[1+(2.3/0.75)]$^{-1}$ = 0.75. We assume off-peak congestion is negligible for our purpose here, so $R^1_{off-peak}$=1.0.

[27] This is based on Henderson's equation (14), in which the desired adjustment factor is the quantity in the square brackets (evaluated at K=K_0) divided by its first term: i.e., R^2_{peak} = $[-(1+\gamma)+\gamma-\gamma/(1+\gamma)] / [-(1+\gamma)]$. In that equation, γ is the exponent in a travel time function $T=T_0(V/V_K)^\gamma$. Small (1992, pp. 69-72) estimated a function of this type, using aggregate observations from Boston expressways, and obtained an exponent $\gamma \approx 3.3$ (which captures the convexity of our travel time function as seen in Figures 2-3). Setting γ=3.3 in Henderson's equation, R^2_{peak} = [-4.3+1-(3.3/4.3)] / [-4.3] = 0.95.

We present the results of a sample of seven urban areas, chosen to cover most of the range of observed speeds on each road type, in Table 5.

Table 5. Investment balance for average road conditions in seven urban areas, 2011

	Very large areas				Large areas		
	Los Angeles	Dallas-Fort Worth	Miami	Boston	Denver	St. Louis	Jackson-ville
Freeways:							
Average no. of lanes	8.7	5.8	6.7	6.4	5.8	6.5	5.8
Free-flow speed, S_F (mi/h)	64.6	64.1	64.0	63.4	62.3	56.0	63.4
Peak speed, S_p (mi/h)	48.6	54.0	56.7	54.2	50.9	44.4	58.9
Peak volume-capacity ratio, V_p/V_K	1.016	1.003	0.994	0.999	1.004	0.995	0.979
Ratio of construction cost elasticities	0.34	0.34	0.34	0.34	0.34	0.34	0.34
Ratio of marginal user costs	1.82	1.20	0.81	1.02	1.19	0.75	0.37
Imbalance (+ favors investment in S_F)	-1.48	-0.86	-0.47	-0.68	-0.85	-0.42	-0.04
Arterials:							
Average no. of lanes	3.6	3.7	4.6	2.3	3.5	3.2	3.7
Free-flow speed, S_F (mi/h)	43.7	39.1	39.2	36.0	38.0	34.9	43.3
Peak speed, S_p (mi/h)	37.4	33.1	31.7	29.5	32.1	29.8	37.4
Peak volume-capacity ratio, V_p/V_K	0.811	0.695	0.758	0.639	0.662	0.534	0.788
Ratio of construction cost elasticities	0.34	0.34	0.34	0.34	0.34	0.34	0.34
Ratio of marginal user costs	0.17	0.15	0.19	0.15	0.14	0.12	0.16
Imbalance (+ favors investment in S_F)	0.16	0.19	0.15	0.18	0.19	0.22	0.18

Note: The imbalance is calculated as the ratio of construction cost elasticities minus the ratio of marginal user costs.
Sources: Schrank et al. (2012b), FHWA (2013), and authors' calculations; see text and Appendix C for more details.

From Table 5, we can see that the overall picture is that freeways demonstrate an over-investment in free-flow speed relative to capacity, whereas for arterials these two dimensions of investment are quite well-balanced. For example, despite its already high capacity, a representative Los Angeles freeway would benefit more from further capacity expansion than from further investment in free-flow speed, due to heavy congestion (second-lowest peak freeway speed among all urban areas). Peak freeway speed is lowest in St. Louis; but so is its free-flow speed, and as a result its investments are much closer to balance although still favoring capacity expansion. To put it differently, the case for giving up some free-flow speed in exchange for more capacity (for example by restriping for narrower lanes) is less strong in St. Louis than in Los Angeles.

For arterials, the imbalance is positive and generally quite close to zero, indicating that there is a slightly greater incremental benefit from improving arterial free-flow speeds than from expanding arterial capacity. Increasing free-flow speed for arterials—which here are assumed to be urban streets with signals—need not necessarily imply upgrading to a higher road type, but could involve targeted upgrades to reduce delays from traffic signals. Such upgrades are analyzed by Samuel (2006, ch. 4), who describes a number of innovative intersection designs that improve both free-flow speed and capacity with modest cost and land requirements. Since these improvements also increase capacity, it is unclear without more detailed analysis what their availability implies for investment balance as defined here.[28]

To test the sensitivity of the results to our assumptions, we consider a refinement of the peaking model. We now assume that there is a two-hour "peak peak" period during which traffic volumes (denoted by V_p^p) are highest, flanked by two "shoulder peak" periods, each with the same traffic volume (V_p^s) and of duration one hour. Full details are provided in Appendix D, including assumptions on the ratio of V_p^p to V_p^s using data from Schrank et al. (2012b) and formulas for the queuing delay. This refined peak model produces a more realistic range of peak volume-capacity ratios across cities: for freeways, they vary from 0.75 in Jacksonville to 0.99 in Los Angeles. As for results, the refined peak model causes the investment imbalances for freeways for the urban areas in our study to become less negative and even slightly positive for

[28] We perform a sensitivity analysis by assuming P=2 and F=14 instead and reestimating the travel time function. Since there are now fewer vehicles affected by congestion and for a given value of v_p, there is also less congestion, many road types now have a higher critical traffic level (defined in Section 5.1), i.e., there are now more instances where incremental investment in S_F rather than V_K is beneficial. As a result, in many urban areas, the freeway imbalance becomes positive though very close to zero; whereas the arterial imbalance is little changed (close to zero). We consider the assumption of P=4 for one-way travel to be more realistic and it is in line with Schrank et al.'s (2012b) definition of peak hours as 6-10 a.m. and 3-7 p.m., but it is useful to keep in mind that the "balance" for a real road depends quite sensitively on the peaking characteristics.

Miami and Jacksonville (thus favoring marginal investment in free-flow speed rather than capacity for the average freeway in these two urban areas), while the arterial imbalances remain very similar.

5.3 Absolute investment criteria

In addition to examining the relative investment criterion, we can analyze the absolute investment criterion for either capacity or free-flow speed, each holding the other constant. The criteria are contained in equations (4a) and (4b), respectively, or equivalently (5a) and (5b)—again with (4a) or (5a) modified by multiplying $(mecc)_t$ by the adjustment factors R_t^1 and R_t^2 . We summarize by calculating the benefit-cost ratio as the travel-time savings from an incremental increase in capacity or free-flow speed divided by the corresponding incremental capital cost. From equation (5a), investment in V_K is warranted if the benefit-cost ratio exceeds one:

$$\frac{B}{C} \equiv \frac{\sum_t q_t V_t \cdot (mecc) \cdot R_t^1 R_t^2}{\rho \varepsilon_{\rho,VK}} > 1. \qquad (15a)$$

Similarly, equation (5b) yields the investment criterion for free-flow speed:

$$\frac{B}{C} \equiv \frac{\alpha \sum_t q_t V_t T_t \left(\varepsilon_{S,SF}\right)_t}{\rho L \varepsilon_{\rho,SF}} > 1 \qquad (15b)$$

where we have replaced per-mile user cost c_t by $\alpha T_t/L$, i.e., travel time per trip divided by trip length L and multiplied by value of time α. The components of these equations can be computed using equations (10), (13), and (14) along with assumptions about amortization, land acquisition, value of time, duration of travel periods, capacities, volume-capacity ratios, and trip length.[29]

[29] In addition to the assumptions mentioned in previous sections, we need values for the interest rate (r), lifetime of the road (Λ) and land acquisition costs as a percentage of total capital cost (x) to calculate ρ using equation (10). Based on Ng and Small (2012), we set r=0.07, Λ=25 years and x=0.183 (since the urban areas in our sample have

One can alternately view this calculation as showing the maximum cost multiplier that could justify the investment under consideration. By "cost multiplier" we mean the incremental cost of expanding either S_F or V_K for a given hypothetical project, divided by the corresponding incremental cost as observed in our Florida cost data.

