Optimizing Road Capacity and Type

Kenneth A. Small
Chen Feng Ng

June 1, 2013

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 most large urban areas provide too high a free-flow speed relative to capacity, thus making the case for reexamination of typical design practice.

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 road capital. While naturally not every issue of interest can be captured with just one additional

---

[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 two cases, however, 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).

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

variable, the advantages of tractability and transparency make this an attractive way to begin bringing the analysis of road types into mainstream transportation economics.

To implement the model, we use empirical data to estimate both investment costs and user costs as functions of the two design variables (capacity and free-flow speed). We estimate a construction-cost function using data on costs of various road types along with their free-flow speeds and capacities. We estimate a user-cost function from information about speeds and flows of different road types, differentiated by free-flow speed,[4] which we supplement with a queuing analysis to account for situations where input flow exceeds capacity.

The result is a continuous, differentiable total cost function which permits standard investment analysis. The model produces the familiar criterion for incremental investment in capacity, and a new criterion for incremental investment in free-flow speed. We combine these criteria to examine how to recognize 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? We examine this balance condition for 24 standard road types under hypothetical conditions, and for representative freeways and arterials for 47 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 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.

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

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, more than three times that for expanding capacity (roughly 1.4 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),[5] 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).

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:

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

where V_K and S_f are design capacity and free-flow design speed, respectively, K 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.

[5] 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.

the time after which the structures and improvements (but not the land) have lost all their value. We assume that K 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 during time interval t consists solely of time costs measured at a constant value of time, α. User time depends both on free-flow speed and on congestion, the latter via the volume-capacity ratio:

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

where t is a time interval (of duration q_t), V_t is traffic volume, 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 and concave in volume-capacity ratio.

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

$$\begin{aligned} C\left(V \mid V_K, S_f\right) &= \rho\left(V_K, S_f\right) + \sum_t q_t U_t\left(V_t \mid V_K, S_f\right) \\ &= \frac{r}{1-e^{-r\Lambda}} K(V_K, S_f) + rA(V_K, S_f) + \alpha \sum_t \frac{q_t V_t}{S_t\left(\frac{V_t}{V_K}, S_f\right)} \end{aligned} \tag{3}$$

where $V=\{V_t\}$ 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:

$$\begin{aligned} \tilde{C}(V) &= \min_{V_K, S_0} C\left(V \mid V_K, S_f\right) \\ &= \min_{V_K, S_f} \left[\rho\left(V_K, S_f\right) + \alpha \sum_t \frac{q_t V_t}{S_t\left(\frac{V_t}{V_K}, S_f\right)} \right]. \end{aligned}$$

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. The first of these investment rules is standard.[6] The second is new to this paper, but obviously follows the same logic.

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).[7] 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:

$$\varepsilon_{\rho,VK} = \frac{1}{\rho} \sum_t q_t V_t \cdot (mecc)_t \equiv \frac{\tilde{R}}{\rho} \tag{5a}$$

[6] 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.

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

$$\varepsilon_{\rho,Sf} = \frac{1}{\rho}\sum_t q_t V_t c_t \cdot \left(\varepsilon_{S,Sf}\right)_t \tag{5b}$$

where $(mecc)_t \equiv V_t \cdot (\partial c_t / \partial V_t)$ 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, and $\varepsilon_{S,Sf}$ is the elasticity of the function $S(\cdot)$ with respect to S_f. (This last elasticity may vary by time period.) The quantity $\tilde{R}$ is imputed revenues from a hypothetical congestion toll set equal to $mecc_t$ in each period when traffic is given by vector V.[8] 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 from these investment criteria. However, we are more confident in their ratio, which is based on 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):

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

This implication of the first-order conditions makes clear that if congestion is large, so that *mecc* exceeds $c \cdot \varepsilon_{S,Sf}$ for a large portion of the time, investment in capacity will be favored relative to that in free-flow speed. On the other hand, if peak traffic congestion is not severe and off-peak travel 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

[8] 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).

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:

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

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. We then have $mecc=\alpha\cdot[(1/S)-(1/S_f)]$, $\varepsilon_{S,Sf} = S/S_f$, 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, $U^0 \equiv \alpha\sum_t q_t V_t / S_f$, 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.[9]

With more realistic models of speed determination, the more general 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 give 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 would provide evidence for this hypothesis.

[9] 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 $mecc=\alpha b\cdot[(1/S)-(1/S_f)]$ 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.

Alternatively, one can consider the tradeoff between free-flow speed and capacity inherent in any particular set of incremental plans or planning guidelines by rewriting (5c) as:

$$\left(-\frac{dS_f / S_f}{dV_K / V_K}\right)_\rho = \frac{\tilde{R}}{\sum_t q_t V_t c_t \cdot \left(\varepsilon_{S,Sf}\right)_t} . \qquad (5c')$$

Suppose, for example, a particular road design could be modified at no change in cost so as to increase free-flow speed 2 percent by sacrificing one percent of capacity. This change would be beneficial if the ratio on the right-hand side of (5c′) (computed with the proposed design in place) is less than 2, whereas a trade in the opposite direction would be beneficial if that ratio is greater than 2. As a reminder, all these types of statements presume that there is a continuum of possible designs and that the resulting costs are smooth functions.

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 discussed in previous sections. 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. Thus, capacity and free-flow speed show sufficient independent variation that we expect to see some possibilities for substitution of the type highlighted in equation (5c′).

Table 1. FDOT cost estimates (in 2011 prices)

Description	No. lanes	Bike lane (width)	Median (width)	Shoulders (inside & outside)	Cost per mile (mill. $)
Undivided arterial	2	4 ft	---	---	4.794
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/specificationsoffice/).

