A joint Initiative of Ludwig-Maximilians-Universität and Ifo Institute for Economic Research

CESifo Working Papers

# URBAN CONGESTION CHARGING: THEORY, PRACTICE AND ENVIRONMENTAL CONSEQUENCES

Georgina Santos
David Newbery

CESifo Working Paper No. 568

September 2001

Presented at the CESifo Workshop on Environmental Economics
and the Economics of Congestion, Venice, June 2001

CESifo
Center for Economic Studies & Ifo Institute for Economic Research
Poschingerstr. 5, 81679 Munich, Germany
Phone: +49 (89) 9224-1410 - Fax: +49 (89) 9224-1409
e-mail: office@CESifo.de
ISSN 1617-9595

CESifo *Working Paper No. 568*
*September 2001*

# URBAN CONGESTION CHARGING: THEORY, PRACTICE AND ENVIRONMENTAL CONSEQUENCES

## Abstract

The theory of road pricing developed for single links suggests time and location varying charges equal to the marginal congestion cost at the efficient level of traffic. The second-best network counterpart is derived, but would be infeasible to implement. Cordon tolls are feasible, and their optimal level computed for eight towns. A cost-benefit study showed that with a suitable choice of location, all schemes were socially profitable, though with wide variations across towns. The environmental benefits of cordon tolls are measured and shown to correlate with optimal congestion tolls, but to be modest in size and not to affect the optimal toll.

JEL Classification: R41, R48, Q28, H23.

Keywords: road traffic congestion, road pricing, congestion charging, cordon tolls, environmental taxes.

*Georgina Santos*
*Department of Applied Economics*
*University of Cambridge*
*Cambridge CB3 9DE*
*United Kingdom*

*David Newbery*
*Department of Applied Economics*
*University of Cambridge*
*Cambridge CB3 9DE*
*United Kingdom*
*dmgn@econ.cam.ac.uk*

# Urban congestion charging: theory, practice and environmental consequences

**Georgina Santos and David Newbery**
**Department of Applied Economics, Cambridge, CB3 9DE, UK**
**August 15 2001**

## INTRODUCTION

The traditional approach to the economics of congestion rests on the standard welfare economic argument that the market failure of congestion requires a corrective charge to internalise the externality. Once traffic flows rise above modest levels, each additional road user lowers the speed of all traffic using the road, imposing external costs that are not taken into account in the private cost-benefit calculus of trip choice. Road taxes go some way to discouraging excessive road use, but are a very blunt instrument, overcharging traffic on less congested roads, but seriously undercharging road users in congested urban areas. In such urban areas the market equilibrium level of traffic can be excessive, particularly during peak hours and the inter-peak shoulder period.

The ideal corrective charge would be equal to the difference between the marginal social cost (MSC) and average private cost (APC) on each link and through each junction at the equilibrium level of traffic corresponding to that set of charges. The obvious difficulty with this ideal is that efficient charges would vary by link, junction and time of the day, and it would be impossible for drivers to make efficient choices even if it were feasible to compute these ideal charges and somehow announce them (either in advance or in real time) to potential travellers. In any case, the technology for such precision pricing is not even available, and would almost certainly not be cost effective even if it were.

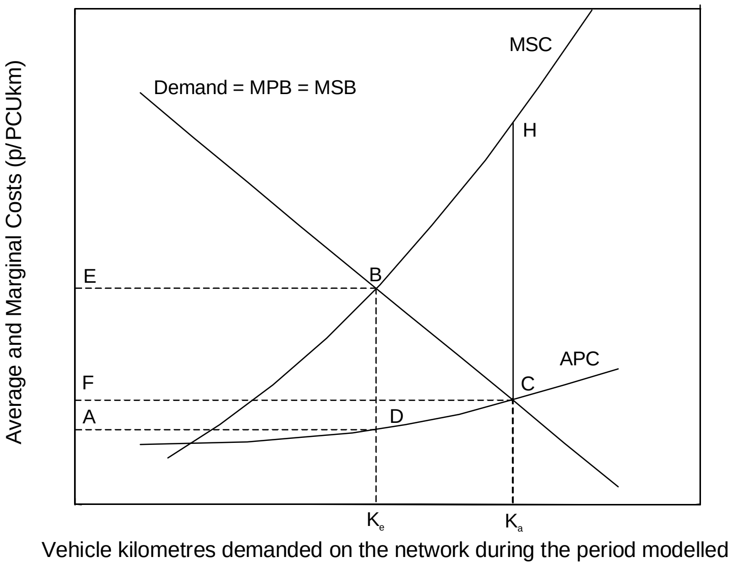
Nevertheless, it is tempting to ask what the benefits of such ideal pricing would be, if only to gain some measure of the social costs of congestion. Even here, the computational problems of estimating these costs are considerable for any realistic model of a congested urban road network. It is, however, reasonably simple to compute the second-best level of traffic that can be supported by a set of trip-based road charges, if we are content to find what uniform reduction in the number of trips from each origin O to each destination D maximises social welfare. The appendix demonstrates that the average of the charges required is exactly equal to the difference between the MSC and APC averaged over the entire set of origin-destination (O-D) trips, provided that all trip users have the same constant elasticity of demand for trips, and respond equivalently to the cost of time (assumed uniform), the vehicle costs per km, and the trip charges. The appendix also shows that in simple two road examples, the average charge is close to the average of the first best efficient road charges, and the increase in social welfare achieved by uniform traffic scaling is close to the theoretical optimum, though it is less likely that this result would survive to more complex and varied urban networks. The procedure for computing the average charge and the resulting measures of welfare gain are summarised in Figure 1.

The idea of computing average charges for actual networks requires a traffic simulation model able to compute total travel costs from which marginal and average costs can then be derived. The model used is SATURN (Simulation and Assignment of Traffic to Urban Road Networks), developed at the Institute for Transport Studies at University of Leeds (Van Vliet and Hall, 1997). The program estimates the ‘generalised cost’ of trips as the sum of both the time cost and the vehicle operating cost:

$$GC_{ij} = VOT \times time_{ij} + VOC \times dist_{ij} \tag{1}$$

where $GC_{ij}$ is generalised cost in pence per PCU to go from origin zone $i$ to destination zone $j$, $VOT$ is value of time in pence per PCUmin, $time_{ij}$ is the time taken to complete the trip in minutes, $VOC$ is vehicle operating cost in pence per PCUkm, and $dist_{ij}$ is the distance travelled to go from origin zone $i$ to destination zone $j$, in km.[2] Time and distance vary according to the route chosen to go from origin zone $i$ to destination zone $j$ but in equilibrium no trip maker can reduce his or her $GC_{ij}$. VOC and VOT in this study were taken as 12.02 pence per PCUkm and 23.4 pence per PCUmin respectively (1998 prices). These values were computed as weighted averages taking into account vehicle and fuel type, vehicle occupations, trip purpose and average wages and value of leisure time, according to guidelines of the Highways Economics Note N°2 (Highways Agency *et al*, 1996).

**FIGURE 1 Average private cost and marginal social cost**



The SATURN model demonstrates that delays at junctions are the major source of urban congestion, in contrast to theoretical models that rely on speed-flow relationships derived from observations on links. In consequence the simulation is much closer to reality than other approaches. SATURN requires a network file and a trip matrix to run the model for a particular town. The network file contains the description of the network, particularly the capacities of links and the characteristics of junctions (priority, roundabouts, traffic signals, etc). The trip matrix is an O-D matrix that contains the number of vehicles (or PCUs) wishing to travel from origin zone $i$ to destination zone $j$ in the time period under consideration (e.g. the peak hour between 8-9am). The software simulates and assigns traffic in urban road networks and iterates until the (Nash) equilibrium is reached as defined above.