Table 6 shows the results for the sample of cities already discussed in Section 5.2 (using the base model with uniform traffic volume during the peak period). Using the figures in Table 6, the case for investment appears strong in both dimensions, in all areas. The variations across cities are not surprising. The case for investment in freeway capacity is extremely strong for a representative road in Los Angeles, with its low average peak freeway speed, and much less so in relatively uncongested Jacksonville. For representative arterials, the case for capacity investment is strongest in Miami: despite the already high arterial capacity there, arterial traffic congestion is also high as indicated by the large differential between free-flow speed and peak speed (see Table 5).

---

populations of 1 million or more). We use the same value of time per passenger vehicle as Schrank et al. (2012b), namely $16.79/hr; they base their figure on McFarland and Chui's (1987) estimate, updated to 2011 dollars, and on assumed average vehicle occupancy of 1.25. We do not account for the value of truck time.

Table 6. Absolute benefit-cost ratios from incremental investments, assuming Florida capital costs, 2011

	Very large areas				Large areas		
	Los Angeles	Dallas-Fort Worth	Miami	Boston	Denver	St. Louis	Jackson-ville
Freeways:							
Free-flow speed, S_F (mi/h)	64.6	64.1	64.0	63.4	62.3	56.0	63.4
Capacity, V_K (veh/h)	18,519	12,307	14,268	13,616	12,382	13,736	12,322
Capital cost, ρ (1000 $ per year per mi)	2,710	2,272	2,409	2,336	2,200	2,016	2,243
B/C: incr. invest. in V_K	49.1	26.5	19.2	24.8	29.1	28.3	7.9
B/C: incr. invest. in S_F	9.1	7.4	8.0	8.2	8.2	12.6	7.1
Arterials:							
Free-flow speed, S_F (mi/h)	43.7	39.1	39.2	36.0	38.0	34.9	43.3
Capacity, V_K (veh/h)	3,216	3,337	4,284	1,589	3,123	2,751	3,393
Capital cost, ρ (1000 $ per year per mi)	823	730	811	487	686	587	832
B/C: incr. invest. in V_K	4.3	4.3	7.7	3.5	4.1	3.0	4.0
B/C: incr. invest. in S_F	8.5	9.9	13.7	7.6	9.7	8.5	8.5

Note: B/C is the benefit cost ratio from incremental investment in capacity (V_K) and free-flow speed (S_F) calculated using equations (15a) and (15b), respectively.

Turning to investment in greater free-flow speed, the case is especially strong for the representative St. Louis freeway, which experiences unusually slow free-flow travel; and for the representative Miami arterial, which carries the most traffic of any of the arterials shown.

6. Conclusion

When free-flow speed is distinguished as an additional dimension of road investment, it becomes possible to analyze some important questions about road design within an optimization framework familiar to economists. Specifically, we can analyze criteria for investment not only in road capacity but in free-flow speed, which effectively means choosing among road types and/or specific design criteria such as lane widths. There is sufficient independence between these two dimensions that one can not only analyze each individually, but consider the optimal balance between them.

Empirically, we find that despite the discreteness of road types, it is feasible to approximate the range of possibilities by describing capital cost and user time costs as functions of capacity and free-flow speed. Doing so will not answer a specific design question for a specific road, but it is useful for broad-brush analyses of road policy, such as occurs in discussions about what type of road network a city needs. Our empirical analysis provides suggestive evidence that in many large congested cities, standard expressway designs are unbalanced in the sense of providing more free-flow speed than is desirable relative to capacity; whereas the same is not true for urban streets and arterial highways. This observation in turn suggests giving greater attention to the possibilities of more low-footprint roads which offer considerable capacity even though speeds are only moderate even at low traffic levels.

There are numerous factors not considered here that would be beneficial to add to this type of analysis. We mention a few here.

First, as emphasized by Ng and Small (2012), these design features have implications for safety which are potentially important but not well understood empirically. Furthermore, these safety implications could change dramatically as technologies, social customs, and legal environments evolve.

Second, some design features that reduce free-flow speed, such as reduced lane or shoulder widths, would be easier to undertake if large trucks are excluded from the road. Therefore, if one wants to use our analysis to reexamine policy toward road design, it would be a good time to also reexamine policy toward separating trucks and cars onto different roads. We conjecture that such policies would be strongly complementary.

Third, a broad policy analysis is likely to affect networks of roads, not just individual roads, which raises the question of how intersections affect costs. Kraus (1981) finds that accounting for the cost of intersections substantially decreases the measured scale economies with respect to capacity, because intersection costs tend to rise more than proportionally to the capacities of the intersecting roads. Whether any similar conclusion would apply for the elasticity of road costs with respect to free-flow speed would be extremely interesting and potentially important to discover.

Fourth, applications to particular road investments need to distinguish a finer time pattern of demand, in order to reduce inaccuracies caused by applying nonlinear relationships to averages. Doing so could also necessitate accounting for demand shifts across times of day. Alternatively, one might consider continuous-time models, such as the “bottleneck model” of Vickrey (1969) and Arnott et al. (1991), which deal with both issues simultaneously.

Finally, the advent of road pricing could substantially change the optimal balance analyzed here. It is difficult to predict the direction. On the one hand, traffic would be more evenly distributed across time, thus favoring more investment in free-flow speed; on the other, the payoff from capacity investment would no longer be limited by induced traffic, and there would be less total traffic to benefit from higher free-flow speeds. We can conclude that a potentially important long-run implication of road pricing would be to change the nature of a desirable urban road network, but pending a model that incorporates pricing, it remains to be seen exactly how.

Given these and other model improvements, we believe our approach to analyzing road investment offers the potential to expand insights and to inform practical analysis, both of which could enhance the efficiency with which roads are provided.

References

Akçelik, Rahmi, 1991. Travel time functions for transport planning purposes: Davidson's function, its time-dependent form and an alternative travel time function. Australian Road Research 21, pp. 49–59.

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.

Arnott, Richard, André de Palma, and Robin Lindsey, 1990. Economics of a bottleneck. Journal of Urban Economics 27: 111-130.

Ben-Akiva, Moshe, Moshe Hirsh and Joseph Prashker, 1985. Probabilistic and economic factors in highway geometric design. Transportation Science, 19, 38-57.

Daganzo, Carlos F., Michael J. Cassidy and Robert L. Bertini, 1999. Possible explanations of phase transitions in highway traffic. *Transportation Research Part A*, 33, 365-379.

FHWA (Federal Highway Administration), 2013. Highway Statistics 2011. http://www.fhwa.dot.gov/policyinformation/statistics/2011/

Henderson, J. Vernon, 1992. Peak shifting and cost-benefit miscalculations. Regional Science and Urban Economics, 22: 103-121.

Jansson, Jan Owen, 1984. Transport System Optimization and Pricing. Chichester, UK: John Wiley & Sons.

Keeler, Theodore E. and Kenneth A. Small, 1977. Optimal peak-load pricing, investment, and service levels on urban expressways, Journal of Political Economy 85, 1-25.

Kraus, Marvin, 1981. Scale economies analysis for urban highway networks, Journal of Urban Economics, 9, 1-22.

Larsen, Odd I., 1993. Road investment with road pricing: Investment criteria and the revenue/cost issue. In: A. Talvitie, D. Hensher, and M. E. Beesley (eds.), Privatization and deregulation in passenger transportation: Second International Conference on Privatization and Deregulation in Passenger Transportation. Espoo, Finland: Viatek Ltd., pp.273-281.

Lindsey, Robin and Erik T. Verhoef, 2000. Congestion modelling. In: David A. Hensher and Kenneth J. Button (eds.), Handbook of Transport Modelling. Amsterdam and New York: Pergamon., pp. 353-373.McFarland, William F. and Margaret Chui, 1987. The Value of Travel Time: New Estimates Developed Using a Speed Choice Model. Transportation Research Record, No. 1116, pp. 15-21.