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.[10] The HCM has separate procedures for freeways, urban streets, and "highways" (which have design standards

[10] 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).[11] 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).[12] 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 $K(V_K,S_f)$ describing initial construction cost.

[11] 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."

[12] 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/ inter-change cost (thousands of $)	Total cost per mile (thousands of $)
2 lanes, undivided	Urban street	12	42.1	35.8	1,277.6	4,794	155	4,949
		11	40.2	34.4	1,245.1	4,647	155	4,802
	Two-lane highway	12	52.5	52.5	3,112.4	4,794	6,716	11,510
		11	47.1	47.1	3,112.4	4,647	6,511	11,158
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	7,190	12,323
		11	49.9	49.9	7,169.7	4,909	6,877	11,786
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	9,979	17,102
		11	51.5	51.5	7,284.6	6,854	9,603	16,457
	Freeway	12	65.5	65.5	8,455.0	8,875	12,433	21,308
		11	63.6	63.6	8,386.8	8,353	11,702	20,055
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	11,189	19,175
		11	51.5	51.5	10,926.9	7,639	10,703	18,342
	Freeway	12	67.0	67.0	12,763.3	9,858	13,811	23,668
		11	65.1	65.1	12,661.0	9,215	12,910	22,125

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.

These estimates imply construction costs per lane-mile, for 12-foot lanes, of roughly \$4.0–5.3 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 translog function to estimate the relationship between construction cost per mile (denoted by K, 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:

$$\begin{aligned}\ln K &= \beta_0 + \beta_1 \ln(S_f / \overline{S_f}) + \beta_2 \ln(V_K / \overline{V_K}) + 0.5\beta_3 \ln(S_f / \overline{S_f})^2 \\ &\quad + 0.5\beta_4 \ln(V_K / \overline{V_K})^2 + \beta_5 \ln(S_f / \overline{S_f}) \ln(V_K / \overline{V_K}) + \varepsilon\end{aligned} \tag{8}$$

The sample means for free-flow speed and capacity are 45.80 mi/h and 5,589 veh/h, respectively.

The regression results, using ordinary least squares on 24 observations, are shown in Table 3. Although none of the second-order terms are statistically significant (at a five-percent level), we prefer the second specification because it allows for varying elasticities, even though the estimated extent of variation is not large. Using that specification, the implied elasticities of construction cost with respect to free-flow-speed and capacity are

$$\varepsilon_{K,Sf} = \beta_1 + \beta_3 \ln\left(S_0 / \overline{S_f}\right) + \beta_5 \ln\left(V_K / \overline{V_K}\right); \quad \varepsilon_{K,VK} = \beta_2 + \beta_4 \ln\left(V_K / \overline{V_K}\right) + \beta_5 \ln\left(S_f / \overline{S_f}\right).$$

As indicated by the first two coefficients of the right column, these elasticities are 1.36 and 0.40, respectively, when calculated at the sample means. 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).[13] 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.

[13] 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.

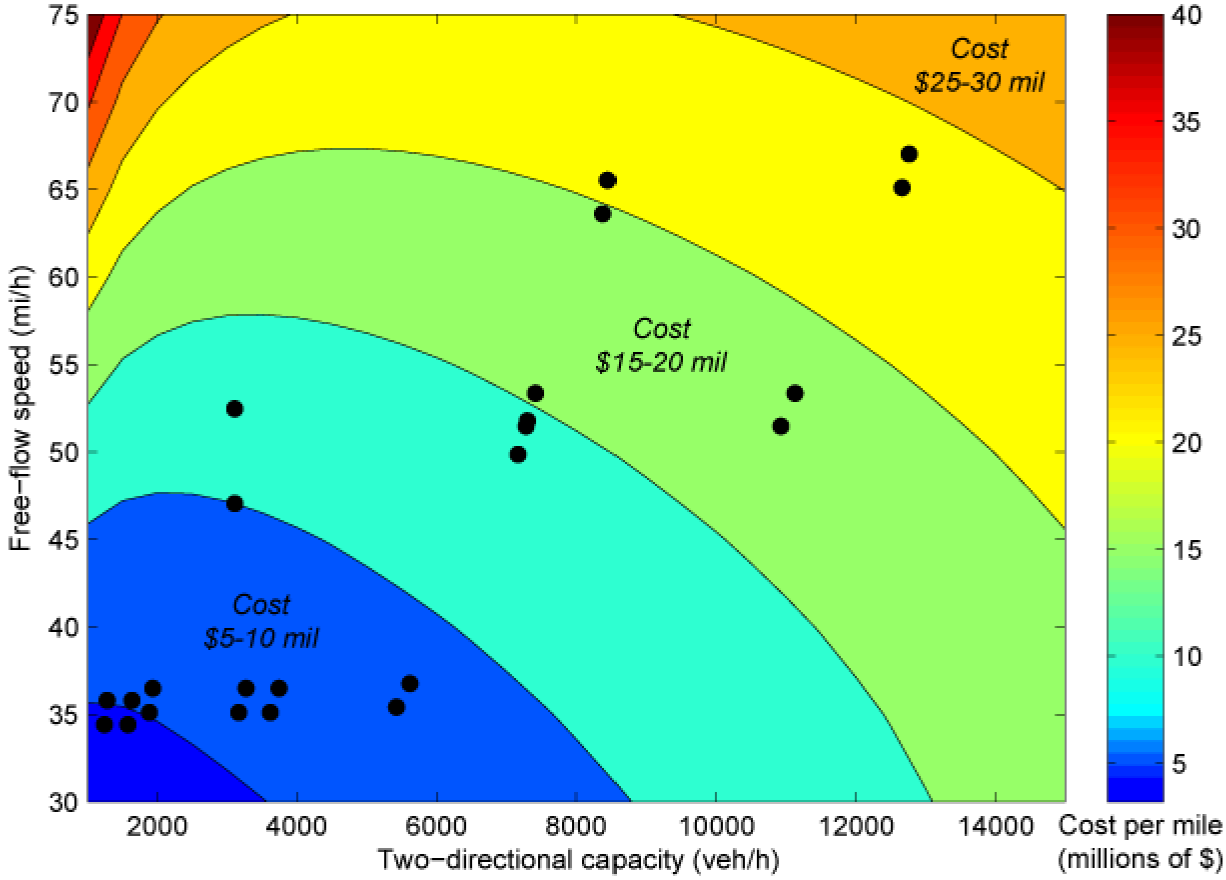
Table 3. Construction cost regression results