The average charge (BD in Figure 1) was computed for Northampton, Kingston upon Hull, Cambridge, Lincoln, Norwich, York, Bedford and Hereford, assuming a constant elasticity demand function:

---

[2] PCU is Passenger Car Units, a measure of the congestive effect of different vehicles. The average is taken over five different vehicle classes in proportion to their use of urban road space.

$$Q(P) = Q_0 * \left(\frac{P}{P_0}\right)^{-\eta}$$

where $Q$ is the demand for trips (measured in PCUs per hour), and $\eta$ is the demand elasticity, defined as a positive number and assumed constant.

Three elasticities were assumed for the calculations: 0.2, 0.4 and 0.7, spanning the plausible range of values in line with elasticity values found in previous studies (Goodwin, 1992; Oum *et al*, 1992; Fowkes *et al*, 1993; DETR, 1998; DETR, 2000; Victoria Transport Policy Institute, 2000). Only the results from the extreme values of 0.2 and 0.7 are presented here.

The average second-best charges computed for each town are presented in Table 1. They were computed as the difference between MSC and APC at the (constrained) efficient level of traffic, defined as the level at which the average marginal social cost and marginal social benefit are equal for an equal proportional reduction of all trips. The actual charges required to reduce each trip by the same proportion would vary by trip (and could be computed if necessary), and if imposed, would result in a reduction of traffic in Figure 1 from $K_a$ to $K_e$. Three areas were considered: the whole town including surrounding motorways and trunk roads, the whole town excluding surrounding motorways and trunk roads, and city centres.

One obvious feature of Table 1 is that the average charges per PCUkm clearly increase as the range of the model is narrowed to the more congested central areas, indicating that it is likely to be inefficient to reduce traffic uniformly over the whole town by the same amount. The table suggests that it would be desirable to reduce trips in the outer areas by a smaller proportion than in the central areas, though how this would be done when many trips that end in the congested centre travel through the less congested outer area is not immediately apparent from this averaging approach. In any case, it is not feasible, even if it were sensible, to impose trip-specific charges just to reduce all trips uniformly, and instead it makes sense to consider a more practical road charging scheme such as cordon tolls.

The final three columns of Table 1 show the results of multiplying the trip length by the charge per km to give the average charge per trip. The length of the trip in the inner areas is measured from the outer boundary that defines the area, so the charge per trip of these inner areas implicitly assumes that there is no road charging until the boundary of the area is reached. These estimates can be compared with the cordon tolls considered in the next section.

**Table 1. Average second-best charges**

| **Town** | **Efficient charge (pence per PCUkm) Area under study** | | | **Implied average trip charge (£) Area under study** | | |
|---|---|---|---|---|---|---|
| | Includes motorways | Excludes Motorways | Central area | Includes motorways | Excludes motorways | Central area |
| ***$\eta$ = 0.2*** | | | | | | |
| Northampton | 141 | 195 | 194 | 7.31 | 5.92 | 2.10 |
| Kingston upon Hull | 84 | Na | 159 | 6.29 | na | 3.21 |
| Cambridge | 42 | 79 | 109 | 2.68 | 2.64 | 2.01 |
| Lincoln | 42 | 42 | 114 | 2.19 | 1.62 | 1.98 |
| Norwich | 11 | 43 | 84 | 1.03 | 1.06 | 0.61 |
| York | 33 | 37 | 108 | 2.13 | 1.10 | 1.42 |
| Bedford | 9 | 40 | 47 | 1.77 | 1.30 | 0.87 |
| Hereford | na | 38 | 120 | na | 1.62 | 1.73 |
| ***$\eta$ = 0.7*** | | | | | | |
| Northampton | 84 | 118 | 110 | 4.29 | 3.53 | 1.18 |
| Kingston upon Hull | 51 | na | 88 | 3.77 | na | 1.77 |
| Cambridge | 25 | 47 | 63 | 1.59 | 1.56 | 1.16 |
| Lincoln | 27 | 26 | 68 | 1.40 | 1.00 | 1.17 |
| Norwich | 9 | 27 | 44 | 0.84 | 0.66 | 0.32 |
| York | 23 | 23 | 60 | 1.46 | 0.68 | 0.78 |
| Bedford | 6 | 25 | 24 | 1.18 | 0.81 | 0.44 |
| Hereford | na | 24 | 64 | na | 1.02 | 0.92 |

Source: SATURN results and own calculations.

## CORDON TOLLS

In a cordon-toll scheme a trip maker is charged to cross the cordon at certain times of the day, possibly charging only for entry in the morning and exit in the evening peaks. The charge does not depend on the time taken or distance travelled within the charged area, nor on levels of prevailing congestion. Only inbound cordons in the morning peak were simulated.

We chose to study a cordon pricing system because it seems to be relatively simple to implement and therefore potentially cost-effective (we present some estimates of cost and benefits below). Cordon tolling is already in use in Singapore, Oslo, Trondheim and Bergen. A few other Norwegian cities are about to implement it as well.

There are already electronic charging technologies available on the market, such as radio frequency tags or smart cards with radio frequency or infrared transponders. We therefore assume that cordon tolls would be collected using electronic road pricing (ERP) technology rather than manually. In the towns studied here, the charged area was defined as the city centre of the town, sometimes delimited by what the local authority defines as inner ring road, sometimes delimited by a judgement of what was the most congested area in that particular town. They reflect the existing pattern of location, residence, work and car ownership in 1998, as revealed by the trip matrices provided by the different local authorities.

The potential impacts of the schemes were estimated using results from SATURN and its batch file procedure to simulate road pricing, SATTAX.[3] SATURN finds the Nash equilibrium, while SATTAX allows the number of trips to respond to the cost of the trips and the routes chosen to depend on the tolls introduced. From these simulations, the reduction in traffic and average travel time were obtained. SATTAX simulates a toll as a time penalty for crossing the cordon. The time penalty required will depend on the value of time assumed (23.4 pence per PCUmin at 1998 prices). Thus a toll of £2.50 per crossing would be modelled as a delay of 641 seconds. SATTAX also simulates the effects of tolls, and models demand responses to changes in trip costs, using the algorithm, SATEASY. The demand response to tolls is modelled by specifying an elasticity of trip demand to the total cost of the trip (including the monetary costs of fuel, the toll, other vehicle ownership costs, and, quantitatively most important, the time cost of the journey).[4] The program thus allows for the impact that some drivers will be ‘tolled-off’ and other drivers will change route, resulting in fewer trips in the tolled (and more congested) area, though possibly with more trips in the immediate neighbourhood. Tolled-off trips will include trips that take place at other (non-charged) periods, or whose users switch mode, or those who decide the journey is not worth the extra cost. We are not able to take account of changes in the destination of trips, e.g. to alternative shopping sites.