Meyer, John R., John F. Kain and Martin Wohl, 1965. The Urban Transportation Problem, Cambridge, MA: Harvard University Press.

Mohring, Herbert, 1965. Urban highway investments. In: Robert Dorfman (ed.) Measuring Benefits of Government Investments, Washington, DC: Brookings Institution, pp. 231-275.

Mohring, Herbert, and Mitchell Harwitz, 1962. Highway Benefits: An Analytical Framework, Evanston, IL: Northwestern University Press.

Ng, Chen Feng and Kenneth A. Small, 2012. Tradeoffs among Free-flow Speed, Capacity, Cost, and Environmental Footprint in Highway Design, Transportation, 39, 1259-1280.

NHTSA (National Highway Traffic Safety Administration), 2012. Corporate Average Fuel Economy for MY 2017-MY 2025 Passenger Cars and Light Trucks: Final Regulatory Impact Analysis. Office of Regulatory Analysis and Evaluation, National Center for Statistics and Analysis, August.

Samuel, Peter, 2006. Innovative Roadway Design: Making Highways More Likeable. Reason Foundation Policy Study 348. http://reason.org/news/show/innovative-roadway-design

Santos, Adella, Nancy McGuckin, Hikari Yukiko Nakamoto, Danielle Gray, and Susan Liss, 2011. Summary of Travel Trends: 2009 National Household Travel Survey. http://nhts.ornl.gov/2009/pub/stt.pdf

Schrank, David, Bill Eisele, and Tim Lomax, 2012a. Adding New Lanes or Roads. http://mobility.tamu.edu/mip/strategies_pdfs/added-capacity/technical-summary/adding-new-lanes-or-roads-4-pg.pdf.

Schrank, David, Bill Eisele, and Tim Lomax, 2012b. Urban Mobility Report 2012. http://mobility.tamu.edu/ums/

Skabardonis, Alexander, and Richard Dowling, 1996. Improved speed-flow relationships for planning application. Transportation Research Record 1572, 18-23.

Small, Kenneth A., 1992. Urban Transportation Economics. Harwood Academic Publishers, Chur Switzerland, now Routledge.

Small, Kenneth A., and Xuehao Chu, 2003. Hypercongestion. *Journal of Transport Economics and Policy* 37, pp. 319-352.

Small, Kenneth A., and Eric T. Verhoef, 2007. The Economics of Urban Transportation. Routledge, Taylor and Francis Group, London and New York.

Strotz, Robert H., 1965. Urban transportation parables. In: Julius Margolis (ed.), The Public Economy of Urban Communities, Washington, DC: Resources for the Future, pp. 127-169.

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

Vickrey, William S., 1969. Congestion theory and transport investment. American Economic Review 59(2): 251-260.

Walters, A.A. (1968). The Economics of Road User Charges. International Bank of Reconstruction and Development, distributed by Johns Hopkins Press, Baltimore.

Wilson, John D. (1983). Optimal road capacity in the presence of unpriced congestion. *Journal of Urban Economics* 13, pp. 337-357.

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. Transportation Research Board, Washington D.C.

Notation

t	Index for time periods, $t = 1,2,\ldots,n$
q_t	Duration of time period t
V_t	Traffic volume at time t
V_K	Capacity
v_t	Volume-capacity ratio (V_t/V_K)
S_F	Free-flow speed (including control delay at zero traffic volume for urban streets)
S_t	Average speed
T_F	Free-flow user time (entire trip)
T_t	Average user time (entire trip)
ρ	Annualized road capital cost (per mile)
r	Interest rate
Λ	Lifetime of road in years
L	Trip length
$\Gamma(\cdot)$	Road construction cost (per mile)
$A(\cdot)$	Right-of-way acquisition cost (per mile)
c_t	Average user cost per vehicle-mile at time t
U_t	Total user cost per road-mile per hour at time t
C	Total agency plus user cost (short run) per road-mile
$\tilde{C}$	Total agency plus user cost (long run) per road-mile
α	Value of time

Appendices for:
Optimizing Road Capacity and Type
Kenneth A. Small
Chen Feng Ng
Feb. 13, 2014

Appendix A. Speeds and capacities from the HCM

This appendix discusses the Highway Capacity Manual's (2000) methodology for calculating speeds and capacities for four types of road: freeways (based on HCM ch. 13, 23), multilane and two-lane highways (based on HCM ch. 12, 20, 21), and signalized urban arterials (based on HCM ch. 10, 15, 16). We focus on urban roads, using the road characteristics specified in the Florida Department of Transportation's cost descriptions as much as possible and otherwise using the default parameters recommended by the HCM. Many of these procedures are identical to those used in Ng and Small (2012).

A.1 Freeways

Free-flow speed

The first step is to estimate free-flow speed (FFS) using equation 23-1 in HCM:

$$FFS = BFFS - f_{LW} - f_{LC} - f_N - f_{ID} \quad \text{(A.1)}$$

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. Note that FFS is not the same as a legal speed limit.

The lane width adjustment, f_{LW}, is 0 and 1.9 respectively for lane width of 12 feet and 11 feet. In all our example freeways, the left and right shoulder widths are 10 feet each, for which f_{LC}=0. Our freeways have either two or three lanes in one direction, for which f_N is 4.5 and 3.0 respectively. Using the default interchange density of 0.5 interchanges per mile gives f_{ID} = 0.

Capacity

Capacity (measured in vehicles per hour) depends on free-flow speed, number of lanes, proportion and types of heavy vehicles, and how familiar drivers are with the road. The calculation proceeds in two steps.

First, the HCM defines "base capacity" *BaseCap* in units of passenger-car-equivalents per hour per lane (pce/h/ln). Its verbal description (p. 23-5), confirmed by the Highway Performance Monitoring System (HPMS) Field Manual (FHWA, 2002, Appendix N), implies:

$$BaseCap = \max\{1700 + 10FFS,\ 2400\} \tag{A.2}$$

which has a maximum of 2,400 pce/h/ln achieved when $FFS \geq 70$ mi/h.

Next, passenger-car equivalents per hour V^{pce} are converted to vehicles V as follows (equation 23-2):

$$V = V^{pce} \cdot PHF \cdot N \cdot f_{HV} \cdot f_p \tag{A.3}$$

where *PHF* is a "peak-hour factor" representing variation in traffic demand within an hour; N is the number of lanes in one direction; f_{HV} is an adjustment factor for heavy vehicles; and f_p is an adjustment factor for driver population (commuters or recreational drivers). For default values HCM in Exhibit 13-5 recommends PHF = 0.92 (for urban areas) and f_p = 1.00 (which applies for commuters). It also recommends a default value of 5% for percentage of heavy vehicles on freeways (Exhibit 13-5); we assume that heavy vehicles consist only of trucks and buses (no recreational vehicles) and that the freeway is on level terrain; this gives f_{HV} = 0.98 (HCM equation 23-3).

To summarize, using the values just listed we can calculate total one-directional capacity in vehicles per hour for each freeway configuration we consider:

$$\textit{One-directional capacity} = \left[\max\{1700 + 10FFS,\ 2400\}\right] \bullet PHF \bullet N \bullet f_{HV} \bullet f_p \tag{A.4}$$

Speed

The HCM gives a speed-flow formula for average passenger-car speed S (mi/h) as a function of per-lane flow rate V^{pce} (pce/h/ln), which applies for $V^{pce} \leq BaseCap$ and for free-flow speeds between 55 and 70 mi/h (HCM, Exhibit 23-3):

$$S = FFS - \max\left\{0, 0.7778 \cdot (\mathrm{FFS} - 48.57) \cdot \left(\frac{V^{pce} + 30FFS - 3400}{40FFS - 1700}\right)^{2.6}\right\}. \tag{A.5}$$

A.2 Multilane highways

In the HCM terminology, "multilane highways" differ from freeways in that highways are not fully access-controlled (i.e. local landowners can access them with driveways), and they can have at-grade road intersections with or without traffic signals if spaced more than two miles apart. Note that we accept the mild inconsistency in the meaning of "highway", which in the HCM is used (even in its title) as a general term for all types of road as well as a specific designation for major roads with characteristics intermediate between freeways and urban streets.