Variables	ln *K*	ln *K*
$\ln S_f - \ln \overline{S_f}$	1.4401***	1.3552***
	(0.136)	(0.153)
$\ln V_K - \ln \overline{V_K}$	0.3314***	0.3997***
	(0.044)	(0.068)
$0.5(\ln S_f - \ln \overline{S_f})^2$		0.7975
		(1.797)
$0.5(\ln V_K - \ln \overline{V_K})^2$		0.3800*
		(0.218)
$(\ln S_f - \ln \overline{S_f})(\ln V_K - \ln \overline{V_K})$		-0.8708
		(0.520)
Constant	9.3192***	9.3261***
	(0.021)	(0.038)
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. An exception occurs at extremely low capacities combined with high free-flow speeds, situations that are unrealistic and for which we neither have observations nor wish to do simulations.

Figure 1. Contour plot of predicted costs using translog 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 (K) 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.[14] Denoting these percentages as x (expressed as a decimal), we can express the right-of-way acquisition cost as a fraction of construction cost: $A = K\cdot[x/(1\text{-}x)]$. The annualized capital cost per mile from equation (1) can therefore be rewritten as:

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

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

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:

$$\begin{aligned} \ln \rho(V_K, S_f) = {} & \ln \kappa + \beta_0 + \beta_1 \ln(S_f / \overline{S_f}) + \beta_2 \ln(V_K / \overline{V_K}) + 0.5\beta_3 \ln(S_f / \overline{S_f})^2 \\ & + 0.5\beta_4 \ln(V_K / \overline{V_K})^2 + \beta_5 \ln(S_f / \overline{S_f}) \cdot \ln(V_K / \overline{V_K}) \end{aligned} \tag{10}$$

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 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. 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.

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 (Federal Highway Administration 2009, 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 (3) can be expressed in terms of travel time (T) for a road of length L, $T_t \equiv L/S_t$, we need travel time to depend 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. We therefore seek a functional form that can adequately represent the results of our more detailed calculations. 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:

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

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.[15]

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 the a single choke point; and (ii) empirically, S_f is positively correlated with (J_a/V_K). The modified Akçelik function is:

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

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).

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). They do this for a variety of road types with 12-foot lanes.

[15] When the "delay parameter" J_a is zero, this equation simplifies to $T=T_f$ for $v\leq1$ and $T=T_f+(1/2)P\cdot[v-1]$ for $v>1$.