## Cost benefit analysis

The criterion used to assess the benefits from a cordon toll was the increase in social surplus. Social surplus is defined as aggregate trip makers’ surplus, defined as the sum of individual utilities less the sum of individual social costs. In the case of a unique origin-destination pair, the utility of driving is the integral under the inverse demand function between zero (or some reference level of traffic, as in the Appendix) and the actual level of traffic. Individual social costs are the generalised costs defined in equation (1) adjusted to make them net of VAT and fuel duties (which are not part of the social costs). They are expressed in equation (2):

$$SC_{ij} = VOT \times time_{ij} + (VOC - VAT - duty) \times dist_{ij} \tag{2}$$

where $SC_{ij}$ is the social cost in pence per PCU to go from origin zone $i$ to destination zone $j$, $VAT$ is a weighted average of the Value Added Tax on fuel and duties and $duty$ is a weighted average of the average fuel duty paid by trip makers exclusive of VAT on duties. VAT and $duty$ in this study were assumed to be 0.82 and 4.3 pence per PCUkm (1998 prices). The sum of all $SC_{ij}$ can also be represented by the integral of the marginal social cost (MSC) between zero and the actual level of traffic.

### *Benefits*

The utility of trips from each origin to each destination is measured (in cash value) by the area under the demand schedule for such trips up to the actual level of traffic. The difference between drivers’ utility before and after the introduction of the toll was computed for each

---

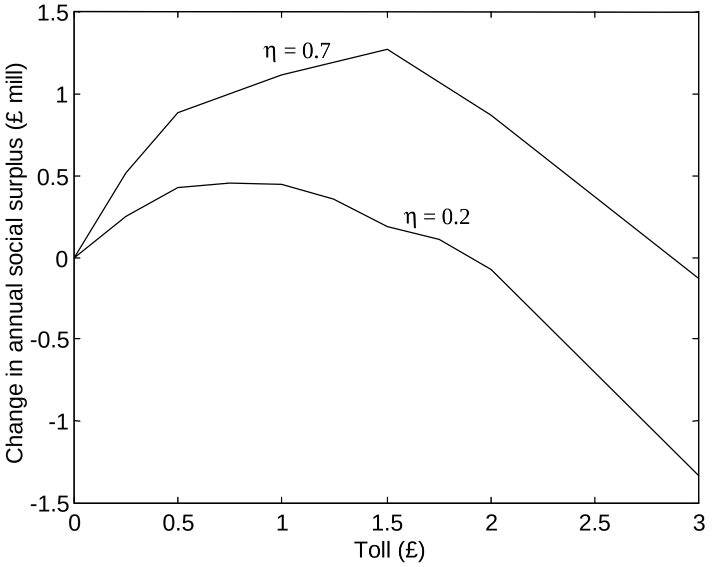
[3] SATURN is a software package developed by ITS, Leeds that allows us to compute the costs of vehicle trips for varying levels of traffic, while SATTAX computes the responses to the cordon tolls (Milne and Van Vliet, 1993).

[4] The elasticity of trip demand is defined as the percentage reduction in trips demanded for a 1% increase in the total trip cost, including the value of time taken.

origin-destination pair and then summed over all such pairs to give the overall change in surplus. The change in total costs was obtained directly from the new cost matrix produced by SATTAX. SATTAX was run for levels of tolls increasing in steps of £0.25 up to £4 and £5. The model was run for the morning peak (8 to 9am).

The optimal toll is the toll for which the change in social surplus, computed as the difference between the change in the sum of individual utilities of making trips *minus* the change in the sum of individual costs, reaches its maximum, as shown on Figure 2, which shows the effects of a high and low elasticity of trip demand. The higher the elasticity the higher the gain at any toll level. When we come to the full cost-benefit analysis, we have chosen to use the values for an elasticity of 0.2, to make sure we err on the side of underestimating rather than exaggerating benefits. If in fact the elasticity is higher, the schemes will be more attractive.

**FIGURE 2 Annual social benefit (£1998 million) for Cambridge**

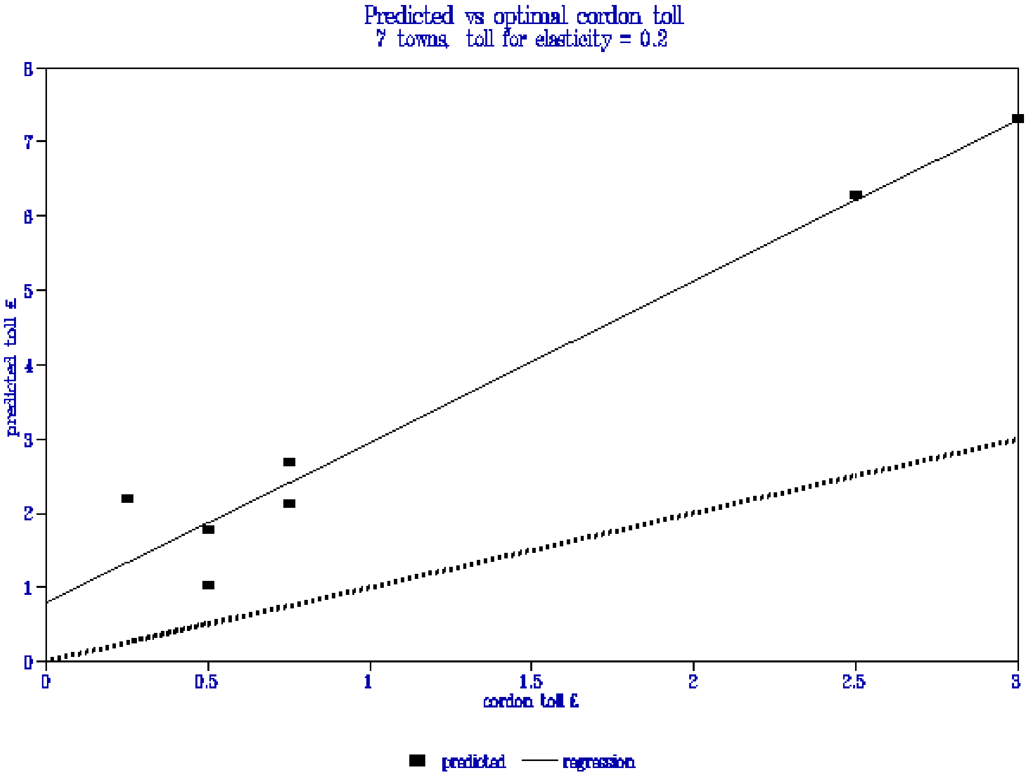


The increase in social surplus, i.e. the gross benefit, that would result from the introduction of this optimal toll was also computed. The increase in social surplus for a whole day was assumed to be three times that from 8-9 am. This is a conservative but reasonable assumption. The inefficiency is almost as high during the evening peak as during the morning peak (Newbery and Santos, 1999). Santos (2000) finds that the dead-weight loss or inefficiency of traffic congestion during the evening peak is typically between 70 and 90% that during the morning peak and can be even higher. The scheme would also improve social welfare in the shoulder peaks, i.e., the congested time-periods that surround the morning and evening peaks. Assuming that the introduction of a cordon toll during the morning and evening peaks and shoulder-peaks would yield an increase of only three times the peak value is therefore probably an underestimate. The results are presented in Table 2.

The annual gross revenues were computed as the number of vehicles that would cross the cordon multiplied by the toll that they would pay and by the number of working days (assumed to be 250) per year. The last four columns in Table 2 show the percentage change that would occur in average kilometres travelled (AKT), average travel time (ATT), total number of trips and number of PCUs crossing the cordon, if the optimal toll were introduced.

