A Stochastic Approach to Traffic Congestion Costs

T. van Woensel*† F. R. B. Cruz‡ §
t.v.Woensel@tm.tue.nl frc@cs.nott.ac.uk

January 7, 2008

Abstract — **The real world is a complex dynamic and stochastic environment. This is especially true for the traffic moving daily on our roads. As such, accurate modeling that correctly considers the real-world dynamics and the inherent stochasticity is very important, especially if government will base its road tax decisions on the outcomes of these models. The contemporary traffic prices, if any, however do not reflect the external congestion costs. In order to induce road users to make the correct decision, marginal external costs should be internalized. To assess these costs, the public sector managers need accurate operational models. We show in this article that using a better representation and characterization of the road traffic, via stochastic queueing models, leads to a more adequate reflection of the congestion costs involved. Using extensive numerical experiments, we show the superiority of the stochastic traffic flow models.**

Keywords — Queueing models; congestion costs; traffic.

1 Introduction

OPTIMAL use of transportation facilities cannot be achieved unless each additional user pays for the additional costs that he imposes on all other users and on the facility itself. As such, quantifying the external congestion costs not only contributes to a socially desirable result, but is necessary to reach such a result.

This basic idea of paying for externalities dates back to Nobel prize winner William Vickrey: in the beginning of the 1950's, he proposed that fares of the New York City subway system should be higher in peak times and in high-traffic sections, and be lower in other sections (Vickrey [1]). Building on this theorem, he later proposed the same idea (*i.e.* fares dependent on time and space) for road congestion (see Vickrey [2] and Vickrey [3]) and he discussed how technological advances could help in the realization of efficient congestion costs and the related prizes (Vickrey [4]). After these seminal works from Vickrey, a large body of literature emerged on road pricing (see e.g. Dewees [5], De Borger et al. [6], Li [7], and many others), but also in other application areas the theorem is observed, such as airport congestion (see Carlin and Park [8]and Davidson [9]), electricity pricing (see Vickrey [10]), *etc.*

The current paper adds to the literature on road congestion costs. We make use of queueing theory applied to traffic flows in order to obtain analytical estimates of the marginal congestion costs, taking into account the inherent stochasticity.

1.1 Scope and Contributions

In this article, the focus will be on the specific area of transportation science, related to congestion costs, certainly a complex, dynamic, and stochastic environment. We deal with estimating marginal congestion costs for uninterrupted traffic flows, such as the delays caused by congestion on major highways. The purpose of this article is to suggest and to illustrate an alternative approach that is based on traffic applications of queueing models. Congestion pricing, as observed by Teodorović and Edara [11], is a way to force drivers to use more rationally the transportation facilities, traveling more during off-peak hours and less, during peak hours. These authors however focus on the online real-time congestion pricing case, while we look into the off-line situation, more appropriate for policy planning.

Queueing models have mainly been used to study congestion for interrupted traffic flows at signalized and unsignalized intersections (examples include Heidemann [12] and Heidemann and Wegmann [13]). However, it has been shown (see Heidemann [14] and Vandaele et al. [15]) that they can also be usefully applied to describe and analyze congestion for uninterrupted traffic flows. Despite of the absence of the development of a formal queue, the above-mentioned papers show that queueing models applied to traffic situations provide an adequate description of the complex dynamic and stochastic environment under study. Additionally, queueing models have been shown to pro-

*Corresponding author.
†Department of Technology Management, Operations, Planning, Accounting and Control, Eindhoven University of Technology, Eindhoven, The Netherlands
‡School of Computer Science and Information Technology, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham NG8 1BB, UK
§On sabbatical leave from Department of Statistics, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil

vide very accurate results (see e.g. van Woensel and Vandaele [16]).

Compared to other approaches in the literature, the main contributions of this paper are:

1. The data requirements for the proposed approach are fairly limited. In principle, traffic count data are sufficient for empirical implementation (no speed observations are needed). This is extremely convenient as empirical approaches need to collect extra data each time the conditions on the road change.

2. Due to the analytical nature of the models used, the proposed approach lends itself very well to detailed sensitivity analysis with respect to exogenous parameters that describe traffic conditions (e.g. free flow speeds, weather conditions, capacity adjustments etc.). Along the same lines, the effects of a number of demand management techniques can be easily simulated, unlike with other empirical methods. As such, the suggested method can be easily integrated in optimal cost models and cost-benefit analysis.

1.2 *Motivation*

In the public sector accurate modeling of operations and logistics functions is a necessary precondition to effective operational planning and control. Public sector managers play a key role in determining and regulating societal externalities. Policy conclusions and regulatory policies based on inaccurate modeling affect the entire economy. Since policy has a fundamental impact on the costs of logistics activities of firms, the private sector will of course also be affected. The consequences of inaccuracy or incomplete modeling in the public sector can be even more significant than in the private sector. In fact, the resulting policy conclusions and regulatory policies will affect the entire economy and not just a single firm. This is perhaps nowhere more evident than in the area of transportation policy in which different models (e.g. for emissions, costs etc.) have led to significant policy pronouncements about everything from ozone precursors to vehicle routing restrictions.

1.3 *Organization*

The structure of the paper is as follows. In Section 2 the methodology underlying the queueing approach to congestion in uninterrupted traffic flows is explained. Section 3 shows how to model the marginal congestion costs within this queueing framework. In Section 4, an application of the methodology is presented, based on empirically observed data flows from Belgium. The queueing models are applied to compute the marginal congestion costs, and a sensitivity analysis of these costs is performed in relation to the various exogenous parameters describing the traffic conditions. The different policy implications are discussed. In Section 5, this article is concluded with final remarks and topics for future research in the area.