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

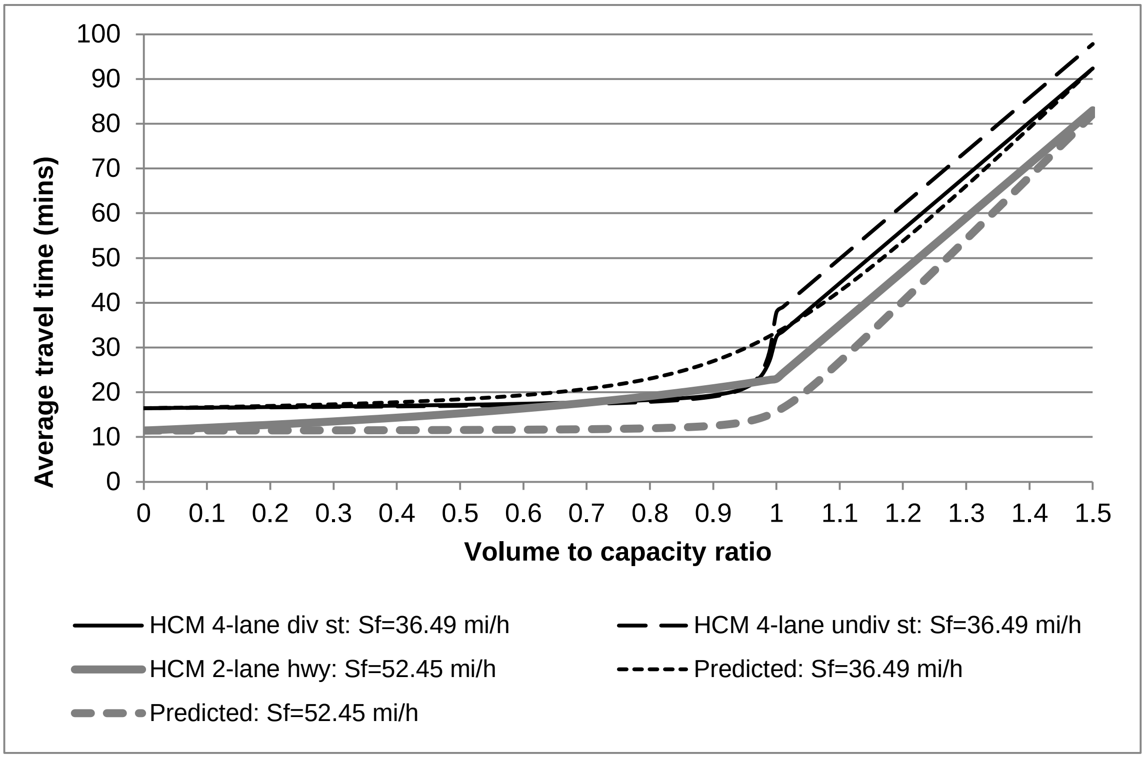


Figure 3. Travel times for a four-lane divided highway and freeway

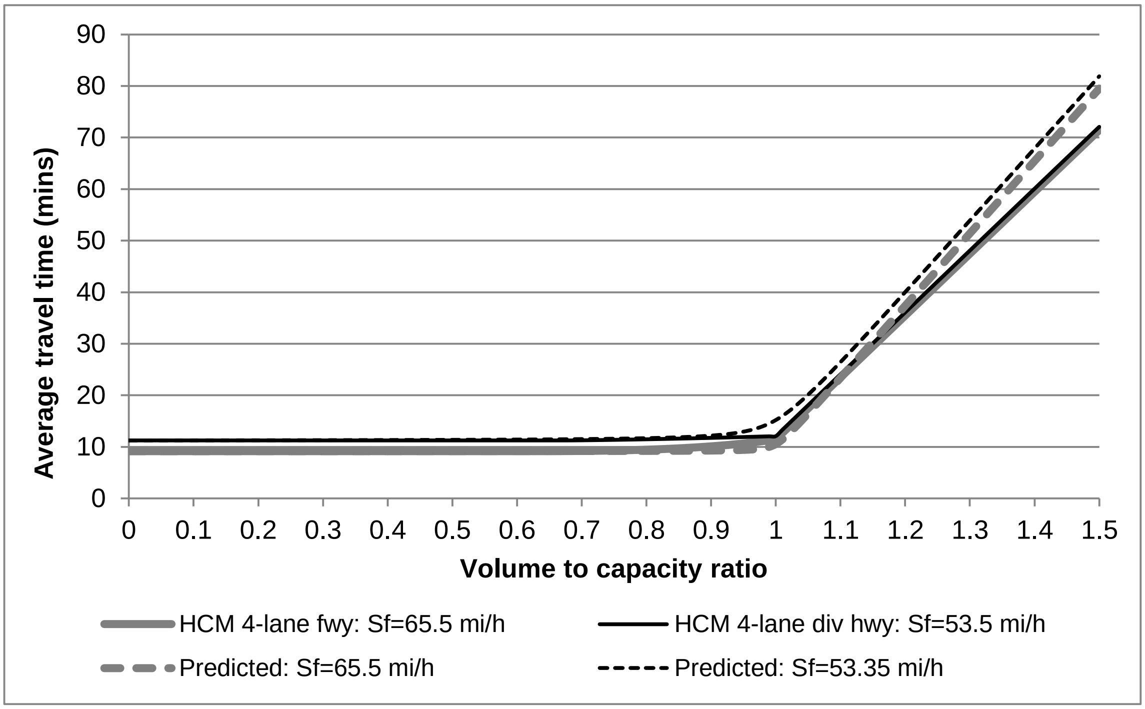
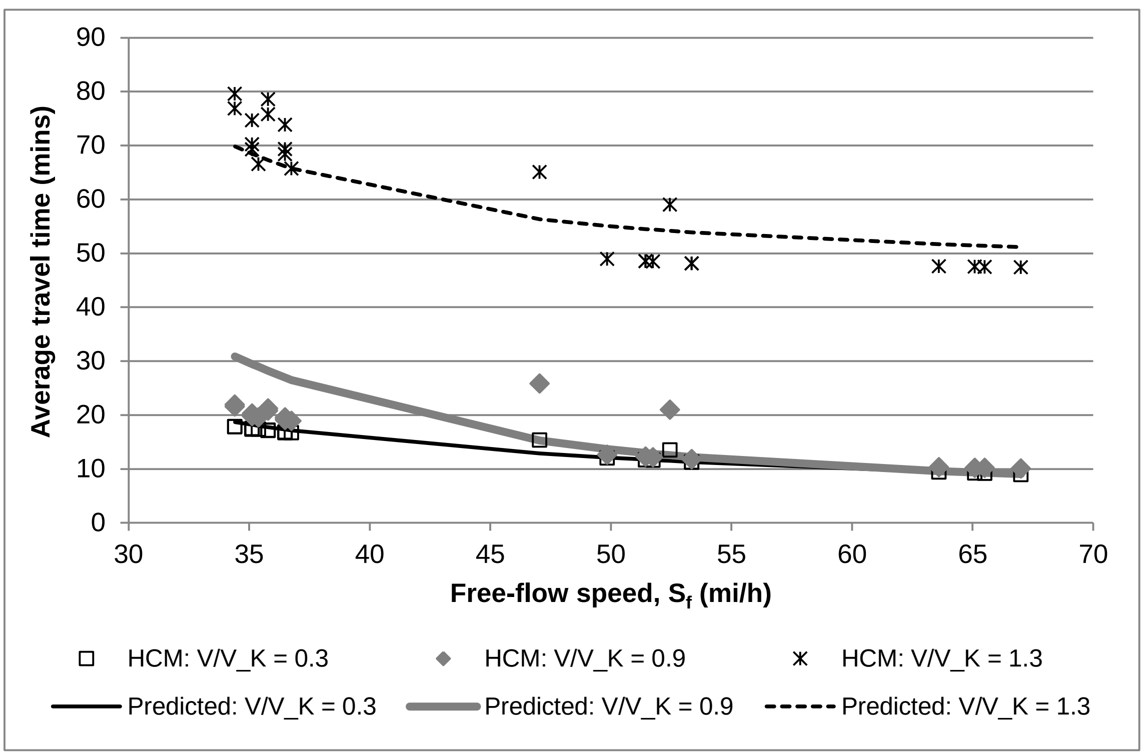


Figure 4. Travel times as a function of free-flow speed, for selected values of volume-capacity ratio



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). This guarantees that second-order conditions for cost minimization are met, so we do not need to explicitly derive and calculate values for those conditions.