It is also interesting to compare the optimal tolls with the trip payments that would be incurred applying the second-best optimal road charges computed in the first part to the trip lengths in the tolled areas (either the whole town or some area within the ring road). The figures in the last three columns of Table 1 can be compared with the optimal toll in Table 2. The best fit is between the total trip costs for the entire town and the cordon toll (possibly because truncating trip lengths arbitrarily at various boundaries tends to underestimate the damage caused by trips that would cross the cordon). Although the correlation (leaving out Hereford, where the cordon toll behaves anomalously with the elasticity) is close for both elasticities (0.2 and 0.7), the average ratio of the trip charge to the optimal cordon charge is 3 for an elasticity of 0.2, and 1.06 for an elasticity of 0.7. Figure 3 shows the relation between trip charges and cordon tolls for an elasticity of 0.2, showing that the regression line is considerably steeper than the 45° line.

**FIGURE 3 Relation between trip charge and optimal cordon toll**



### *Costs*

The capital and operating costs of a cordon toll scheme for each town considered in this study are presented in Table 3. To estimate the costs the toll was assumed to operate from 7 to 10am and from 4 to 7pm. The toll should depend on the level of congestion and should be set lower at the shoulder-peak hours, 7-8 and 9-10am, and 4-5 and 6-7pm.[5] The number of vehicles crossing the cordon during this time period was deduced from the number of vehicles crossing the cordon between 8-9am using SATTAX and from the daily traffic

---

[5] A sophisticated toll would increase gradually from zero to the prescribed level and back over the shoulder periods to prevent bunching of trips around the beginning and end of the charged periods.

distribution inbound on Cambridge radial routes.[6] The number of transactions per day was estimated to be 3.7 times the number of transactions between 8-9am. The number of intra-vehicular units (IVUs) to be installed was assumed to be three times the daily number of cordon crossings. The IVUs' implementation costs, £15 for the tag, which is currently the cheapest option available on the market (Cheese and Klein, 1999), were multiplied by the number of IVUs required in each case. Infrastructure costs, of £45,300 per point (Cheese and Klein, 1999)[7], were multiplied by the number of cordon points. One fourth of the cordon points were assumed to be dual lane, and would therefore require gantries. According to Cheese and Klein (1999), the cost of one gantry is £97,000.

Operating costs, which include all costs of running the tolls, such as labour costs, costs of maintenance and costs of operating the infrastructure, were estimated using data from the Norwegian Public Roads Administration. These are in the order of 7 pence (10 US cents) per transaction at most. This figure was therefore multiplied by the number of transactions per day and by the number of days on which the scheme would operate per year, assumed to be 250. The IVUs were assumed to have a life of six years. The electronic devices in the infrastructure were assumed to have a life of five years and a value of £18,765 per cordon point (Cheese and Klein, 1999). The rest of the infrastructure was assumed to have the same life as the equipment and the scheme was assumed to last 30 years. In 30 years, the IVUs would have to be replaced four times, and the infrastructure, five. The net present value (NPV) of the costs is presented in the final column of Table 3.

The surprisingly lower benefit-cost ratio for the heavily congested town Cambridge (Table 4) prompted further investigation. The first experiment introduced an additional outer cordon. It was found that while one cordon around the city centre would be marginal, a double cordon scheme, one in the city centre and one for virtually the whole town, would be worth implementing, with benefits being considerably higher than costs. That suggested considering just the outer cordon by itself, and it was found that this would perform even better, with both social surplus and benefit-cost ratio increasing. This somewhat surprising result suggests that if the tolls for the inner and outer cordon could be separately optimised, then it should be possible to increase benefits further, though whether by enough to offset the higher operating costs is not clear.[8] We tried a few different combinations of inner and outer tolls for an elasticity of 0.2. We found that lowering the outer cordon to £4.75 while raising the inner cordon from zero to a modest level (£0.50) increased the daily benefit for the morning peak by a further 12%. This suggests the need for further investigation on combination of optimal tolls and also suggests that the precise location of the cordon may be critical in determining the benefits and hence the attractiveness of the scheme.

---

[6] We are indebted to James Lindsay, from WS Atkins, who provided us with data on vehicle counts in Cambridge.

[7] MVA (1995) estimates infrastructure costs at £110,000 per point. Cheese and Klein's (1999) estimate was chosen instead because it is more recent and prices for this type of equipment are likely to decrease with time and technological progress.

[8] In later research we plan to investigate optimising the combination of individually differentiated tolls of double cordons, though this involves substantially more computation.

## Table 2. Optimal cordon tolls for high and low trip demand elasticities

| Town | Elasticity at the original level of traffic | Optimal toll (£ to cross the cordon) | Increase in Individual surplus (p/PCUkm) | Benefit (£ mill./year) | Gross Revenues (£ mill./yr) | Ratio Revenue: Benefit | Changes in | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | AKT (%) | ATT (%) | No. of trips (%) | No. of PCU cordon crossings (%) |
| Northampton | 0.2 | 3.00 | 3.5 | 2.37 | 8.25 | 3.5 | 0.90 | -4.7 | -1.5 | -23 |
| | 0.7 | 3.50 | 7.3 | 4.85 | 8.32 | 1.7 | 0.20 | -9.9 | -2.8 | -33 |
| Kingston upon Hull | 0.2 | 2.50 | 3.7 | 3.24 | 7.69 | 2.4 | -0.60 | -6.1 | -1.1 | -15 |
| | 0.7 | 3.50 | 6.0 | 5.14 | 9.05 | 1.8 | -1.60 | -10.9 | -3.2 | -29 |
| Cambridge | 0.2 | 0.75 | 0.7 | 0.45 | 1.75 | 3.9 | 0.50 | -1.9 | -0.8 | -12 |
| | 0.7 | 1.50 | 1.9 | 1.27 | 2.80 | 2.2 | 1.10 | -6.3 | -3.0 | -29 |
| Lincoln | 0.2 | 0.25 | 1.2 | 0.44 | 0.52 | 1.2 | 0.06 | -2.6 | -0.2 | -9 |
| | 0.7 | 1.00 | 1.5 | 0.54 | 1.56 | 2.9 | 0.60 | -3.7 | -3.2 | -31 |
| Norwich | 0.2 | 0.50 | 0.9 | 0.90 | 1.24 | 1.4 | 0.90 | -2.5 | -0.9 | -18 |
| | 0.7 | 0.75 | 1.3 | 1.29 | 1.67 | 1.3 | 1.60 | -3.3 | -2.2 | -33 |
| York | 0.2 | 0.75 | 1.3 | 0.72 | 1.21 | 1.7 | 1.10 | -4.2 | -0.8 | -20 |
| | 0.7 | 1.50 | 1.5 | 0.87 | 2.23 | 2.6 | 2.50 | -6.2 | -3.6 | -39 |
| Bedford | 0.2 | 0.50 | 2.7 | 0.52 | 1.21 | 2.3 | -0.20 | -5.3 | -1.0 | -6 |
| | 0.7 | 1.50 | 1.9 | 0.35 | 2.70 | 7.7 | -1.00 | -7.0 | -8.4 | -30 |
| Hereford | 0.2 | 3.50 | 3.3 | 0.53 | 4.26 | 8.0 | 3.50 | -13.0 | -4.8 | -25 |
| | 0.4 | 1.75 | 4.8 | 0.77 | 2.20 | 2.9 | 2.30 | -14.0 | -4.8 | -23 |
| | 0.7 | 1.50 | 5.5 | 0.85 | 1.82 | 2.1 | 1.40 | -15.5 | -6.1 | -25 |