The capacity and free-flow speed of a multilane highway are calculated using the procedures outlined in Chapters 12 and 21 of the HCM, which are very similar to the freeway calculations.

Free-flow speed

Free-flow speed is estimated using HCM equation (21-1):

$$FFS = BFFS - f_{LW} - f_{LC} - f_M - f_A \tag{A.6}$$

where *BFFS* is the base free-flow speed (60 mi/h as stated in HCM Exhibit 12-3), f_{LW} is an adjustment for lane width, f_{LC} is an adjustment for lateral clearance, f_M is an adjustment for median type, and f_A is an adjustment for access density (Exhibits 21-4 to 21-7). The lane width adjustment is identical to that of the freeway case. The lateral clearance adjustment is based on the right and left lateral clearances from the travel lanes to roadside obstructions such as light standards, signs, trees, etc; a standard raised curb is not considered an obstruction. The right lateral clearance for our example highways is 4 feet (based on the right shoulder width) and the left lateral clearance is 6 feet (for both undivided and divided roads), leading to f_{LC} = 0.4. The median adjustment, f_M, is 1.6 for undivided highways and zero for divided highways. For access adjustment, we use the HCM's default value of 25 access points per mile for a high-density suburb (Exhibit 21-4), which implies f_A = 6.25.

Capacity

Capacity is calculated in the same manner as for freeways, except that (A.2) is replaced by:

$$BaseCap = \max\{1000 + 20FFS,\ 2200\} \tag{A.7}$$

Speed

The speed-flow functions in HCM Exhibit 21-3 are used to estimate speed depending on the traffic volume. For our free-flow speeds, they imply that

$$S = \begin{cases} FFS - 0.1658 \cdot \left[\left(FFS - 32.21 \right) \left(\dfrac{V^{pce} - 1400}{34.2FFS - 1181} \right)^{1.31} \right] \text{ if } 50 < FFS \leq 55 \\ FFS - 0.2326 \cdot \left[\left(FFS - 35 \right) \left(\dfrac{V^{pce} - 1400}{33FFS - 1050} \right)^{1.31} \right] \text{ if } 45 < FFS \leq 50 \end{cases} \tag{A.8}$$

or S=FFS, whichever is smaller.

A.3 Two-lane highways

Our calculations are based on HCM, ch. 20. Two-lane highways here assumed undivided and have *FFS* between 45 and 65 mi/h, estimated as follows (HCM equation 20-2):

$$FFS = BFFS - f_{LS} - f_A \tag{A.9}$$

where f_{LS} and f_A are adjustments for lane/shoulder width and access points, respectively (Exhibits 20-5 to 20-6). The highways in this paper have shoulder widths of 4 feet, giving f_{LS}=1.3 and f_{LS}=1.7 for lane width 12 feet and 11 feet, respectively. The adjustment for access density is identical to that for multilane highways, in our case f_A = 6.25. However, the HCM gives no guidance for *BFFS*; we assume it is 60 and 55 mi/h for 12-foot and 11-foot lane widths, respectively. These values yield *FFS* of 52.45 mi/h for 12-foot lanes and 47.05 mi/h for 11-foot lanes.

Two-lane highways have a fixed capacity of 1,700 pce/h for each direction of travel. To convert passenger-car equivalent flow rates (V^{pce}) to volumes in terms of vehicles per hour, we apply HCM equation 20-3:

$$V = V^{pce} \cdot PHF \cdot f_G \cdot f_{HV} \tag{A.10}$$

where the peak hour factor (*PHF*) is 0.92 as in the case of freeways and the grade adjustment factor (f_G) is 1 for level terrain. The heavy vehicle adjustment factor, again assuming 5% heavy vehicles and no recreational vehicles, is f_{HV} = 1/[1+0.05(E_T–1)] where:

$$E_T = \begin{cases} 1.7 \text{ if } V^{pce} \le 300 \\ 1.2 \text{ if } 300 < V^{pce} \le 600 \\ 1.1 \text{ if } V^{pce} > 600 \end{cases} \qquad \text{(A.11)}$$

based on HCM equation 20-4 and Exhibit 20-9. At capacity, the one-directional flow rate exceeds 600 pce/h, so f_{HV} = 0.995. These assumptions and setting V^{pce} = 1,700 pce/h yield a one-directional capacity of 1,556.22 veh/h.

Average travel speed is estimated using HCM equation 20-5:

$$S = FFS - 0.00776(V^{pce,a} + V^{pce,o}) - f_{np} \qquad \text{(A.12)}$$

where $V^{pce,a}$ and $V^{pce,o}$ are respectively the approach and opposing flow rates (in pce/h) and f_{np} is the adjustment for percentage of no-passing zones.[30] For simplicity, we assume $V^{pce,a}$ = $V^{pce,o}$ and an absence of no-passing zones, so f_{np} = 0 (Exhibit 20-11). Note that when converting volumes (in veh/h) to passenger-car equivalent flow rates to calculate average travel speed, E_T depends on V^{pce} and vice versa as seen in equations (A.10) and (A.11). As a result, the iterative procedure recommended by the HCM on p. 20-9 is used where V^{pce} is initially estimated as *V*/*PHF*, then the appropriate E_T is selected from equation (A.11) and third step is to recalculate V^{pce} using equation (A.10). If V^{pce} from step three exceeds the flow-rate range from which E_T was chosen in step two, E_T is now selected from the higher flow-rate category and the process is repeated until an acceptable value of V^{pce} is found.

A.4 Urban streets

The urban streets in our paper are assumed to be suburban principal arterials (design category 2), with one signalized intersection per mile, speed limits of 40-45 mi/h, no parking,

[30] For two-lane highways and urban streets with permitted left turns (two- and four-lane undivided streets), explicit assumptions are needed for the opposing traffic flow. Opposing flows affect speed on two-lane highways and the capacities and therefore control delay of urban streets with permitted left turns. We assume that the opposing flow rate is equal to approach flow rate, which is a useful simplification for the general case, but this does cause travel times to be slightly underestimated (overestimated) if the actual opposing flow is less (greater) than the approach flow.

and little pedestrian activity. The HCM's definition of "free-flow speed" does not consider control delay at signalized intersections (details below); henceforth we call this "unimpeded speed" and our paper's use of "free-flow speed" for urban streets is based on the unimpeded speed and control delay when traffic volume is zero (since we want to relate free-flow speed to travel time). The HCM procedures for urban streets are detailed in chapters 10, 15, and 16 but the HCM provides little guidance for estimating unimpeded speeds when field measurements are not available. We use the procedure recommended by Zegeer *et al* (2008, pp. 66-73) where unimpeded speed is determined by the speed limit of the road (45 mi/h and 40 mi/h for the roads with 12-foot lanes and 11-foot lanes, respectively). The Florida Department of Transportation's cost descriptions specify that the urban streets have curbs and gutters, and we assume that access density is 25 per mile as in the case of highways. These assumptions lead to the free-flow speeds seen in Table 2 of the text.

A vehicle's travel time on an urban street (ignoring queuing due to volumes exceeding capacity, computed separately) consists of running time plus control delay. Based on Exhibit 15-3 of the HCM, running time for an urban street one mile or longer is calculated as simply the length divided by the unimpeded speed. We assume that the capacity of the urban street is equal to the capacity of the signalized intersections (described below), and queuing when volume exceeds capacity occurs only at the entrance to the road, prior to the first signal.

Control delay is the 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- or right-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, the control delay calculations will focus on lane groups with through lanes (which could be shared turn lanes).