2 Travel Time Functions

An important issue in the modeling and optimization of transportation networks is the characterization of the travel time functions, leading to acceptable estimates of the travel times on the road under a wide range of conditions (e.g. rain, two lanes etc.). In this section, we briefly review the research so-far in traffic flow modeling practice. For a detailed comparison of some of the travel time functions mentioned here, the reader is also referred to van Woensel et al. [17].

2.1 *Empirical approaches*

It is often observed that the speed for a certain time period tends to be reproduced whenever the same flow is observed. Based on this observation, it seems reasonable to postulate that, if traffic conditions on a given road are stationary, there should be a relationship between flow, speed, and density. This relationship results in the concept of speed-flow-density diagrams. These diagrams describe the interdependence of traffic flow (q), density (k) and speed (v). The seminal work on speed-flow diagrams was the paper by Greenshields [18]. Among the existing methods to describe traffic flows and the travel times the most popular is by far the use of empirical speed-flow relations (for more recent applications see, among others, O'Mahony and Kirwan [19]) and Li [7]).

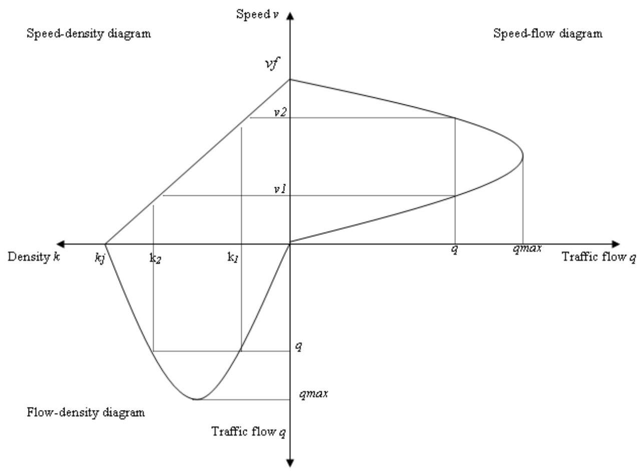

Figure 1: The relations between the speed-flow, speed-density, and flow-density diagrams

Figure 1 illustrates that although every speed v has a corresponding traffic flow q the reverse is not true. There are two speeds for every traffic flow, an upper branch (v_2), in which the speed decreases as the flow increases, and a lower branch (v_1), in which the speed increases. Intuitively it is clear that, as the flow moves from 0 (at free flow speed v_f) to $q_{\max}$, the congestion increases but the flow rises because the decline in speed is over-compensated by the higher traffic density; if traffic tends to grow past $q_{\max}$, flow falls again because the decline in speed more than offsets the additional vehicle numbers, further increasing congestion [20]. The flow-density diagram and the speed-density diagrams are an equivalent representation and can be interpreted in the same way.

Two standard procedures exist for determining the needed parameters of these relations (see Daganzo [20]). First, they can be estimated by econometric models by using data on average traffic volumes and observed average speeds, selecting the functional form that best fits the data (see e.g. Mayeres [21]). Secondly, assuming particular nonlinear functional forms with a limited number of unknown parameters, the latter can be calibrated on the basis of just a few observations, typically including the free flow speed at zero traffic flow (see e.g. De Borger and Proost [22]). Obviously, both methods have shortcomings:

1. Estimating speed-flow relations by econometrics requires data on average speeds-average flows (or volume versus time per kilometer). Typically traffic volumes are available from standard traffic counts, but in many cases information on the corresponding average speed of the traffic flow is not recorded. To cope with this problem, other data generation processes have to be used (see e.g. O'Mahony and Kirwan [19]).

2. Calibration of speed-flow relations on the basis of just a few observations leads to obvious prediction problems for speeds that substantially deviate from those used to calibrate the parameters.

3. Both methods are of limited use to study the effect on congestion of changes in a number of important determinants such as road capacity, free flow speeds, weather conditions, demand management techniques etc.

2.2 *Analytical approaches*

The analytical approaches are preferred over simulation techniques, which usually are computationally intensive, or econometric models, which are strongly dependent on the underlying data. Of course, such a choice may lead to very complicated and highly intractable models. There is of course always the question of finding an acceptable trade off between approximate analytical and simulation. The resulting analytical queueing models are transparent and usable enough to motivate the use of these models (see Vandaele et al. [15]). Instead of presenting a formal overview of the queueing methodology and its application to traffic flows, the remainder of this section will offer an intuitive approach and a selection of the crucial formulas based on the work of Vandaele et al. [15], van Woensel et al. [23], and Heidemann [14], which is sufficient to understand the remaining of this article.

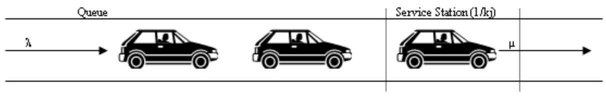

Figure 2: Queueing representation of traffic flows

In a queueing approach to traffic flow analysis, roads are subdivided into segments, with length equal to the minimal space needed by one vehicle on that road, as seen in Figure 2. Let us define k_j as the maximum traffic density (i.e. maximum number of cars on a road segment). This segment length is then equal to $1/k_j$ and matches the minimal space needed by one vehicle on that road. Each road segment is considered as a service station, in which the vehicles arrive at a certain rate λ (i.e. the demand) and they are served at another rate μ (see Vandaele et al. [15]).