Source: Santos, Newbery and Rojey (2001)

Note: AKT: average kilometres travelled per trip, ATT: average travel time

**Table 3. Annual costs of implementing a cordon toll in different towns (£ 1998)**

| Town | | N° of crossings between 8 and 9 AM | Number of crossings per day | Number of IVUs | Number of cordon points | Implementation costs (£ million at 1998 prices) | | Operating costs (£ million at 1998 prices) | PDV of total costs (£ million at 1998 prices) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | IVUs Tag | Infrastructure | | |
| Cambridge | *Inner cordon* | 10,527 | 38,950 | 116,850 | 16 | 1.75 | 1.11 | 0.68 | 16.1 |
| | *Two cordons* | 21,407 | 79,206 | 158,412 | 26 | 2.38 | 1.88 | 0.92 | 22.3 |
| | *Outer cordon* | 10,880 | 40,256 | 120,768 | 10 | 1.81 | 0.70 | 0.70 | 15.9 |
| Northampton | | 14,189 | 52,499 | 157,498 | 12 | 2.36 | 0.83 | 0.92 | 20.6 |
| Kingston upon Hull | | 14,529 | 53,757 | 161,272 | 14 | 2.42 | 0.97 | 0.94 | 21.3 |
| Hereford | | 6,494 | 24,028 | 72,083 | 8 | 1.08 | 0.56 | 0.42 | 9.7 |
| Lincoln | | 9,074 | 33,574 | 100,721 | 20 | 1.51 | 1.39 | 0.59 | 14.5 |
| Bedford | | 10,335 | 38,240 | 114,719 | 14 | 1.72 | 0.97 | 0.67 | 15.6 |
| Norwich | | 12,164 | 45,007 | 135,020 | 22 | 2.03 | 1.53 | 0.79 | 19.0 |
| York | | 8,005 | 29,619 | 88,856 | 21 | 1.33 | 1.46 | 0.52 | 13.2 |

Source: See text

Note: IVUs and infrastructure costs are capital one-off costs that take place in year zero. IVUs and infrastructure will need to be replaced every five and six years. Operating costs are annual costs. The number of IVUs for the double cordon case was computed as the number of outer cordon crossings multiplied by 0.33 plus the number of inner cordon crossings, and like in the other cases, all multiplied by 3.

### *Comparison of costs and benefits*

Cordon tolling would be worthwhile only if the net present value (NPV) of benefits less costs were positive. Revenues are transfers, not benefits, and should not be part of the cost benefit analysis, though they are clearly of central interest to the charging authority and are the mechanism by which the costs are covered. There are additional benefits linked to the reduction in emissions, discussed below, but not included in these estimates of NPV. In addition, there would perhaps be a benefit resulting from fewer accidents. Traffic accidents, however, are related to both speed and traffic volume. If road charging increased speeds, accidents would increase, but since road pricing reduces traffic volumes, accidents would decrease. If a cordon toll increased speeds not by increasing running speed but by reducing queuing delays, the number of accidents would probably decrease (MVA, 1995), but the remaining accidents would (perhaps) be more severe. We have ignored this benefit as we believe this would require separate study to reach firm conclusions.

If there are distortions elsewhere in the urban economy, there would be a case for extending the analysis to value the impacts of transport changes on these distorted sectors, but this would be a major undertaking in its own right and not one we have considered. One obvious limitation of our approach is that SATURN is a medium-run model that holds car ownership and the O-D pattern trips constant. We have assumed that the changes in social surplus computed for the first year would hold during the whole life of the project. This is of course unrealistic. In the long run higher elasticities should be used as people and businesses might relocate and in addition, there may be changes in the local authorities' land use plans in response to changing transport demands. The model does not allow for such longer-term responses. To assess the full impact of any road-pricing scheme a more complex transport and land-use model would be needed, augmented with some view about likely local authority responses. It might be possible to forecast traffic growth, but it is likely that traffic management arrangements would be adapted to deal with such growth and the existing model would then no longer represent the network correctly. Our defence of the simplifying assumption of constant traffic is that the long-run impacts of relocation caused by road pricing are likely to reduce traffic, while economic development is likely to increase traffic, making a no-change assumption not unreasonable. If anything, it is likely to underestimate the benefits of road pricing.

Table 4 summarises the costs and benefits for cordon tolls in the eight towns.

**Table 4. Net Present Value of a cordon toll in different towns (£1998 million)**

| Town | | Total cost | Benefit | Net Present Value | Benefit/Cost |
|---|---|---|---|---|---|
| Cambridge | *Inner cordon* | 16.1 | 17.6 | 1.5 | 1.1 |
| | *Two cordons* | 22.3 | 63.6 | 41.4 | 2.9 |
| | *Outer cordon* | 15.9 | 89.6 | 73.7 | 5.6 |
| Northampton | | 20.6 | 90.0 | 69.4 | 4.4 |
| Kingston upon Hull | | 21.3 | 123.5 | 102.2 | 5.8 |
| Hereford | | 9.7 | 19.5 | 9.8 | 2.0 |
| Lincoln | | 14.5 | 17.0 | 2.4 | 1.2 |
| Bedford | | 15.6 | 19.9 | 4.3 | 1.3 |
| Norwich | | 19.0 | 23.9 | 4.9 | 1.3 |
| York | | 13.2 | 27.4 | 14.2 | 2.1 |

Source: Own calculations

Note: discount rate: 6%, benefits assumed to be constant throughout the 30 years

The 1998 Treasury test discount rate of 6% was used.[9] A comparison of costs and benefits indicates that road pricing would be very beneficial in Kingston upon Hull and Northampton, and to a lesser extent in Hereford and York, on our conservative estimates. These schemes would become more beneficial if costs proved to be lower or if the elasticity of the demand proved to be higher than 0.2.

Two cordons in Cambridge, one inner and one outer, each charging £1.50 per crossing, yield a benefit-cost ratio almost three times as high that of an inner cordon scheme, thus making it attractive. A single outer cordon charging £5 increases the benefit-cost ratio to more than five times that of a single inner cordon. An outer cordon charging £4.75 combined with an inner cordon charging £0.50 yields higher benefits than an outer cordon implemented on its own. The costs are also higher because there are two cordons instead of one, but the net benefit is still increased, making this the preferred option.

This shows that where a single inner cordon scheme might not be worthwhile, an outer cordon scheme could be, and a double cordon, but with each charge optimally set, might do even better (though if the tolls are not carefully set, much of the benefit may be lost). Varying the location and possibly the number of cordons is therefore likely to be worth investigating before deciding on the desirability of a road pricing scheme in any one town, for even where the viability of a single cordon is not in doubt, the benefits of one at a different location, possibly combined with an additional cordon, may greatly improve the outcome.

### ***Environmental impacts of cordon tolls***

The next issue to address is whether the choice of cordon tolls is likely to be significantly affected by the environmental consequences of reduced traffic (which should be beneficial) and possibly increased speeds (which may improve fuel efficiency, further enhancing the benefits, but which may increase some pollutants, and hence have adverse effects). Although cordon tolling may encourage more and/or longer trips (Richardson and Bae, 1998), the simulations for the eight towns show that in all cases there would be positive environmental benefits, at least for the major health and global warming impacts. Unfortunately, the range of environmental cost estimates of road transport emissions is very wide, though recent work is doing much to narrow the range. The high estimates of the total environmental costs presented below are about 15 times as high as the low estimates (showing the considerable uncertainty attached to the figures). Fortunately, this uncertainty has little practical effect, as even the high cost estimates are modest compared to the traffic efficiency gains.