The derivatives of the modified Akçelik function lead to the following values needed to calculate equations (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.[16]

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 equations (5c). In Section 5.1 we consider a wide selection of roads and traffic levels, in order to explore the range of conditions when each type of road is appropriate. In Section 5.2 we look at empirical data to see whether representative roads in various cities would better serve their areas with a different type of design. In Section 5.3, we go further and examine the absolute criteria for investing in capacity or free-flow speed, i.e. equations (5a-b), for the same sample of cities and for a hypothetical example illustrating the possibility of trading off free-flow speed against capacity.

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 for peak volume-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), the ratio of peak to off-peak volume (V_p/V_o=1.25), and other assumptions taken from Ng and Small (2012).[17] We believe these assumptions are relatively favorable to investment in free-flow speed; in particular, many congested cities probably have considerably higher values of V_p/V_o.[18]

---

[16] 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 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.

[17] 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.

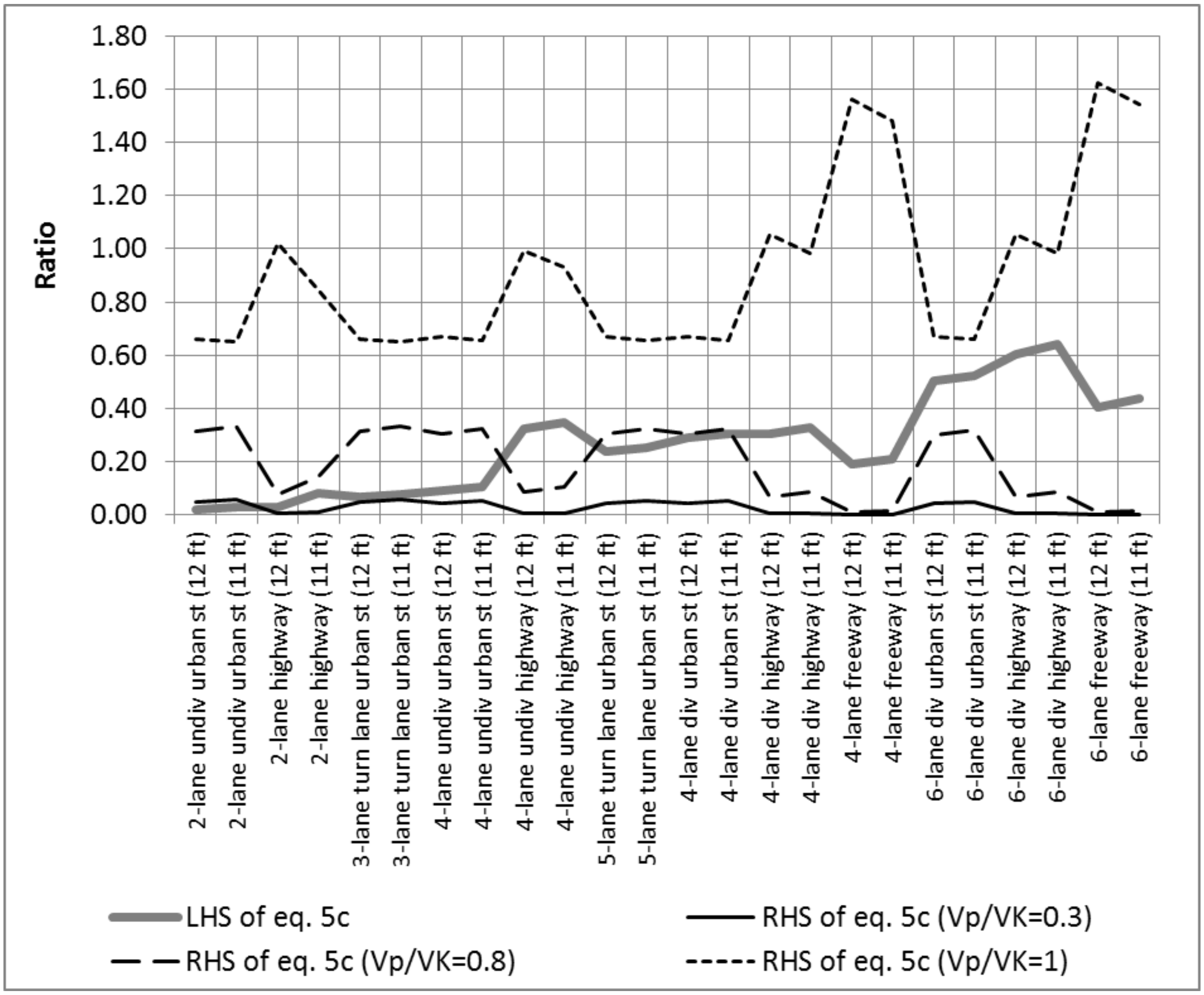
[18] According to Hu and Reuscher (2004), 59 percent of all national person trips occur during the twelve off-peak hours defined by 9 a.m. – 1 p.m. and 4–10 p.m. If it is evenly divided in direction, this amounts to about 5 percent of trips per hour on a one-directional roadway. Another 37 percent, or 6 percent per hour, occur within the six peak hours 6–9 a.m. and 1–4 p.m. This would imply a national average peaking ratio of V_p/V_o =6/5=1.2 if the peak trips

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 $V_p/V_K = 0.3$, investing in S_f is more beneficial for all types of roads except two-lane urban streets. 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 intermediate case where $V_p/V_K = 0.8$ is illuminating. With this level of peak traffic, all the highways and expressways of four lanes or more offer inefficiently high free-flow speeds relative to their capacity; whereas two-lane highways and two- to five-lane urban streets would benefit relatively more from expanding free-flow speed. 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 the most highways and expressways exhibit over-investment in free-flow speed under the design standards embedded in the Florida cost data.

are distributed evenly across directions, or 9/5=1.8 if half of the peak trips are concentrated in one direction (inbound in the morning, outbound in the afternoon).