Following Heidemann [14], the arrival rate λ is defined as the product of the traffic density k and the free flow speed v_f, or $\lambda = k \times v_f$. Similarly, the service rate μ is defined as the product of free flow speed v_f with the maximum traffic density k_j, or $\mu = k_j \times v_f$ (Heidemann [14]). The interaction of λ and μ results in a certain realized flow q which will be a function of the road characteristics (e.g. k_j and v_f) and the queueing model parameters (e.g. the variability). Vandaele et al. [15] and Heidemann [14] showed that the speed v can be calculated by dividing the length of the road segment $(1/k_j)$ by the total time in the system (W), or:

$$v = \frac{1/k_j}{W}. \qquad (1)$$

The total time in the system W is different depending upon the queueing model used. This time is the sum of the time spend in the queue W_q and the time being in service W_p. The average time in service is always equal to the mean of the distribution of the service times. The determination of the time spend in the queue W_q is however dependent upon the specific queueing model under investigation. Vandaele et al. [15] developed different queueing models[1]. This article will be limited to the $M/M/1$ and $M/G/1$ queueing models, as the results for these models can be obtained in closed form (closed form expressions for the time in the queue W_q for these basic queueing models are readily available in the literature, see e.g. Hillier and Lieberman [24]). The methodology developed for these models can however be extended to more general models (e.g. $GI/G/1$ models for which W_q needs to be approximated), but the results are only obtained numerically. Table 1 gives

[1] In order to refer to the queueing models, we use Kendall's notation, in which $M/G/1$ means a queueing system with **M**arkovian arrival rates, **G**eneral service times, and 1 server in the system.

Table 1: The time in the queue W_q

Model	W_q
$M/M/1$	$v = \frac{k_j \times v_f \pm \sqrt{k_j \times v_f \times [k_j \times v_f - 4q]}}{2k_j}$
$M/G/1$	$v = \frac{2 \times k_j \times v_f - q\left(c_s^2 - 1\right) \pm \sqrt{[q(c_s^2-1) - 2 \times k_j \times v_f]^2 - 16 \times k_j \times v_f \times q}}{4k_j}$

the relevant formulas for the speed v for the models under consideration.

The validity of the queueing approach was proved to be a representation of the traffic flow reality in a number of papers, both using empirical data and simulation (see e.g. van Woensel and Vandaele [16], van Woensel et al. [25]). Validation results showed that the developed queueing models can be adequately used to model uninterrupted traffic flows. As such, these models can be used to evaluate potential improvements in the existing traffic conditions. Starting from the existing traffic state, potential improvements are easily quantified and compared with one and another. Potential improvements that can be evaluated using this queueing approach are for example, congestion costs (the scope of this paper), environmental impact of traffic (see van Woensel et al. [23]), optimal number of lanes, investment analysis, routing decisions (see van Woensel et al. [26]) etc.

2.3 *Discussion*

The traditional empirical approach is limited in terms of predictive power and sensitivity analysis. Vandaele et al. [15] and Heidemann [14], showed that queueing models can also be used to explain uninterrupted traffic flows and thus offering a more practical approach, useful for sensitivity analysis, forecasting etc. It is important to mention that the queueing models developed all assume steady-state conditions, i.e. such that the same behavior is reproduced and observed every time with the same probability. The steady-state queueing models are most appropriate in design and policy recommendations, which is the main focus of this paper. This assumption has a number of consequences:

1. The traffic flows observed are stationary meaning that all vehicles will always be driving (no matter how slow) and never come to a full stop. The non-stationary traffic experiences stop's and go's and would be more suitable for modeling bottlenecks. These non-stationary traffic flows can be modeled using e.g. transient queueing models. Heidemann [27] showed that under non-stationary conditions, the speed-flow-density results deviate from the ones obtained with stationary queueing models. [27] also demonstrated that the non-stationary flow-density diagrams converge to the stationary ones when the time period considered in the non-stationary models grows to infinity. Moreover, the transient queueing models are more useful in specific control situations for relatively small networks.

2. At some places in this paper, the term *congestion* will be used in a strictly queueing theory sense, meaning more than one customer in the system leading to traffic intensity strictly larger than 0. When considering getting stuck in traffic, the term *traffic jam* will be used. Note also that although the queueing theory terminology is used when talking about *waiting time*, a more appropriate term would be *delay*. The first term would imply in traffic flow terminology a full stop, while the latter one assumes that vehicles still move but at a lower rate.

3. The observed traffic flows can never be larger than the capacity, as in this case, the traffic intensity would grow larger than 1 which would mean in terms of queueing theory that the system is explosive (more arrivals than the service stations can handle). In queueing practice, this means that the arrivals are lost (and not *back-ordered*). Hence, this is different from the bottleneck model, If the incoming flow exceeds capacity, the excess flow accumulates in the form of a queue propagating upstream (see Daganzo [20]).

Of course, any model (e.g. econometric models, simulation etc.) is an abstraction of reality. A queueing model is a mathematical model and by definition, an abstraction of reality. It has been shown however that this abstraction does not mean that the queueing approach is inadequate to model uninterrupted traffic flows.

Finally, it has to be mentioned that other queueing approaches do exist. Jain and Smith [28] described in their paper a state-dependent $M/G/C/C$ queueing model for traffic flows. Part of their logic is used to extend our queueing models to state-dependent ones. Also a lot of research is done on a travel time-flow model originating from Davidson [9]. The model is based on some concepts of queueing theory but a direct derivation has not been clearly demonstrated (see Akçelik [29, 30]).