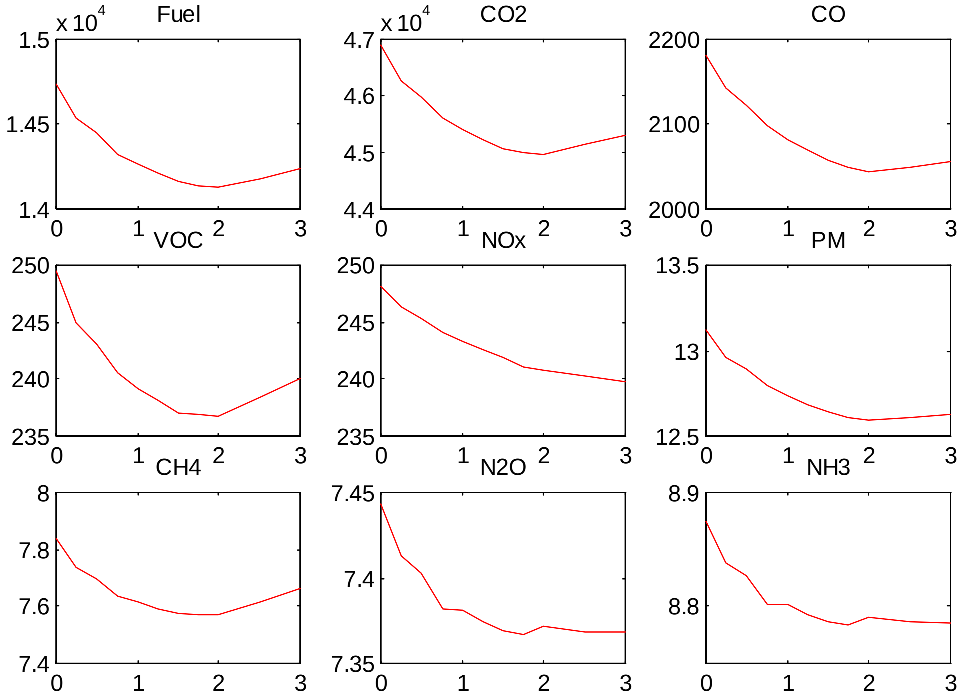
Santos, Rojey and Newbery (2000) present the results of valuing the emissions, and they are summarised here. The evaluation of emissions was based on the methodology described in Chapter 7 of European Environment Agency (2000). The emissions factors were obtained by applying these formulae to the average speed obtained from SATURN for each trip defined by its origin and destination. The pollutants we considered are carbon dioxide ($CO_2$), carbon monoxide (CO), volatile organic compounds (VOC), nitrogen oxide ($NO_x$), particulate matter (PM), methane ($CH_4$), nitrous oxide ($N_2O$) and amonia ($CH_3$), though only health impacts and global warming effects were valued in the cost analysis. As an example, we present the emissions (kilograms of pollutant) for Cambridge when the trip elasticity was assumed to be 0.4 (i.e. in the middle of the range). Emissions were found to vary with the level of tolls introduced. The results are reported on Figure 4. Tolls in £/PCU are shown on the *x*-axes, and the *y*-axes are truncated.

---

[9] The Treasury was, in early 2001, reconsidering the test discount rate and may reduce it somewhat. If so, the benefit-cost ratio would be increased.

The next step is to value each emission, and then sum to give the relationship between the total cost of emissions and the level of tolls, to see to what extent the traffic benefits are correlated with the environmental benefits.

**FIGURE 4 Reduction of emissions for each pollutant at different levels of tolls**



The corresponding monetary value of the reduction in emissions in these eight towns and their estimates are presented in Table 5. The table shows that, even when using the highest estimate for pollution costs, the increase in benefit caused by the reduction in emissions is small (typically less than 10%) compared to gains from improved traffic efficiency and time saved. Figure 5 below shows that the emissions benefits (for an elasticity of 0.2) correlate well with the transport benefits of cordon tolls in Northampton. In all cases the optimum toll for transport purposes was the same as the optimal toll for minimising environmental costs.