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 mentioned above, the initial queue occurs only once at the road entrance before the first signal since the traffic volume arriving at each intersection is never greater than the intersection's capacity. This queuing delay is calculated separately using the bottleneck queuing model described in the text, and 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:

$$d = \frac{0.5Z(1-g/Z)^2}{1-[\min(1,X)(g/Z)]} \cdot \phi + 900q\left[(X-1) + \sqrt{(X-1)^2 + \frac{8kIX}{cap \cdot q}}\right] \tag{A.13}$$

where Z is the cycle length, g is effective green time, X is the volume-capacity ratio of that lane group, ϕ is the progression adjustment factor, q is the duration of the analysis period (in hours), k is the incremental delay factor, I is the upstream filtering factor (equal to 1 since the upstream signal is more than a mile away), and *cap* is lane-group capacity. Time durations Z, g, and hence d are all conventionally measured in seconds. The first term in equation (A.13) is the uniform control delay while the second term is incremental delay.

The progression adjustment factor, ϕ, 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 $\phi = 1$. 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 through and shared turn/through lane groups have identical values of g/Z and that traffic distributes across lanes so that they have identical values of X. We also assume that vehicles are not allowed to turn right during red signal phases. Therefore these lane groups have the same delay, given by equation (A.13). The total control delay 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 for X, with one-directional capacity defined appropriately as we now describe.

The urban street's capacity is based on the saturation flow rates, s_i, of the through/shared through lane groups, along with the fraction of time the signal is green and the proportion of traffic at each intersection that is making turns if there are any exclusive turn lanes. (It is assumed that the exclusive turn lanes have ample capacity; a reasonable assumption since we are using the HCM default value of 10% of total traffic each turning left and right.) 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 as cap_i) is:

$$cap_i = s_i \cdot (g_i / Z) \tag{A.14}$$

where the effective green ratio g_i/Z is here taken to be identical for through/shared through lane groups.

Table A.1 shows the number of turn lanes for each type of urban street in our paper and signal phasing for left-turn lanes (permitted or protected); these characteristics are not specified in the Florida Department of Transportation cost estimates and turn lane configurations are determined based on the road width. The table also shows how one-directional capacity is calculated for each road type as a function of lane-group capacity c_i and of the fractions τ_L and τ_R of traffic turning left and right, respectively.

Table A.1: Turn lane configurations and capacities for urban streets

Two directional no. of lanes	**One-directional lane configuration (L: left turn, R: right turn, T: through)**	**Signal phasing for left turns**	**One-directional capacity**
2 lanes, undivided	1 shared L/R/T	Permitted	cap_{LRT}
2 lanes, plus center turn lane	1 exclusive L 1 shared R/T	Protected	$(1-\tau_L)^{-1}(cap_{RT})$
4 lanes, undivided	1 shared L/T 1 shared R/T	Permitted	cap_{LRT}
4 lanes, plus center turn lane	1 exclusive L 1 exclusive T 1 shared R/T	Protected	$(1-\tau_L)^{-1}(cap_T + cap_{RT})$
4 lanes, divided*	1 exclusive L 2 exclusive T 1 exclusive R	Protected	$(1-\tau_L-\tau_R)^{-1}(cap_T)$
6 lanes, divided*	1 exclusive L 3 exclusive T 1 exclusive R	Protected	$(1-\tau_L-\tau_R)^{-1}(cap_T)$

Notes: Lane groupings for capacity determination are based on the guidelines in HCM Exhibit 16-5. cap_i is the capacity of each lane group, and τ_L and τ_R are the percentage of total traffic volume turning left and right, respectively. It is assumed that lane configurations are the same whether a road has 11-foot lanes or 12-foot lanes.
* Divided roads have a 22-foot median and combined with the 4-foot right shoulder, this results in sufficient width for the lane configurations shown above.

The saturation flow rates needed for equation (A.14) 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). 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:

$$s = s_0 N f_w f_{HV} f_{RT} f_{LT} PHF \tag{A.15}$$

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} and f_{LT} are the right-turn and left-turn adjustment factors, respectively (applicable only if vehicles in that lane group can make turns, to account 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 0.97 when lane width is 11 ft. Again, the percentage of heavy vehicles is assumed to be 5% as in the case of the other roads, leading to $f_{HV} = 0.95$. For exclusive through lane groups, $f_{RT} = f_{LT} = 1$; otherwise, the adjustment factor for turns is calculated based on Exhibit 16-7 and in the case of permitted phasing for left turns, Appendix C of HCM Chapter 16.[31] In general, f_{RT} and f_{LT} never exceed 1 and they decrease as the proportion of traffic in that lane group making turns in that direction increases. As in the case of freeways, the peak hour factor, PHF, is assumed to be 0.92.

Appendix B. Florida Department of Transportation (FDOT) road construction costs

B.1 Modification to road costs

The costs listed in Table 1 were modified slightly from the original Florida Department of Transportation data. In the original source data, the 2-lane undivided arterial had a cost per mile of $4,793,671.06, but this was oddly higher than the cost of a 3-lane undivided arterial with a center turn lane ($4,768,947.38). In examining the detailed data listing the quantities and prices of each input and comparing them between different road types, we found a typo in the number of 42” pipe culverts: the 2-lane undivided arterial required 5,056 units while the other undivided arterials required only 56 units. After reducing the number of 42” pipe culverts to 56 for the 2-lane undivided arterial, its cost per mile fell to $4,179,218.22.

[31] Calculating the left turn adjustment factor requires assumptions on the opposing traffic flow since this determines the opportunity for cars to make left turns; as in the case of two-lane highways, it is assumed that the opposing flow is equal to the approach flow.

B.2 Costs of traffic signals and interchanges

In our example roads, we assume that urban streets have signalized intersections every 1.0 mile and highways and freeways have interchanges with urban streets every 2.0 miles. When calculating construction costs, specific assumptions are required regarding the type of signal and interchange. Although there are no statewide estimates for signals from the Florida Department of Transportation, two of its district offices (District 3 and District 7) provide 2011 cost estimates for mast arm signals—where the signals are mounted on poles extending over the roadway—for two-lane, four-lane and six-lane roads. To be consistent with the statewide cost estimates for roads, we include costs related only to construction, maintenance of traffic, and mobilization. Since District 3's cost estimates for urban roads are lower (averaging about 91 percent of the statewide estimates) while District 7's cost estimates are higher (about 113 percent of the statewide estimates), the average of the two districts' cost estimates is used. The signal costs are assumed to be the same for roads with 12-foot and 11-foot lane widths and are listed in Table 2 of the paper.

Only District 7 provides an up-to-date cost estimate for interchanges between freeways/highways and urban streets, specifically a single point urban interchange (SPUI) that costs $21,467,980. According to the St. Louis District of the Missouri Department of Transportation:

> This interchange, also know as an X-interchange or an Urban Diamond is being used extensively in the reconstruction of existing freeways as well as constructing new freeways… The name "Single Point" refers to the fact that all through traffic on the arterial street, as well as the traffic turning left onto or off the interchange, can be controlled from a single set of traffic signals.[32]

Since District 7's cost estimates for urban roads are 113 percent of the statewide estimates, on average, the SPUI cost estimate is divided by 113 percent and halved to arrive at a per mile cost (since it is assumed that there is an interchange every two miles). There is no guidance as to what type of road the SPUI cost estimate applies to, so we use this per mile cost estimate for the four-lane divided highway with 12-foot lanes. This cost is adjusted accordingly for the highways/freeways with two- and six-lanes by the ratio of those roads' construction costs (see Table 2 of the paper). For roads with 11-foot lane widths, we multiply the SPUI costs of the

[32] Missouri Department of Transportation (2013).

corresponding 12-foot roads by the same factor used to adjust road construction costs as mentioned in the paper.

Appendix C. Assumptions and detailed results of the investment balance computations

Table C1 gives the detailed results portrayed in Figures 5 and 6 of the text, where we depict the investment balance condition for the 24 road types below.