3 Toward Traffic Congestion Costs

Following the conventional approach in the literature (see De Borger et al. [6] and Li [7]), the total cost of a delay, denoted by TC_q (EURO per kilometer per hour), for a given flow q, is defined as

$$\mathrm{TC}_q = q \times C_q, \quad (2)$$

with $C_q = \mathrm{VOT} \times \frac{1}{v_q}$, in which C_q is the cost for a traveler to take a trip (given a certain flow q) and VOT is the value of time. The speed function v_q is then given by the queueing formulas presented in the previous sections. Note again that in the literature this C_q is coming from regression fits of empirical data, while here in this article the analytical queueing approach is followed.

The marginal congestion cost is defined as the extra cost due to a structural exogenous increase of traffic demand by 1 vehicle. In order to obtain the marginal congestion cost, the marginal change of TC_q, Eq. (2), over q is derived, which leads to

$$\mathrm{MC}_q = \frac{\partial \mathrm{TC}_q}{\partial q}, \quad (3)$$

$$= \left[C_q + q \times \frac{\partial C_q}{\partial q} \right]. \quad (4)$$

The first component, C_q or $\left[\mathrm{VOT} \times \frac{1}{v_q}\right]$ in Eq. (4), is the average cost per vehicle per kilometer experienced by the marginal traveler and the second component, $q \times \frac{\partial C_q}{\partial q}$ in Eq. (4), is the increase in time cost per kilometer due to the increase in the flow q. The first component is thus the internal cost MC_q^I. The second one is the external cost MC_q^E imposed by the driver. The specific value of both components depends upon the queueing model chosen. Table 2 summarizes the functions for different queueing models.

Table 2: External cost component for different queueing models

Model	Marginal congestion cost MC_q
$M/M/1$	$q \times \frac{4k_j^2 v_f}{\sqrt{k_j v_f (k_j v_f - 4q)} \left(k_j v_f \pm \sqrt{k_j v_f (k_j v_f - 4q)} \right)^2}$
$M/G/1$	$q \times \frac{4k_j \left((c_s^2 - 1) + \frac{(c_s^2-1)q - 2(c_s^2+3)k_j v_f}{2\sqrt{((c_s^2-1)q - 2k_j v_f)^2 - 16 k_j v_f q}} \right)}{\left((c_s^2-1)q + 2k_j v_f \pm \sqrt{((c_s^2-1)q - 2k_j v_f)^2 - 16k_j v_f q} \right)^2}$

4 Results and Applications

In this section, results are presented based on the integrated queueing-congestion costs approach. Based on the described queueing models and the proposed methodology, the marginal external congestion costs are calculated. We first elaborate on the congestion costs for traffic in a single node. Secondly, this analysis is extended toward network models.

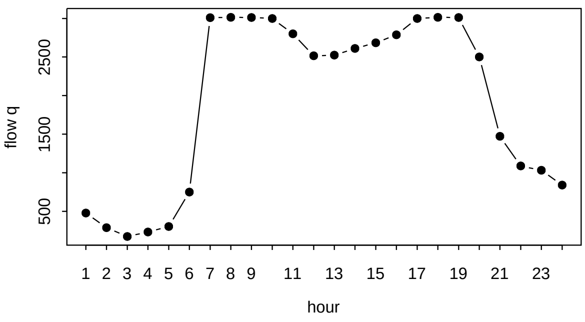

Figure 3: The hourly flow over the day

The data are collected by the Department of Environment and Infrastructure of the Ministry of the Flemish Community. This department is responsible for collecting and reporting the counting data for all Flemish roads. The dataset consists of observations of the number of vehicles per hour for 273 days of the A1/E19 highway, section Mechelen North–Rumst, headed to Brussels (i.e. counter 19012). There are three lanes which are aggregated into one lane (see also Button and Hensher [31]). All vehicle observations are recalculated into passenger car equivalents (one truck counting for two passenger cars). The data used can be seen in Figure 3.

4.1 *Traffic at a single node*

Using the flow data obtained for one counting point on the mentioned highway, we are able to evaluate the magnitude of the congestion costs for a number of parameter settings. We set the free flow speed v_f equal to 120 km/h, which is approximated by the maximum speed allowed on this stretch. The jam density k_j is set equal to 201, which is obtained based on the maximum historical flow ever observed on this stretch of road (see van Woensel and Vandaele [16] for more information). The VOT is conventionally set equal to 6.67EUR (see Png et al. [32]).

Figure 4 shows the speed-flow diagrams (upper graphs) for different settings of the coefficient of variation c_s. Note that when $c_s = 1.0$, it refers to a Markovian $M/M/1$ queueing model. The other values result in a $M/G/1$ queueing model. The lower graphs represent the shape of the C_q function, or $C_q = \mathrm{VOT} \times \frac{1}{v_q}$. These curve have a positively sloped part which corresponds to the congested branch (upper branch) of the speed-flow diagram and a negatively sloped backward-bending part corresponding to the hyper-congested branch (lower branch) of the speed-flow diagram. Note that these curves generated here are analytical counterparts of the empirical curves as described in e.g. Lindsey and Verhoef [33]. Also depending upon the variability (as a function of c_s), the maximum flow that can