**FIGURE 5 Emission benefits and transport benefits vs. cordon tolls for Northampton**

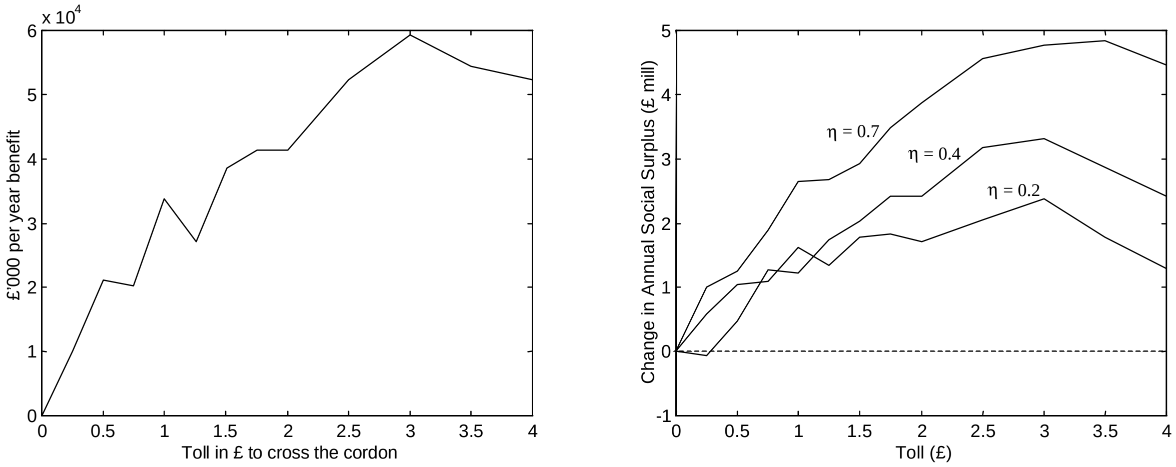


**Table 5. Reduction in environmental costs from optimal cordon tolls**

| Town | Elasticity | Estimate | Environmental benefit £'000 per year | | | As % of transport benefits |
|---|---|---|---|---|---|---|
| | | | Health | Global Warming | Total | |
| Northampton | 0.2 | Low | 2.7 | 1.3 | 4.0 | 0.2 |
| | | High | 39.9 | 19.5 | 59.4 | 2.5 |
| | 0.7 | Low | 8.1 | 4.7 | 12.8 | 0.3 |
| | | High | 122.3 | 69.7 | 192.0 | 4.0 |
| Kingston upon Hull | 0.2 | Low | 5.9 | 4.2 | 10.1 | 0.3 |
| | | High | 90.0 | 62.2 | 152.2 | 6.8 |
| | 0.7 | Low | 13.2 | 8.9 | 22.1 | 0.4 |
| | | High | 201.5 | 132.2 | 333.6 | 6.5 |
| Cambridge | 0.2 | Low | 1.5 | 0.8 | 2.3 | 0.5 |
| | | High | 23.2 | 11.6 | 34.9 | 7.8 |
| | 0.7 | Low | 5.2 | 3.1 | 8.3 | 0.7 |
| | | High | 80.6 | 45.6 | 126.2 | 9.9 |
| Lincoln | 0.2 | Low | 2.4 | 1.3 | 3.7 | 0.8 |
| | | High | 37.0 | 18.7 | 55.7 | 12.7 |
| | 0.7 | Low | 1.5 | 0.9 | 2.4 | 0.4 |
| | | High | 23.8 | 12.9 | 36.7 | 6.8 |
| Norwich | 0.2 | Low | 1.5 | 1.5 | 3.0 | 0.3 |
| | | High | 22.4 | 21.6 | 44.1 | 4.9 |
| | 0.7 | Low | 3.2 | 2.6 | 5.8 | 0.4 |
| | | High | 48.3 | 38.3 | 86.6 | 6.7 |
| York | 0.2 | Low | 1.2 | 0.8 | 2.0 | 0.3 |
| | | High | 17.6 | 12.6 | 30.2 | 4.2 |
| | 0.7 | Low | 3.2 | 2.1 | 5.3 | 0.6 |
| | | High | 48.8 | 31.0 | 79.7 | 9.2 |
| Bedford | 0.2 | Low | 0.8 | 0.5 | 1.3 | 0.3 |
| | | High | 12.4 | 7.1 | 19.5 | 3.8 |
| | 0.7 | Low | 4.2 | 2.3 | 6.5 | 1.9 |
| | | High | 64.3 | 34.5 | 98.8 | 28.3 |
| Hereford | 0.2 | Low | 2.1 | 1.5 | 3.6 | 0.7 |
| | | High | 31.1 | 22.3 | 53.4 | 10.0 |
| | 0.7 | Low | 3.1 | 2.0 | 5.1 | 0.6 |
| | | High | 46.5 | 30.2 | 76.7 | 9.1 |

Source: Santos, Rojey and Newbery (2000)

## CONCLUSIONS

The paper started from the traditional method of estimating congestion charges used on links and adapted it to networks. This allowed a ready derivation of the second-best average toll that would reduce all trips by the optimal uniform amount. While this is a good approximation to the average first best toll for simple (two-road) systems, it is less likely to be a good approximation for complex networks. The method does not give the individual optimal road charges, and in any case their imposition would be neither feasible nor sensible. The calculations may be useful for indicating the severity of congestion in individual towns, and thus suggesting where more practical solutions may be desirable. One such practical form of congestion pricing is cordon tolling, already in use in a number of countries. We computed the optimal cordon toll that maximises social surplus after allowing for driver response (who may decide to reduce the number of trips, change the time of their trips to avoid tolls and congested periods, or use different non-tolled routes), in eight English towns.

The cost-benefit analysis shows that even with the cheapest available technology (assumed in this study) there were borderline cases with benefit-cost ratios only slightly above one.

This is the case of Lincoln, Bedford, Norwich and Cambridge (with a single cordon). It appears to be possible to considerably raise the benefit-cost ratio by replacing it by an outer cordon, at least judging by the example of Cambridge, where with one cordon only around the city centre tolls are at best marginal, whereas with an outer cordon, the benefit-cost ratio rises to 5.6. York and Hereford are towns with positive but modest net present values (and benefit-cost ratios of around 2). Finally, there are two unambiguous cases: Northampton and Kingston upon Hull, where the benefit-cost ratio is well above 4, and the net benefits are sizable.

Back-office costs are the most significant usage-related cost and the one on which the viability of the schemes appears to depend most sensitively. It may be that improved electronic processing can substantially reduce these costs. If so, more schemes will become socially desirable. In the mean time, the most important lesson to draw from this study is that trials should be carefully targeted at the few towns (such as Kingston upon Hull and Northampton) that appear to have a considerable excess of social benefits over costs. These trials should be studied closely, to validate the travel responses upon which the benefits and cost so sensitively depend and to validate the estimates of implementation and operating costs. That knowledge will allow subsequent schemes to be better designed and subsequent cost-benefit analyses more accurately undertaken.

Cordon tolls also yield environmental benefits that in all cases are closely correlated with the transport benefits. Although it is hard to measure their absolute magnitude at all accurately, even very high values give additional benefits that are rarely greater than 10% of the traffic benefits. As the environmental benefits are both modest and closely correlated with the transport benefits there is little advantage in modeling them separately, as they are unlikely either to change the optimal toll or the viability of the scheme.

## ACKNOWLEDGEMENTS

Support from the Economic and Social Research Council (ESRC) under Grant R000223117, ‘Road Pricing and Urban Congestion Costs’, and from the Department of the Environment, Transport and the Regions (DETR), under Contract N° PPAD 9/99/28, is gratefully acknowledged. Any views expressed in this paper are not necessarily those of the ESRC, DETR, or University of Cambridge.

The authors are grateful to Laurent Rojey, with whom they co-authored two papers that gave origin to the present study, and to Prof. Dirck Van Vliet and Dave Milne of the Institute for Transport Studies at University of Leeds for help with SATURN and SATTAX. Thanks are also due to Markus Kuhn from the Computer Lab at Cambridge University for computer assistance.

Provision of data by the following local authorities and consultancies working for them is gratefully acknowledged: Director of the Environment and Transport Department of Cambridgeshire County Council, WS Atkins, Herefordshire Council, Lincolnshire County Council, Symonds Group, York City Council, Northamptonshire County Council, Kingston upon Hull City Council, Bedfordshire County Council, and Norfolk County Council.

## REFERENCES

Cheese, J. and G. Klein (1999), *Charging Ahead: Making Road User Charging Work in the UK*, A Trafficflow Project Report, The Smith Group Ltd, Guildford, April.

Department of the Environment, Transport and the Regions (1998), *An Approach to Estimating the Welfare Costs of Congestion through the NRTF 1997 Framework*, DETR, London, April.

Department of the Environment, Transport and the Regions (2000), Modelling using the National Road Traffic Forecasting Framework for Tackling Congestion and Pollution and Transport 2010: The 10 Year Plan, Technical Report, DETR, London, December.

European Environment Agency (2000), EMIP/CORINAIR Atmospheric Emission Inventory Guidebook, European Environment Agency, Second Edition http://binary.eea.eu.int:80/g/group07.pdf

Fowkes, A. S., Sherwood, N. and C. A. Nash (1993), 'Segmentation of the Travel Market in London: Estimates of Elasticities and Values of Travel Time', *Working Paper 345,* Institute for Transport Studies, University of Leeds, Leeds, May.

Goodwin, P. B. (1992), 'A Review of New Demand Elasticities with Reference to Short and Long Run Effects to Price Changes', *Journal of Transport Economics and Policy*, Vol. 26, N° 2, pp. 155-169.

Highways Agency, Scottish Office Development Department, The Welsh Office, The Department of the Environment for Northern Ireland and the Department of Transport (1996), *Highway Economics Note N°2,* in *Design Manual for Roads and Bridges*, Vol. 13, HMSO, London.

Milne, D. and D. Van Vliet (1993), 'Implementing Road User Charging in SATURN' *ITS Working Paper 410,* Institute for Transport Studies, University of Leeds, Leeds, December.

MVA (1995), *The London Congestion Charging Research Programme,* Final Report, Vol. 1: Text, Government Office for London, HMSO, London.

Newbery, D. M. (2000a), 'Measuring Congestion Costs Geometrically', *Mimeo,* Department of Applied Economics, Cambridge, December.