Table C1. Application of first-order cost-minimizing conditions for 24 road types

No. of lanes (two-directions)	Road type	Lane width (feet)	Free-flow speed (mi/h)	$\frac{\varepsilon_{\rho,VK}}{\varepsilon_{\rho,SF}}$	$\frac{\sum_t q_t V_t \cdot (mecc)_t}{\sum_t q_t V_t c_t \cdot (\varepsilon_{S,SF})_t}$			Critical value of V/V_K	Avg peak speed at critical value (mi/h)**
					V_p/V_K =0.3	V_p/V_K =0.8	V_p/V_K =1.0		
2 lanes, undivided	Urban street	12	35.8	0.34	0.05*	0.32*	0.66	0.81	24.59
		11	34.4	0.34	0.06*	0.33*	0.65	0.80	22.80
	Two-lane highway	12	52.5	0.34	0.01*	0.08*	1.02	0.93	46.38
		11	47.1	0.34	0.01*	0.14*	0.85	0.90	39.30
2 lanes, ctr turn lane	Urban street	12	35.8	0.34	0.05*	0.32*	0.66	0.81	24.59
		11	34.4	0.34	0.06*	0.33*	0.65	0.80	22.80
4 lanes, undivided	Urban street	12	36.5	0.34	0.04*	0.31*	0.67	0.82	25.35
		11	35.1	0.34	0.05*	0.32*	0.66	0.80	23.88
	Multilane highway	12	51.8	0.34	0.01*	0.08*	1.00	0.93	45.26
		11	49.9	0.34	0.01*	0.11*	0.93	0.92	42.79
4 lanes, ctr turn lane	Urban street	12	36.5	0.34	0.04*	0.31*	0.67	0.82	25.35
		11	35.1	0.34	0.05*	0.32*	0.66	0.80	23.88
4 lanes, divided	Urban street	12	36.5	0.34	0.04*	0.31*	0.67	0.82	25.35
		11	35.1	0.34	0.05*	0.32*	0.66	0.80	23.88
	Multilane highway	12	53.4	0.34	0.00*	0.07*	1.05	0.93	47.81
		11	51.5	0.34	0.01*	0.09*	0.98	0.93	44.77
	Freeway	12	65.5	0.34	0.00*	0.01*	1.56	0.97	62.74
		11	63.6	0.34	0.00*	0.02*	1.48	0.96	60.87
6 lanes, divided	Urban street	12	36.8	0.34	0.04*	0.30*	0.67	0.82	25.77
		11	35.4	0.34	0.05*	0.32*	0.66	0.81	23.98
	Multilane highway	12	53.4	0.34	0.00*	0.07*	1.05	0.93	47.81
		11	51.5	0.34	0.01*	0.09*	0.98	0.93	44.77
	Freeway	12	67.0	0.34	0.00*	0.01*	1.63	0.97	64.71
		11	65.1	0.34	0.00*	0.01*	1.54	0.97	62.21

Notes: * Indicates that the ratio of construction cost elasticities exceeds the ratio of marginal user costs (a situation favoring investment in capacity relative to that in free-flow speed).
** Calculated from the estimated travel time function using equation (12) and the parameters in Table 4.

In Section 5.2, where we examine the investment balance condition for "freeways" and "arterials" in various urban areas, we combine data on road mileage from the Federal Highway Administration (2013) with data on lane-miles from Schrank et al. (2012b) to obtain the average number of lanes. To match up the roads in the two datasets, we assume that "freeways" are what the FHWA classifies as "Interstates" and "Other freeways and expressways", while "arterials" are the FHWA's "Other principal arterial" and "Minor arterial". Both datasets report vehicle miles traveled (VMT) and we use this to see if there are any large discrepancies between the two datasets; for most urban areas, the difference in VMT between the two datasets is very small (less than 5%) but for several urban areas (including Chicago, San Francisco and Washington D.C.) the discrepancy is more than 10% for either freeways, arterials, or both.

The average number of lanes for freeways, averaged across urban areas, is 7.1 and 6.1 for very large and large urban areas, respectively, and for arterials it is approximately 3.3 for both types of urban areas. We use the average number of lanes for each urban area to estimate capacity, using the HCM capacities calculated from previous sections as a basis. Specifically, freeway capacity for each urban area is extrapolated from the capacity of a six-lane freeway with 12-foot lanes:

Two-directional capacity = (12,763.31)·(*Avg. no. of lanes*)/6

For arterials, a similar procedure is implemented where if the average number of lanes for an urban area is between 2 and 4, capacity is interpolated based on the HCM capacities of two-lane undivided and four-lane divided urban streets, and if the average number of lanes exceeds four, capacity is interpolated from four-lane divided and six-lane divided urban streets. The only exceptions were if the difference in VMT was greater than 10% as mentioned earlier; for these urban areas we used the capacities of a six-lane freeway and a four-lane divided urban street.

Appendix D. Sensitivity analysis using a refined peak model

We conduct a sensitivity analysis by modifying our assumptions on the peaking pattern. Instead of a constant volume V_p during peak hours, we now assume that there is a "peak peak" period during which traffic volumes are highest, flanked by two "shoulder peak" periods, each with the same traffic volume. The off-peak period is still assumed to have constant traffic volume V_o, with duration F = 12 hours. We assume that the duration of the period encompassing the peak peak and shoulder peak is still P = 4 hours, with V_p now representing the average traffic volume during this period. The duration of the peak peak is set at 2 hours, and the duration of each shoulder is 1 hour. Again using the notation that $v = V/V_K$, we have

$$v_p \equiv \frac{V_p}{V_K} = \frac{V_p^s + V_p^p}{2V_K} \tag{D.1}$$

where the superscripts s and p indicate the shoulder and peak, respectively, within the peak period.

Defining $\sigma = V_p^p / V_p^s$, we use Exhibits A-2 to A-4 in Schrank et al. (2012b) – which depict traffic distribution profiles over the course of a weekday – to approximate σ for three possible cases: “low congestion” (σ = 10/7), “moderate congestion” (9/7), and “severe congestion” (8/7). To classify roads into one of these three cases, Schrank et al. calculate the speed reduction factor as the ratio of the average peak period speed to free-flow speed. Freeways are classified under “low to no congestion” if the speed reduction factor is 90-100%, “moderate” if it is 75-90%, and “severe” if it is below 75%. The corresponding ranges for arterials are 80-100%, 65-80% and below 65%.

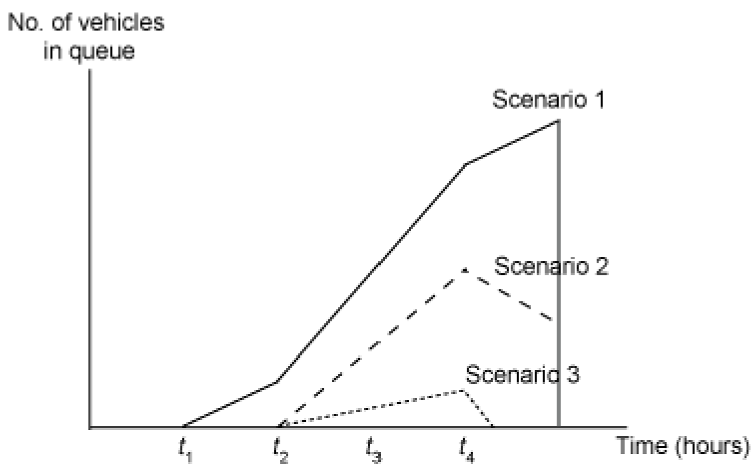
Using data on average free-flow speed and peak speed for “very large” and “large” urban areas as seen in Table 5, we apply the above classifications to these urban areas. It should be noted that the classifications from Schrank et al. (2012b) are for individual roads, but we are applying them to the urban area as whole for lack of better data. It turns out that representative freeways in most urban areas are classified as “moderate congestion.” (Los Angeles just misses the definition of “severe congestion”; thus, we can see that averaging over the entire urban area can make even a notoriously-congested city seem not as congested). Jacksonville is the only urban area of the ones we show that has a “low congestion” classification for freeways. Arterials in all seven urban areas are classified under “low congestion”.