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



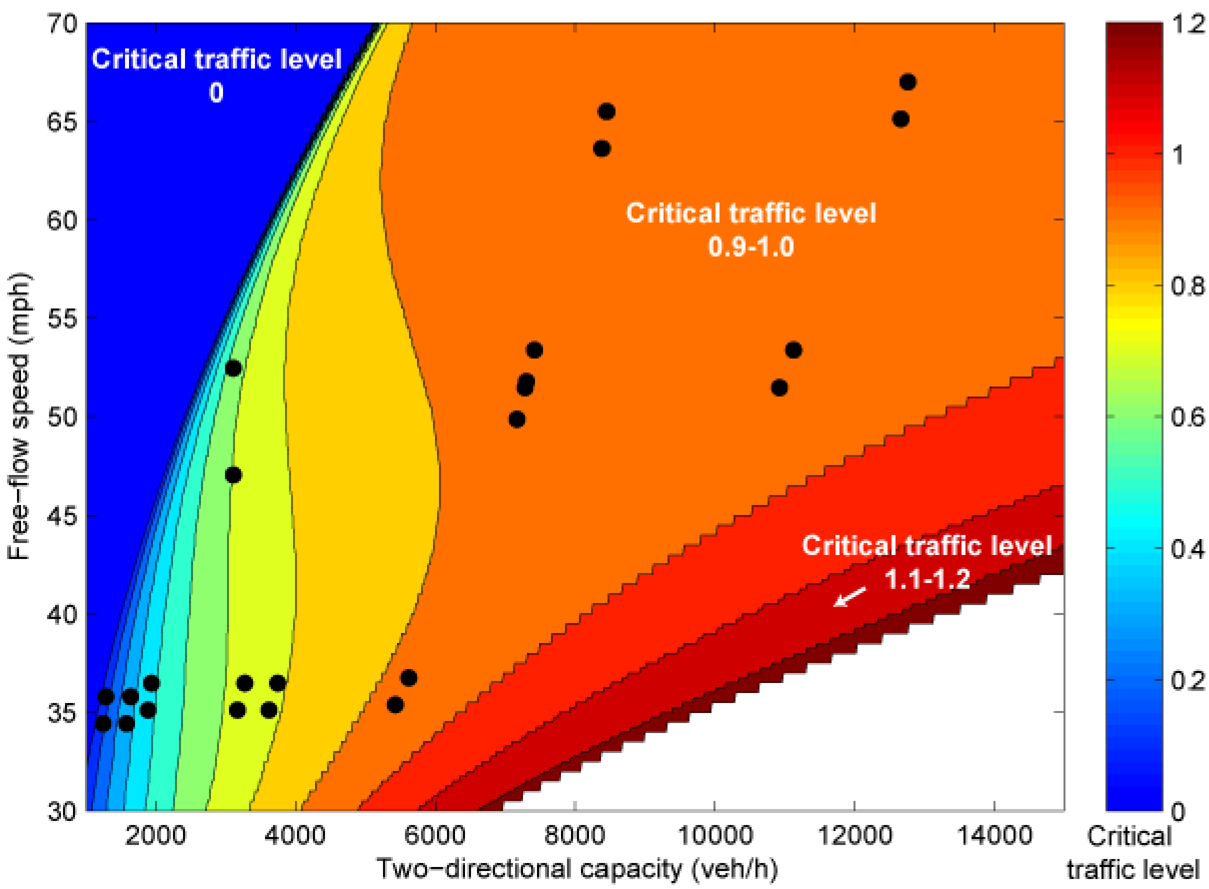
Note: Investment in S_f is favored relative to that in V_K when the LHS (ratio of construction cost elasticities: thick line) exceeds the RHS (ratio of marginal user costs: thin and dashed lines).

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.[19] As we shall see, however, the analysis of a large discrete change can be more sensitive to this assumed ratio.

[19] 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.

Figure 6 broadens the computations to a wide range of free-flow speeds and capacities. For each combination of these two investment variables, it displays the "critical traffic level," defined as the maximum value of V_p/V_K for which the ratio of construction 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.

Figure 6. Critical traffic levels for various free-flow speeds and capacities, 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 0.5 mi/h increments of free-flow speed and 20 veh/h increments of two-directional capacity.

In the upper left portion of the figure, with high free-flow speed but low capacity, the critical traffic level is zero: investment in capacity instead of free-flow speed is strongly preferred. As free-flow speeds and capacities rise, in general the critical traffic level increases; for many types of roads, it is between 0.9 and 1.0 (just before queuing begins), which is intuitive because queuing causes the marginal external congestion cost to rise significantly, making the case for capacity investment much more compelling. In the unshaded lower right portion of the figure, the critical traffic level is not calculated but is probably greater than 1.25;[20] these are high-capacity roads with low free-flow speed that would strongly benefit from incremental investment in free-flow speed.

For the road types in our sample, shown as black dots in the figure, the critical traffic levels range from 0.1 to 0.5 for urban streets of less than five lanes, and from 0.6 to almost 1.0 for all other road types. Corresponding average peak speeds for these critical traffic levels, shown in Appendix C, range from 28 to 36 mi/h for urban streets and two-lane highways, and from 47 to 56 mi/h for multilane highways and 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. We use the average free-flow speed and average peak speed for “freeways” and “arterials”, as compiled by the Schrank et al. (2012b), for “very large” and “large” urban areas.[21]

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

[20] The critical values are not calculated explicitly here because this region violates our model’s assumption that $V_o/V_K < 1$ (i.e., off-peak volumes do not encounter queuing).