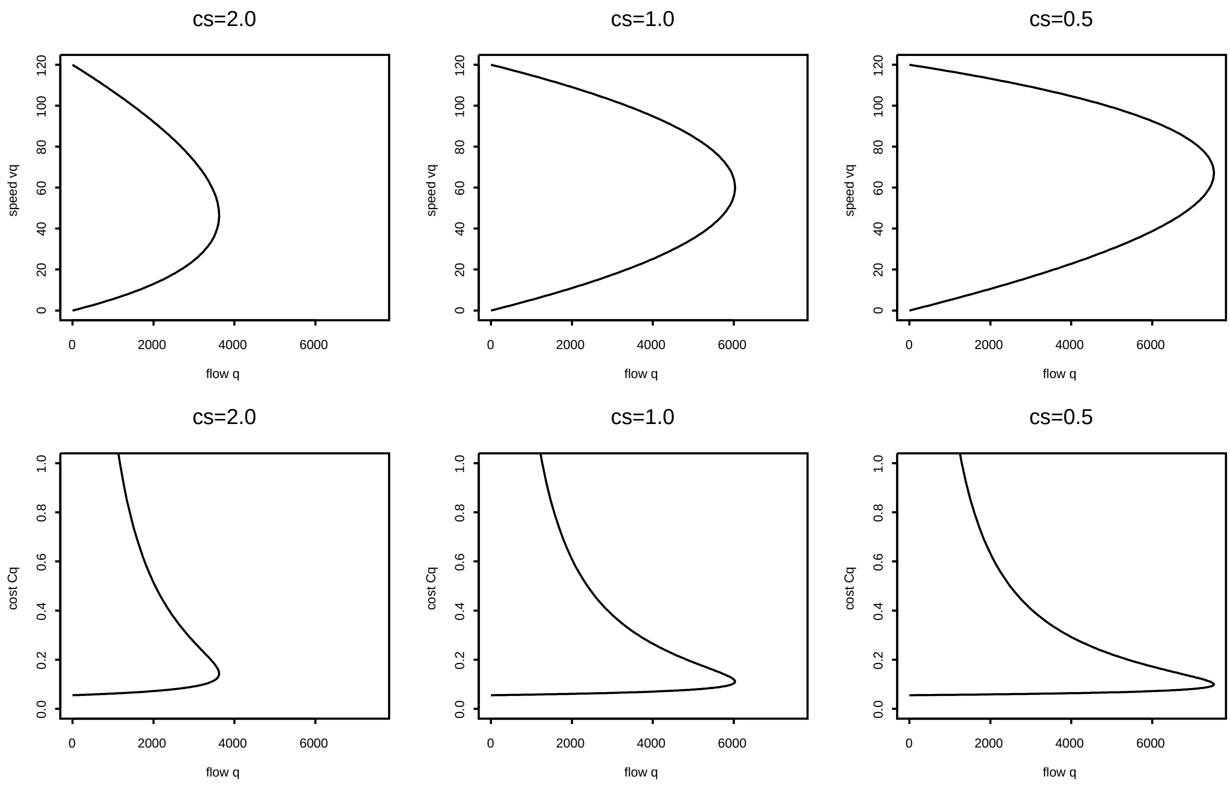

Figure 4: Speed-flow diagrams and the resulting cost for a traveler, C_q

be realized is different. It should be clear that the maximum capacity of the road system depends on the circumstances in which traffic realizes itself (e.g. snow versus sunny). This is clearly reflected in the speed-flow diagram where the maximum flow ranges between 3,600 and 8,200 vehicles.

Several authors (see e.g. Chu and Small [34] and Lindsey and Verhoef [33]) argued that the hyper-congested branch of the speed-flow density diagram is not stable. These authors concluded that this part of the speed-flow curve is the result of transient demand fluctuations and is considered to be unsuitable for building policy plans. As such, for economic cost analysis and policy applications, it is reasonable to ignore the hyper-congested branch of the speed-flow curve and only focus on the congested (steady-state) branch of the speed-flow curve (see also Button [35]). We will adopt the same approach in this paper, as the prime purpose is to develop generic insights for policy making which should be based on steady-state rather than on transient behavior.

Figure 5 shows the marginal congestion costs MC_q^I and MC_q^E for different coefficients of variation c_s, but considering only the upper (i.e. congested) branch of the speed-flow diagram. Both the internal cost component as the external cost component are an increasing function in flow q. However, the external cost MC_q^E tends to be significantly higher than the internal cost MC_q^I for medium to high flows.

Figure 6 gives the respective internal cost component and the external cost component, for a specific flow profile for a random day (depicted in Figure 3). One can also see that the congestion costs during the morning peak (from 7–10) are significantly higher than the ones in the evening peak (from 16–19). This of course is due to the observation that the traffic intensity in the morning peak is higher than in the evening peak, which tends to be spread over a longer time. In the off-peak moments however, the total costs are mainly comprised of the internal costs. The external costs are limited. It should be clear that if one decide to use the road on highly congested moments, the user pays more. This has been recognized already in the literature mentioned before. On the other hand, note again the influence of the stochasticity on the external costs. For high traffic intensities, the costs tend to be much higher for days with high variability compared to low variability. Obtaining the relevant costs for these situations is only possible when using analytical models (rather than empirical models).