With this refined peak model, there are now three possible scenarios for which the average queuing delay for the entire four-hour period can be calculated, as illustrated in Figure D1. The first is when the queue begins during the first shoulder and continues building up throughout the peak peak and second shoulder periods ($V_p^p > V_p^s > V_K$). The average queuing delay is

$$\overline{D}_p = 2(v_p - 1)\,. \tag{D.2}$$

This happens to be the same delay as that in the original peaking model in the paper (see Section 4), but this is a coincidence due to our specific assumptions regarding the durations of the peak peak and shoulder periods. As in the original model, it is assumed that off-peak travelers do not experience any queuing delay.

Figure D1. Queuing scenarios (refined peak model)



Note: t_1, t_2 and t_4 indicate the starting times of the first shoulder, the peak peak, and the second shoulder, respectively.

The second scenario occurs when there is no queuing during the first shoulder, the queue begins during the peak peak period, and the queue does not fully dissipate by the end of the second shoulder. Using $y \equiv 2(v_p^p - 1)/(1 - v_p^s) + 2$ to denote the duration of the queue (see Ng and Small [2012] for details on how this is derived), this scenario occurs when $y > 3$, i.e., the duration of the queue is greater than the duration of the peak peak plus second shoulder periods. Therefore, this scenario occurs when $V_p^p > V_K > V_p^s$ and $2(v_p^p - 1)/(1 - v_p^s) > 1$. The average queuing delay to peak-period travelers under this scenario is

$$\overline{D}_p = \left(\frac{2\sigma v_p}{1+\sigma} - 1\right) + \left(\frac{1}{4(1+\sigma)}\right)\left(\frac{2v_p}{1+\sigma} - 1\right). \tag{D.2}$$

Finally, in the third scenario, there is no queuing during the first shoulder, queuing starts during the peak peak period, but the queue now dissipates by the end of the second shoulder. That is, $V_p^p > V_K > V_p^s$ and $2(v_p^p - 1)/(1 - v_p^s) < 1$. The average queuing delay is.

$$\overline{D}_p = \left[\frac{2\sigma v_p - (1+\sigma)}{(1+\sigma)^2}\right]\left[\sigma + \left(\frac{2\sigma v_p - (1+\sigma)}{1+\sigma - 2v_p}\right)\right]. \tag{D.2}$$

As with the original peaking model with uniform traffic volume, travel times are calculated for each of the 24 road types listed in Table 2 using the HCM procedures detailed in Appendix A. Whenever there is queuing (i.e. whenever $V_p^p > V_K$), the average queuing delay from the refined peak model is added to these travel times. Since only the 'low" and "moderate" congestion classifications are relevant to the urban areas in our study, we estimate two versions of the travel time function (equation [12]); the estimated parameters are shown in Table D1.

Table D1. Estimates of modified Akçelik function (refined peak model)

	"Low" congestion (σ = 10/7)		"Moderate" congestion (σ = 9/7)	
Parameter	Estimate	Standard error	Estimate	Standard error
γ_1	0.2791	0.00096	0.2858	0.00096
γ_2	11.754	1.0391	24.4105	3.3244
γ_3	-0.0942	0.0024	-0.1202	0.0038
Observations	3,624		3,624	
R-squared	0.9903		0.9895	

Figures D2 and D3 compare predicted travel times from the refined peak model at "low" and "moderate" congestion to those from the original model used in the paper, for a four-lane divided urban street (Figure D2) and a four-lane freeway (Figure D3). We can see that the differences among the three prediction models are not very large. They are most marked around v_p = 1; at any given v_p, the "low" congestion model has the highest travel time because it has the highest ratio of peak peak volume to shoulder peak volume. This may seem counterintuitive, but it may make more sense if we infer v_p from a given travel time (as we do to obtain the peak volume-capacity ratios in Table D2 below). For example, if the average travel time is 30 minutes on a four-lane divided urban street, the "low" congestion model would give us the lowest v_p, while the original peak model with constant volume throughout the four-hour peak period would give us the highest v_p.

Figure D2. Travel times for a four-lane divided urban street (S_F = 36.5 mi/h) for the refined and original peak models

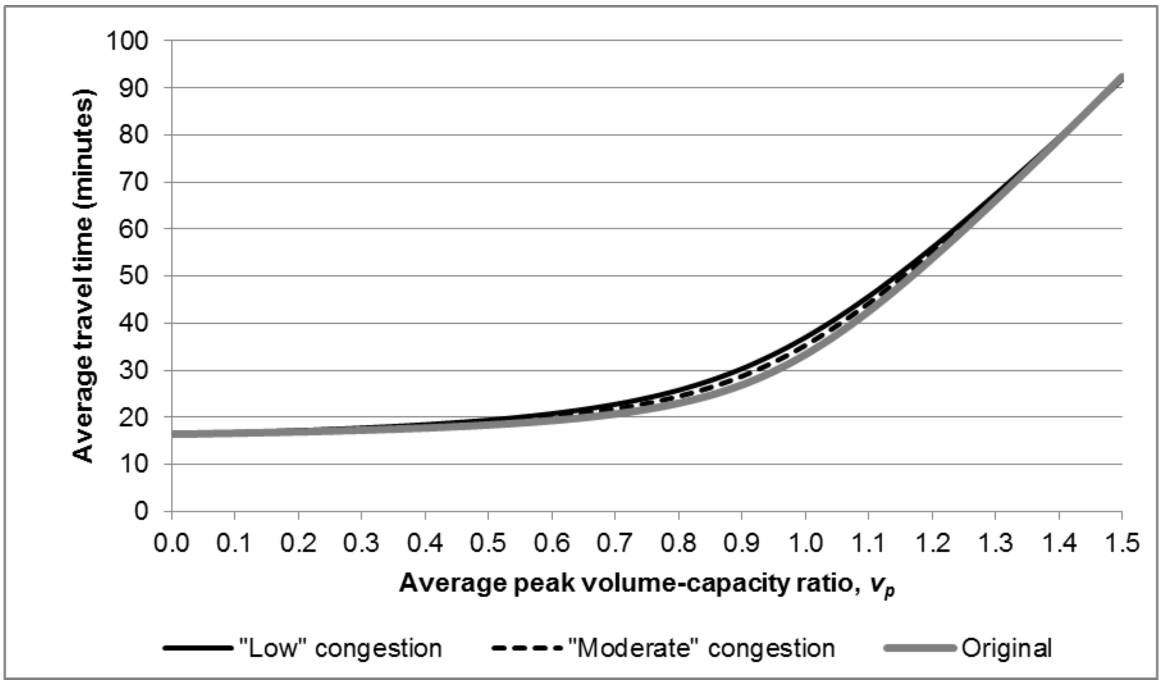
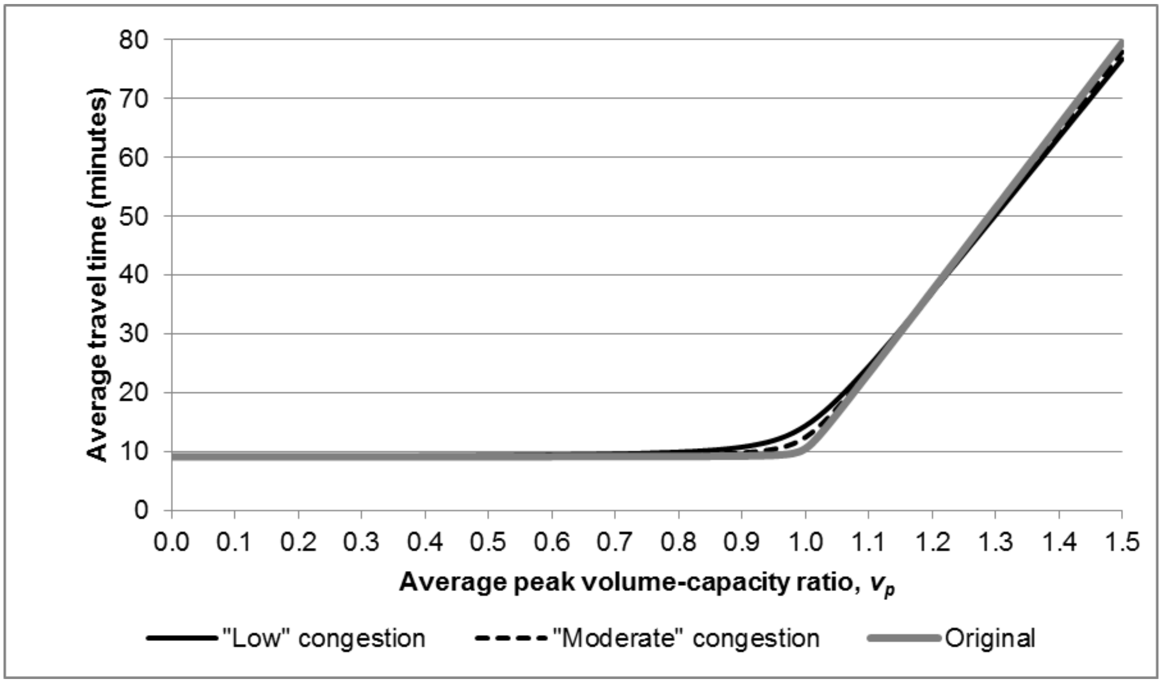