[21] These areas 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.

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 ratio. Thus, for each urban area we have a representative "average" road (either a freeway or arterial) with unique free-flow speed, capacity, and peak/off-peak volume-capacity ratio; 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, the investment balance for a representative road does not necessarily apply to the entire urban area.

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	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	0.993	0.976
Ratio of construction cost elasticities	0.95	0.43	0.58	0.54	0.48	0.77	0.45
Ratio of marginal user costs	2.55	1.67	1.12	1.42	1.66	0.99	0.46
Imbalance (+ favors investment in S_f)	-1.60	-1.23	-0.54	-0.88	-1.18	-0.21	-0.01
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	0.13	0.20	0.30	0.06	0.20	0.21	0.15
Ratio of marginal user costs	0.20	0.17	0.22	0.17	0.16	0.13	0.19
Imbalance (+ favors investment in S_f)	-0.08	0.04	0.08	-0.12	0.04	0.08	-0.04

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.[22]

For arterials, the imbalance is generally quite close to zero. The biggest imbalance is in Boston, for which an unusually small average lane width and high congestion imply a relative preference for capacity. In Miami and St. Louis, there is a slightly greater incremental benefit from improving arterial free-flow speeds than for 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.

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). We summarize by calculating the benefit-cost ratio as the travel time savings from an incremental increase in 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:

[22] 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, in contrast to the case of P=4 where nearly all of the imbalances were negative; whereas the arterial imbalance is still fairly similar (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.

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

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

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

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.[23]

One can alternately view this calculation as the maximum cost multiplier that could justify the investment under consideration, where 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. Even so, this calculation should not be taken too literally, because it does not account for induced traffic: the tendency of greater capacity to attract new users. As a result, it will exaggerate the benefit-cost ratio that could be achieved in reality, as demonstrated by SACTRA (1994). In addition, we reiterate that we have less confidence in the absolute than in the relative calculations.

Table 6 shows the results for the sample of cities already discussed in Section 5.2. Using these figures, the case for investment is strong in both dimensions, in all areas. The variations across cities are not surprising. The case for investment in freeway capacity is extremely strong in Los Angeles, with its low average peak freeway speed, and much less so in relatively uncongested Jacksonville. For arterials, the case for capacity investment is strongest in Boston and weakest in St. Louis. The case for investment in greater free-flow speed is strongest for St.

[23] 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 populations of 1 million or more). We use the same value of time per vehicle as Schrank et al. (2012b), namely \$16.79/hr, who base their figure on McFarland and Chui's (1987) estimate, updated to 2011 dollars, and on assumed average vehicle occupancy of 1.25.

Louis freeways and Miami arterials, while weakest for Jacksonville freeways and Boston arterials.

Table 6. Absolute benefit-cost ratios from incremental investments, assuming Florida capital costs and no induced traffic

	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)	2,789	2,278	2,426	2,356	2,224	2,147	2,256
B/C: incr. invest. in V_K	49.2	37.0	23.4	30.8	37.8	25.0	9.4
B/C: incr. invest. in S_f	18.3	9.6	12.0	11.7	10.9	19.6	9.2
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)	879	732	810	522	682	563	877
B/C: incr. invest. in V_K	8.6	5.9	8.4	11.4	5.6	3.9	7.1
B/C: incr. invest. in S_f	5.4	7.2	11.4	3.8	7.0	6.2	5.7

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.

Finally, we present an example of a situation where one can trade off an increase in capacity for a decrease in free-flow speed by choosing among two road types. Here we depart from our incremental analysis using continuous functions, and instead perform straightforward cost-benefit calculations. Each calculation considers replacing plans for a standard six-lane freeway by instead building two undivided four-lane highways with below-standard lane widths. The two highways combined are slightly more expensive to build and provide 12 percent more capacity, but at a cost of 26 percent lower free-flow speed. For this example, we assume the freeway would encounter peak travel time of just under 30 minutes for a 10-mile trip, which is associated with a peak volume-capacity ratio of 1.15. Having the same number of vehicles distributed evenly across the two highways would give each of these roads a peak volume-capacity ratio of 1.02.

Results are shown in Table 7. Using the same peaking assumption as before, that the ratio of peak to off-peak volume is 1.25, building the two highways instead of the freeway saves more than 10 minutes per peak trip, but adds nearly 4 minutes per off-peak trip. Thus, the six-lane freeway is preferred since both its capital cost and total user time cost are lower. However, if we assume instead that V_p/V_o=1.5, i.e., we have the same peak volume as before but there are now fewer vehicles during off-peak hours, then off-peak travel time on the highways increases by just 3.5 minutes relative to that on the freeway and total user time actually decreases. As it happens, the value of this time savings is worth more than the extra capital cost, yielding a benefit-cost ratio of 2.64.

Table 7. Example of tradeoff between free-flow speed and capacity

	V_p/V_o = 1.25		V_p/V_o = 1.50	
	6-lane freeway (12 ft)	Two 4-lane undiv hwys (11 ft)	6-lane freeway (12 ft)	Two 4-lane undiv hwys (11 ft)
Free-flow speed, S_f (mi/h)	67.0	49.9	67.0	49.9
Capacity, V_K (veh/h)	12,763	14,339	12,763	14,339
V_p/V_K	1.15	1.02	1.15	1.02
V_o/V_K	0.92	0.82	0.76	0.68
Average peak travel time, T_p (min)	29.6	19.0	29.6	19.0
Average off-peak travel time, T_o (min)	9.1	12.9	9.0	12.5
Capital cost, ρ (million $ per mi)	2.41	2.78	2.41	2.78
Total user time cost (million $ per mi)	28.66	29.10	26.29	25.30
Total cost (million $ per mi)	31.07	31.88	28.69	28.08
Incremental benefits, B (million $ per mi)		-0.43		0.99
Incremental capital cost, C (million $ per mi)		0.37		0.37
B/C		-1.15		2.64

Note: All benefits and costs are per year, and the incremental benefits/capital cost are calculated based on building two four-lane undivided highways instead of one six-lane freeway.