4.2 *Sensitivity Analysis*

A sensitivity analysis is done for the different parameters in the model. As pointed out earlier, this is one of the major contributions of the proposed queueing-congestion costs approach compared to other

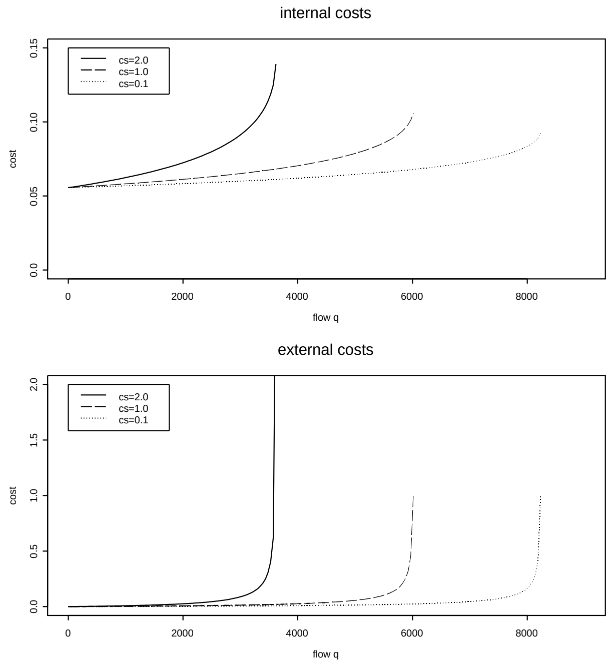

Figure 5: The marginal congestion cost MC_q for different coefficients of variation c_s

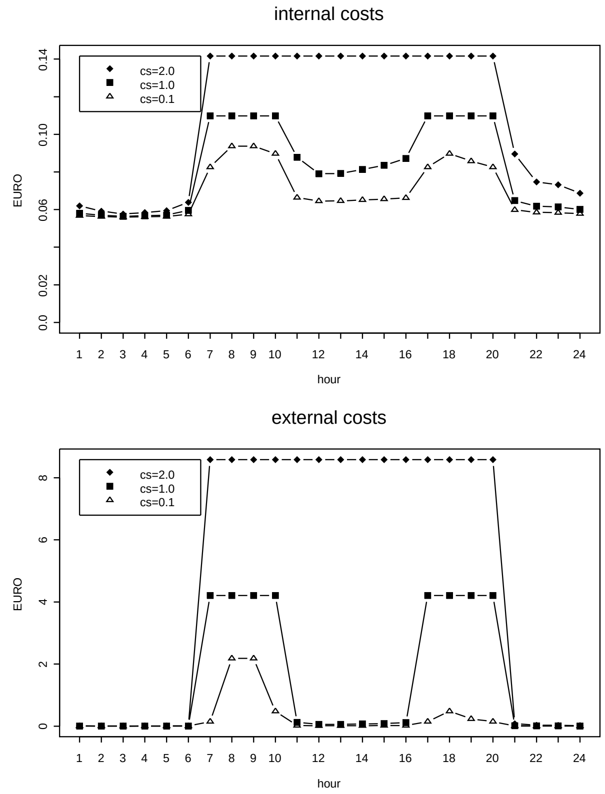

Figure 6: The marginal congestion costs MC_q over a day

approaches. In order to perform a sensitivity analysis, partial derivatives of the marginal external congestion cost to the parameters v_f, k_j and c_s are calculated, keeping all other parameters equal. The rationale behind changing the jam density is to determine the influence of adding lanes on the congestion costs. This can be done as the jam density k_j can be seen as a proxy for the number of lanes. For reasons of simplicity, it is assumed that the jam density is a linear function of the number of lanes. The coefficient of variation of the service times can be used to see the effect of the different factors that influence the traffic, e.g. weather circumstances. The analytical formulas for the different queueing models can easily be obtained.

In general, the effect (sign) for v_f and k_j is negative and for c_s is positive. In other words, increasing the free flow speed v_f or the jam density k_j will result in lower costs (due to the increase in capacity), while an increase in the coefficient of variation c_s will result in higher costs (due to the decrease in capacity).

In Figure 7, the external cost component is always shown for the maximum flow that can be observed for the parameter setting (i.e. a highly congested situation) as the results of the previous section suggested that the external cost component only becomes significant for high congestion. One should note that this is not the same as the above partial derivatives of the marginal external congestion cost to the parameters v_f, k_j and c_s. There it was assumed that the flow q remained the same; in the remainder of this section the effect is examined as if the flow on the highway is always equal the maximum flow able to use the highway (which depends upon the parameter settings). The rationale behind is that if, for example the number of lanes is increased, the number of cars eventually using the road will also increase. As such the sensitivity results presented in the remainder of the section can be seen as long term effects.

The sensitivity analysis shows that the external costs for the maximum flows are strongly depending on the parameters. It is clear that there is a significant stochastic effect on the external costs. Ignoring this important influence and treating this complex phenomenon as deterministic (as done in many empirical approaches) is bound to lead to inaccurate policy making.

4.3 Policy Implications

In the literature, the estimated marginal external congestion costs vary dramatically: for a congested highway the approximations of Mayeres [21], De Borger et al. [6] and Li [7] range from 0.2040EUR to 1.8751EUR per kilometer per vehicle. The range of the former results is remarkable, since the assumed hourly brute income and the value of time (VOT) expressed as a fraction of the hourly brute income do not differ significantly (the hourly brute income is in general assumed to be equal to 10EUR, and a recent estimate (Png et al. [32]) presents the VOT as 67% of the hourly brute income). For the queueing-congestion costs approach the cost estimates are presented in Table 3 (with the speed-flow profile and parameter setting discussed in the pre-