Figure D3. Travel times for a four-lane freeway (S_F = 65.5 mi/h) for the refined and original peak models



We now redo the calculations for Tables 5 and 6 using the estimates from Table D1, once again applying the adjustment factors discussed in Section 5.2 of the paper. The capital cost parameters and assumptions on road length, ratio of peak to off-peak volumes, value of time, etc, are the same as those in the paper. Table D2 is the same as Table 5 of the paper except it uses calculations from the refined peak model. Compared to the original model, the refined peak model leads to several differences in results. Urban areas now are calibrated to have lower peak volume-capacity ratios, with more variation between the different urban areas. The ratio of marginal user costs – calculated from the RHS of equation (5c) – is generally lower (though not always, e.g., in the case of St. Louis arterials), as the refined peak travel time functions usually lead to lower values in the numerator (the *mecc* calculations) and higher values in the denominator (the travel time multiplied by the elasticity of speed with respect to free-flow speed calculations).

Table D2. Investment balance for average road conditions in seven urban areas, 2011 (refined peak model)

	Very large areas	Large areas

		Los Angeles	Dallas-Fort Worth	Miami	Boston	Denver	St. Louis	Jackson-ville
Freeways:								
	Average no. of lanes	8.7	5.8	6.7	6.4	5.8	6.5	5.8
	Free-flow speed, S_F (mi/h)	64.6	64.1	64.0	63.4	62.3	56.0	63.4
	Peak speed, S_p (mi/h)	48.6	54.0	56.7	54.2	50.9	44.4	58.9
*	Congestion classification	Mod.	Mod.	Mod.	Mod.	Mod.	Mod.	Low
*	Peak volume-capacity ratio, V_p/V_K	0.993	0.961	0.935	0.951	0.964	0.937	0.747
	Ratio of construction cost elasticities	0.34	0.34	0.34	0.34	0.34	0.34	0.34
*	Ratio of marginal user costs	0.83	0.43	0.27	0.36	0.47	0.39	0.10
*	Imbalance (+ favors investment in S_F)	-0.50	-0.09	0.07	-0.03	-0.14	-0.05	0.24
Arterials:								
	Average no. of lanes	3.6	3.7	4.6	2.3	3.5	3.2	3.7
	Free-flow speed, S_F (mi/h)	43.7	39.1	39.2	36.0	38.0	34.9	43.3
	Peak speed, S_p (mi/h)	37.4	33.1	31.7	29.5	32.1	29.8	37.4
*	Congestion classification	Low	Low	Low	Low	Low	Low	Low
*	Peak volume-capacity ratio, V_p/V_K	0.601	0.540	0.613	0.540	0.525	0.455	0.577
	Ratio of construction cost elasticities	0.34	0.34	0.34	0.34	0.34	0.34	0.34
*	Ratio of marginal user costs	0.15	0.15	0.19	0.17	0.15	0.13	0.14
*	Imbalance (+ favors investment in S_F)	0.19	0.19	0.14	0.16	0.19	0.20	0.20

Notes: * indicates that this row differs from Table 5 in the paper as a result of the refined peak model.
The imbalance is calculated as the ratio of construction cost elasticities minus the ratio of marginal user costs. Urban areas are classified as having either "moderate congestion" or "low congestion".
Sources: Schrank et al. (2012b), FHWA (2013), and authors' calculations.

As a result, the imbalance for freeways becomes less negative; it is even positive for the average freeway in Miami and Jacksonville, thus favoring incremental investment in free-flow speed rather than capacity for the average freeway in these two urban areas. The case for capacity expansion is still strongest for Los Angeles, and weakest for Jacksonville (as argued in the paper). For arterials, the imbalances here are very similar to those in Table 5, largely because arterials have lower v_p than freeways, and differences in travel time (and hence differences in marginal user costs) are smaller across the different peaking models when v_p is low.

The same patterns are shown in Table D3, which can be compared to Table 6 of the paper. Compared to the results in Table 6, the benefit-cost ratios for freeway capacity expansion with the refined peak model are substantially lower (about 20-60% of those in the original model), while the benefit-cost ratio for free-flow speed investment is higher for representative freeways in all urban areas except Jacksonville. Meanwhile, the benefit-cost ratios for arterials (for both capacity and free-flow speed investment) are generally about half or more of those from the original model.

Table D3. Absolute benefit-cost ratios from incremental investments, assuming Florida capital costs, 2011 (refined peak model)

		Very large areas				Large areas		
		Los Angeles	Dallas-Fort Worth	Miami	Boston	Denver	St. Louis	Jackson-ville
Freeways:								
	Free-flow speed, S_F (mi/h)	64.6	64.1	64.0	63.4	62.3	56.0	63.4
	Capacity, V_K (veh/h)	18,519	12,307	14,268	13,616	12,382	13,736	12,322
	Capital cost, ρ (1000 $ per year per mi)	2,710	2,272	2,409	2,336	2,200	2,016	2,243
*	B/C: incr. invest. in V_K	27.9	10.4	6.5	9.4	12.8	15.1	1.6
*	B/C: incr. invest. in S_F	11.2	8.2	8.2	8.7	9.1	13.1	5.6
Arterials:								
	Free-flow speed, S_F (mi/h)	43.7	39.1	39.2	36.0	38.0	34.9	43.3
	Capacity, V_K (veh/h)	3,216	3,337	4,284	1,589	3,123	2,751	3,393
	Capital cost, ρ (1000 $ per year per mi)	823	730	811	487	686	587	832
*	B/C: incr. invest. in V_K	2.4	2.8	5.1	2.6	2.8	2.3	2.2
*	B/C: incr. invest. in S_F	5.4	6.4	9.0	5.1	6.3	5.9	5.4

Notes: * indicates that this row differs from Table 6 in the paper as a result of the refined peak model.
B/C is the benefit cost ratio from incremental investment in capacity (V_K) and free-flow speed (S_F) calculated using equations (15a) and (15b), respectively.

Thus, we can see that refining and disaggregating the peak period generally moves the analysis in favor of investment in free-flow speed rather than capacity for freeways, while the differences in the results from the two models are smaller for arterials for the reasons stated earlier.

Additional references (aside from those listed in the paper)

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

Missouri Department of Transportation, 2013. Single Point Urban Interchanges. St. Louis District. http://www.modot.org/stlouis/links/SinglePointUrbanInterchanges.htm