Newbery, D. M. and G. Santos (1999), *Quantifying the Costs of Congestion*, End of Award Report, ESRC Grant R000222352, Department of Applied Economics, University of Cambridge, Cambridge, December.

Oum, T. H., Waters II, W. G. and J-S. Yong (1992), 'Concepts of Price Elasticities of Transport Demand and Recent Empirical Estimates', *Journal of Transport Economics and Policy,* Vol. 26, N° 2, pp. 139-154.

Richardson, H. W. and C-H. C. Bae (1998), 'The Equity Impacts of Road Congestion Pricing', *in* Button and Verhoef (1998).

Santos, G. (2000), 'Comparison of the Social Costs of Congestion during the Morning and Evening Peaks in Four English Towns', *Mimeo,* Department of Applied Economics, Cambridge, October.

Santos, G., Newbery, D. and L. Rojey (2001), 'Static Vs. Demand Sensitive Models and the Estimation of Efficient Cordon Tolls: An Exercise for Eight English Towns', *Transportation Research Board Record,* forthcoming. Also published in the CD ROM of the 80th Annual Meeting of the Transportation Research Board, Washington DC, USA, January 7-11 '01.

Santos, G., Rojey, L. and D. Newbery (2000), 'The Environmental Benefits from Road Pricing', *DAE Working Paper 0020,* Department of Applied Economics, Cambridge, December.

Van Vliet, D. and M. Hall (1997), *SATURN 9.3 - User Manual*, The Institute for Transport Studies, University of Leeds, Leeds.

Victoria Transport Policy Institute (2000), 'Transportation Elasticities', *in TDM Encyclopedi*a. http://www.vtpi.org/tdm/tdm11.htm. Accessed on Dec. 14, 2000.

## Appendix: Proof of average second best toll

Figure 1 is standard for representing congestion in the literature, but it only applies when traffic can be measured by a scalar variable (in the figure, traffic flow per lane). If the diagram is to be interpreted for urban road networks, we need some method of converting the complex pattern of flows on links into a scalar measure of overall traffic. SATURN provides a simple method of scaling flows by changing the number of trips between each origin and destination (OD) pair in proportion. We can number the set of OD trips by *i*, where $i = 1,2,3,...m.n$, if there are *m* origins and *n* destinations.

We now need to define analogues of the network average and marginal costs and demand prices. To that end, define the following notation:

- the social cost for the trip *i* is $c_i$ pence/PCU,
- the private cost of trip *i* is $p_i$ pence/PCU (the perceived effective total cost of time and distance),
- the length in km of trip *i* is $d_i$ km,
- the number of trips *i* is $q_i$ PCU, $q_i = q_i(p_i)$,
- the consumer *utility* (measured in cash terms) of the total number of trips *i* is $U^i(q_i)$, where $dU/dq = f_i(q_i) = p_i$ in equilibrium.

If the level of traffic relative to the equilibrium is measured by the scalar θ, then the actual number of OD trips *i* is $q_i = \theta q_{i0}$ for all *i*. The consumer utility can be also expressed as an integral of the inverse demand schedule:

$$U_i(q_i) = U(\theta q_{i0}) = \int_{Q_m}^{\theta Q_{i0}} f_i(q)dq$$

where $q_m$ is some arbitrary but fixed minimum level, possibly zero, to ensure the finiteness of the integral.

Social welfare at traffic level θ will be $W(\theta) = \Sigma U^i(\theta q_{i0}) - C(\theta)$, where $C(\theta)$ is the total social cost of trips, $\Sigma\theta q_{i0}.c_i(\theta q_{i0})$. The uniform reduction that maximises social welfare (the constrained or second-best optimum)[10] can be found by differentiating $W(\theta)$ (using the integral form of consumer utility):

$$\frac{dW}{d\theta} = 0 \Leftrightarrow \frac{dC}{d\theta} = \sum_i q_{i0} f_i(q_i).$$

To relate this to the various curves in figure 1 we need to express prices and costs in pence/PCUkm. The total PCU km travelled (PCUKT) on trip *i* is $d_i q_i$, and total PCUKT is $D = \Sigma d_i q_i$. The average private cost, *a*, and the marginal social cost, MSC, *m*, are then:

$$a = \frac{\sum_i p_i q_i}{D}; m = \frac{d}{dD}\left(\sum_i c_i q_i\right).$$

To a close approximation, $D = \theta D_0$, where $D_0 = \Sigma q_{i0} d_{i0}$, so at the optimum

$$m = \frac{dC}{dD} = \frac{1}{D_0}\frac{dC}{d\theta} = \frac{\sum_i q_{i0} f_i(q_i)}{\sum_i q_{i0} d_{i0}}.$$

---

[10] Second best because all trips are constrained to be equiproportionately reduced, rather than all trips being varied to maximise social welfare.

It remains to define the average price such that at the optimum, the MSC is equal to the price. Fortunately, the condition will be satisfied with the natural interpretation that the average price is the trip weighted price per PCUkm:

$$P = F(\theta) = \frac{\sum_i q_i f_i(q_i)}{D} = \frac{\sum_i \theta q_{i0} f_i(q_i)}{\theta D_0},$$

as required for (second-best) optimality. It also follows that the deadweight loss is correctly measured by the area between the demand schedule, the MSC schedule, and the vertical line through the market equilibrium at θ = 1.

The final question is whether the required tax in figure 1, $P(\theta)$-$a(\theta)$, is the average of the taxes required to reduce each trip to a fraction θ of its original level. The tax required on trip $i$ is $t_i = f_i(q_i) - p_i(q_i)$, so the average tax per PCUkm, $t$, is given by

$$t = \frac{\sum_i t_i q_i}{D} = \frac{\sum_i q_i f_i(q_i)}{D} - \frac{\sum_i q_i p_i(q_i)}{D} = m - a.$$

Consequently, the measured average corrective tax derived in figure 1 is equal to the road charges that would have to be levied on each trip to reduce demand by the desired amount. Note that these road charges would have to be levied solely as a function of origin and destination to avoid influencing the choice of route (other than in response to the changed level of traffic). Such road charges are doubly inefficient relative to the first best, in that they do not discourage traffic from the most congested parts of town, nor do they penalise trips with higher external costs more heavily than those with low external costs do. They do, however, allow one to place meaning on the various costs, externalities and corrective charges in figure 1.

## Approximations to efficient tolls

The next question to address is how well the average second-best congestion charge computed above (i.e. the average of the taxes required to optimally reduce traffic on each link by the same proportion), corresponds to the average of the (first-best) efficient charges (i.e. those that are such that at the efficient set of link and junction charges, the MSC of each trip is equal to the demand price at that level of traffic). This is computationally demanding to estimate for a complex network, but it is straightforward to simulate a two-road example. With the same values for time and distance costs as in the paper, and a (constant) demand elasticity of 0.3, two roads with (untaxed) equilibrium lane flows that are quite widely different (e.g. 400 and 1000 PCU/hr) and quite different optimal traffic reductions (6% and 16%) achieve 99.8% of the theoretical maximum social welfare at a reduction of 15%, and the average of the efficient taxes is 97.5% of the second-best average tax. The reduction in dead-weight loss achieved by the uniform reduction is 97.8% of the theoretical maximum achievable with optimal charges. At least in simple configurations, the average tax computed by this simple method seems a reasonable approximation to the average efficient tax.