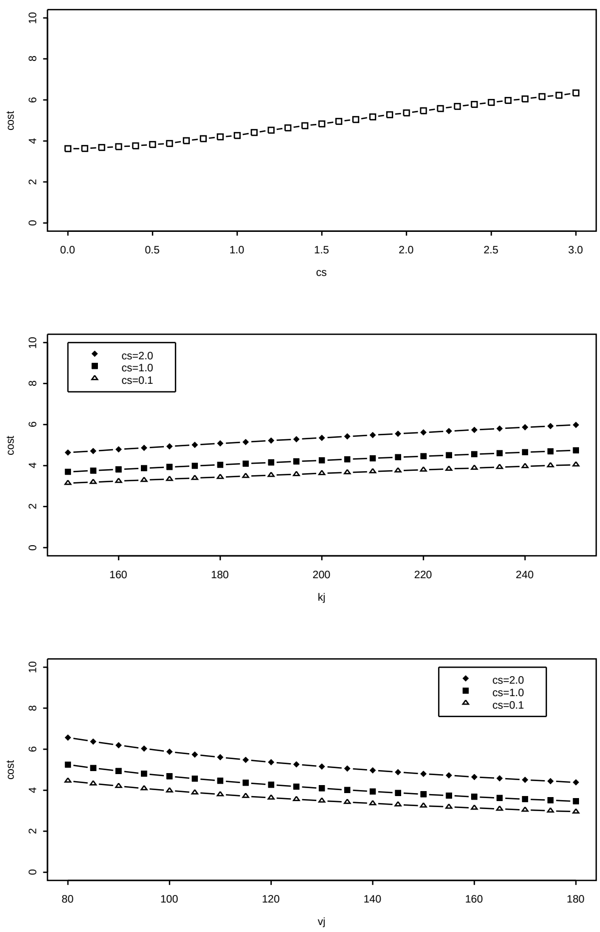

Figure 7: Sensitivity of the parameters on the external costs

vious section).

Table 3: Overview of peak and off-peak external costs

	$c_s = 0.1$	$c_s = 1$	$c_s = 2$
Morning peak	1.24EUR	2.35EUR	4.66EUR
Evening peak	0.29EUR	1.66EUR	3.38EUR
Off-peak	0.02EUR	0.03EUR	0.05EUR

As one can see, these numbers are not substantially different from the results reported in the literature. There are however significant differences when looking at the different levels of variability. Self-evidently, to explain the differences, one has to bear in mind the outdated traffic flow data used in early literature. Next to this, it should be noted that the results reported in the literature are based on observed speed-flow density diagrams with an implicit coefficient of variation of the service times (which is unknown). The presented approach can thus be seen as a framework in which all past observations regarding costs can fit in.

In the literature, most studies rely on comparing the existing traffic conditions against a notional 'base' in which the traffic volumes are at the same high levels, but all vehicles all deemed to travel at completely congestion-free speeds. This situation could never exist in reality is it reasonable to encourage public opinion to imagine that this is an achievable aim of transport policy. In this paper, the idea of a totally congestion-free target is ignored, and emphasis is put on the change in congestion that would be realistically achievable as a result of implementing specific more or less ambitious transport policies, such as road building, public transport improvements, and transport prices.

Since the policy made by public sector managers, mainly depends on the models they use, accurate operational models are mandatory. Suboptimal decisions affect the entire economy and drive companies into decision making which can rather be worse than better for the society as a whole and other stake-holders. Therefore government should use accurate and complete tools to support their operational policy conclusions, especially when it comes to regulating traffic, one of the very important characteristics of contemporary logistical activities. In addition, these regulations can significantly impact the way companies conduct their transportation planning, distribution organization, deliveries, shipments, etc.

These observations are crucial when it comes to the development of traffic measures. The decisions, made by strategic and operational public sector managers clearly affect the private sector. Thus, the dynamic queueing approach proposed has the potential for making important contributions to improving the decision making of both public sector managers and private companies related to logistics and traffic movements. Through the use of the sensitivity analysis, the public sector manager can get useful insights in the effects of his measures. Moreover, not only short term effects (where flow is unchanged) are quantified but also long term effects (where flow increases) can be quantified.

5 Conclusions and Topics for Future Research

As documented in the literature, optimal use of a transportation facility cannot be achieved unless each additional user pays for the additional costs that this user imposes on all other users and on the facility itself. As such, a congestion toll not only contributes to a socially desirable result, but is necessary to reach such a result. This paper makes use of the existing queueing theory for traffic flows in order to get an estimate of marginal congestion costs.

The main advantage of the methodology described in this paper, consists in the possibility to derive the marginal congestion costs in an analytical way taking into account the inherent stochasticity of the real world. This approach relies less on the availability of required data: only data on the flows is needed, contrary to the traditional approaches where also data for the speed is needed. The marginal congestion cost appears readily easy to calculate, and can straightforwardly be decomposed in internal and external costs. It is demonstrated that using the queueing approach to the speed-flow-density relationship leads to a more controllable environment to conduct what-if analysis and perform sensitivity analysis. Based on the experiments, the effects of a change the free flow speed (e.g. changing speed limit), the jam density (e.g. add an extra lane) and the coefficient of variation (e.g. snowy day versus sunny day) are shown.

Future research mainly involves extending the approach toward multiple user types (e.g. passenger cars versus trucks) and finding the equilibrium conditions in a (queueing) network setting for the proposed approach.

Acknowledgments

The research of prof. Frederico Cruz has been partially funded by CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) of the Ministry for Science and Technology of Brazil, grants 201046/1994-6, 301809/1996-8, 307702/2004-9, 472066/2004-8, and 472877/2006-2, by CAPES (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*), grant BEX-0522/07-4, by FAPEMIG (*Fundação de Amparo à Pesquisa do Estado de Minas Gerais*), grants CEX-289/98, CEX-855/98, and TEC-875/07, and by PRPq-UFMG, grant 4081-UFMG/RTR/FUNDO/PRPq/99.