Intuitively, because the two highways offer more total capacity at the expense of free-flow speed, they are beneficial to peak travelers at the expense of off-peak travelers. In general, we would expect that this type of tradeoff would be more favorable to the higher-capacity option when V_p/V_o is high.

This example is motivated in part by Samuel (2006), who argues that most US cities have major roads that are too wide and too sparsely spaced. Samuel argues the point from a different perspective, involving the engineering inefficiencies of intersections between very wide roads. Our approach, which recognizes explicitly the tradeoff between the needs of peak and off-peak travelers, thus complements his. While our earlier analysis of investment balance does not strictly apply to this discrete example, it does give some clues. In this example, the "investment balance" for the freeway (not shown in the table) is -4.5 at the higher ratio of peak- to off-peak traffic; that is, at the margin, the freeway offers too high a free-flow speed relative to capacity. The highway, by contrast, is much closer to balance, with value -1.0. Thus, it is perhaps not surprising that the freeway investment turns out unfavorable in this case.[24]

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 analytical functions 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

[24] However, the investment balance, an incremental criterion, is not nearly as sensitive to V_p/V_o as is the benefit-cost criterion for this discrete investment example: at V_p/V_o=1.25, the balance is -3.7 for freeways and -0.8 for arterials.

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.

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, 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.

Fifth, our analysis does not include induced demand, i.e., the tendency of a road improvement to attract new traffic. This might well affect investment balance as well as the absolute investment criteria. To analyze this, one would need to have a more microscopic picture of induced demand than is common, relating it specifically to increases in average speed by time period.

Finally, the potential for road pricing to reduce congestion would substantially change the optimal balance analyzed here, probably in favor of less capacity and more free-flow speed. Thus, our model suggests another potentially important long-run implication of road pricing: changing the nature as well as the capacity of a desirable urban road network.

With these and other improvements, we believe our approach to modeling road investment offers the potential for expanding insights and increasingly sophisticated practical analysis, all 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.

Federal Highway Administration, 2011. Summary of Travel Trends: 2009 National Household Travel Survey. http://nhts.ornl.gov/2009/pub/stt.pdf

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

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

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

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 Harwitz, Mitchell, 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.

SACTRA (Standing Committee on Trunk Road Assessment), 1994. Trunk Roads and the Generation of Traffic, London: Her Majesty's Stationery Office.

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

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, A., and Dowling, R., 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 Verhoef, Eric T., 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.

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_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
$K(\cdot)$	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 (to be made available to readers, not intended for publication)

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:

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

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

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

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

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

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:

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

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:

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

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.[25] 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, 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

[25] 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.

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.[26] 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.

[26] 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.

Appendix B. 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 88 percent of the statewide estimates) while District 7's cost estimates are higher (about 108 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[27]:

> "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."

Since District 7's cost estimates for urban roads are 108 percent of the statewide estimates, on average, the SPUI cost estimate is divided by 108 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

[27] See http://www.modot.org/stlouis/links/SinglePointUrbanInterchanges.htm.

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 C.1 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	0.02	0.05	0.32	0.66	0.15	35.4
		11	34.4	0.03	0.06	0.33	0.65	0.18	34.0
	Two-lane highway	12	52.5	0.03	0.01*	0.08	1.02	0.64	35.6
		11	47.1	0.08	0.01*	0.14	0.85	0.70	28.6
2 lanes, ctr turn lane	Urban street	12	35.8	0.07	0.05*	0.32	0.66	0.37	34.7
		11	34.4	0.08	0.06*	0.33	0.65	0.36	33.4
4 lanes, undivided	Urban street	12	36.5	0.09	0.04*	0.31	0.67	0.47	35.4
		11	35.1	0.11	0.05*	0.32	0.66	0.47	34.1
	Multilane highway	12	51.8	0.32	0.01*	0.08*	1.00	0.92	49.3
		11	49.9	0.35	0.01*	0.11*	0.93	0.92	47.5
4 lanes, ctr turn lane	Urban street	12	36.5	0.24	0.04*	0.31	0.67	0.73	33.5
		11	35.1	0.25	0.05*	0.32	0.66	0.72	32.4
4 lanes, divided	Urban street	12	36.5	0.29	0.04*	0.31	0.67	0.78	33.1
		11	35.1	0.31	0.05*	0.32	0.66	0.78	31.9
	Multilane highway	12	53.4	0.30	0.00*	0.07*	1.05	0.93	50.7
		11	51.5	0.33	0.01*	0.09*	0.98	0.92	49.0
	Freeway	12	65.5	0.19	0.00*	0.01*	1.56	0.96	55.6
		11	63.6	0.21	0.00*	0.02*	1.48	0.95	55.7
6 lanes, divided	Urban street	12	36.8	0.50	0.04*	0.30*	0.67	0.92	31.3
		11	35.4	0.52	0.05*	0.32*	0.66	0.93	30.0
	Multilane highway	12	53.4	0.60	0.00*	0.07*	1.05	0.97	50.1
		11	51.5	0.64	0.01*	0.09*	0.98	0.97	48.3
	Freeway	12	67.0	0.41	0.00*	0.01*	1.63	0.97	55.2
		11	65.1	0.44	0.00*	0.01*	1.54	0.97	54.7

* 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).

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. (2012) 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 vehicles 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.

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.