References

[1] Vickrey WS. A proposal for revising New York's subway fare structure. Journal of the Operations Research Society of America 1955;3 (1): 38–68.

[2] Vickrey WS. Pricing in urban and suburban transport. The American Economic Review 1963;53 (2): 452–465.

[3] Vickrey WS. Congestion theory and transport investment. The American Economic Review 1969;59 (2): 251–260.

[4] Vickrey WS. Statement to the joint committee on Washington, DC, metropolitan problems (with a foreword by Richard Arnott and Marvin Kraus). Journal of Urban Economics 1994;36 (1): 42–65.

[5] Dewees DN. Estimating the time costs of highway congestion. Econometrica 1979;47 (6): 1499–1512.

[6] De Borger B, Mayeres I, Proost S, Wouters S. Optimal pricing of urban passenger transport: A simulation exercise for Belgium, Journal of Transports Economics and Policy 1996; 31–54.

[7] Li MZF. The role of the speed-flow relationship in congestion pricing implementation with an application to Singapore. Transportation Research Part B 2002;36: 731–754.

[8] Carlin A, Park RE. Marginal cost pricing of airport runway capacity. American Economic Review 1970;60 (3): 310–319.

[9] Davidson KB. The theoretical basis of a flow - Travel time relationship for use in transportation planning. Australian Road Research 1978;8 (1): 32–35.

[10] Vickrey WS. Efficient pricing of electric power service: Some innovative solutions. Resources and Energy 1992;14 (1-2): 157–174.

[11] Teodorović D, Edara P. A real-time road pricing system: The case of a two-link parallel network. Computers & Operations Research 2007;34 (1): 2–27.

[12] Heidemann D. Queue length and delay distributions at traffic signals. Transportation Research Part B 1994;28: 377–389.

[13] Heidemann D, Wegmann, H. Queueing at unsignalized intersections. Transportation Research Part B 1997;31: 239–263.

[14] Heidemann D. A queueing theory approach to speed-flow-density relationships. In: Proceedings of the 13th International Symposium on Transportation and Traffic Theory. Lyon, France 1996.

[15] Vandaele N, van Woensel T, Verbruggen A. A queueing based traffic flow model. Transportation Research Part D 2000;5 (2): 121–135.

[16] van Woensel T, Vandaele N. Empirical validation of a queueing approach to uninterrupted traffic flows. 4OR 2006;4 (1): 59–72.

[17] van Woensel T, Kerbache L, Peremans H, Vandaele N. Vehicle routing with dynamic travel times: A queueing approach. European Journal of Operational Research 2008;186 (3): 990–1007.

[18] Greenshields BD. A study of traffic capacity. Highway Research Board Proceedings 1935;14: 448–477.

[19] O'Mahony M, Kirwan KJ. Speed-flow relationship and feasibility of roadpricing. Edward Elgar Publishing Ltd, Ch. 10 - Reforming Transport Pricing in the European Union - A Modeling Approach 2001.

[20] Daganzo CF. Fundamentals of Transportation and Traffic Operations. Elsevier Science 1997.

[21] Mayeres I. The marginal external cost of car use - with an application to Belgium. Tijdschrift voor Economie en Management 1993;XXXVIII (3): 225–258.

[22] De Borger B, Proost S. Reforming Transport Pricing in the European Union: A Modeling Approach. Edward Elgar 2001.

[23] van Woensel T, Creten R, Vandaele N. Managing the environmental externalities of traffic logistics: The issue of emissions, Journal of POM, Special issue on Environmental Management and Operations 2001;10 (2): 1–29.

[24] Hillier FS, Lieberman GL. Introduction to Operations Research, McGraw-Hill, 1995.

[25] van Woensel T, Wuyts B, Vandaele N. Validating state-dependent queueing models for uninterrupted traffic flows using simulation, 4OR 2006;4 (2): 159–174.

[26] van Woensel T, Kerbache L, Peremans H, Vandaele N. A queueing framework for routing problems with time-dependent travel times, Journal of Mathematical Modelling and Algorithms, Special Issue "Quantitative Aspects of Transportation and Logistics" 2007;6 (1): 151–173.

[27] Heidemann D. Non-stationary traffic flow from a queueing theory viewpoint. In: Proceedings of the 14th International Symposium on Transportation and Traffic Theory. Jerusalem, Israel 1999.

[28] Jain R, Smith JM. Modeling vehicular traffic flow using $M/G/C/C$ state dependent queueing models. Transportation Science 1997;31 (4): 324–336.

[29] Akçelik R. Travel time function for transport planning purposes: Davidson's function, its time dependent form and an alternative travel time function. Australian Road Research 1991;21 (3): 49–59.

[30] Akçelik R. Relating flow, density, speed and travel time models for uninterrupted and interrupted traffic. Traffic Engineering and Control 1996;37 (9): 511–516.

[31] Button KJ, Hensher DA (Eds.). Handbook of Transport Systems and Traffic Control. Handbooks in Transport. Pergamon 2001.

[32] Png GH, Olszewki P, Menon APG. Estimation of the value of travel time of motorists. Journal of Institute of Engineers 1994;34: 9–13.

[33] Lindsey R, Verhoef E. Traffic congestion and congestion pricing. Discussion Paper TI 2000-101/3 2000.

[34] Chu X, Small KA. Hypercongestion. Journal of Transport Economics and Policy 2003;37: 319–352.

[35] Button KJ. Transport Economics. Edgar Elgar Publishing 1992.