Efficiency and equity of congestion charges

Ida Kristoffersson – CTS
Leonid Engelson – CTS


Abstract
Efficiency of congestion charging schemes has been extensively studied in road pricing literature. However, few studies analyze both efficiency and equity of congestion charging schemes. This paper shows the importance of conducting an equity analysis as a complement to an efficiency analysis. Comparing different charging scenarios for Stockholm, the paper shows that changing the location of the charging stations may alter the system from progressive to regressive. In the paper, the most efficient scenario is the least equitable. Indeed, the results of this paper show that moving towards a more efficient scheme design, where amounts are more closely related to congestion level, the charging system turns from progressive to regressive. The reason is the uneven distribution of workplaces and residential areas in Stockholm. Combined with richer socio-economic groups to a larger extent living in the part of Stockholm with more workplaces, this leads to a trade-off between charging for congestion and designing an equitable system. The paper concludes that congestion charging cannot be said to be progressive or regressive per se, rather it varies between cities and even between different scheme designs for the same city. Furthermore, results of the mesoscopic simulations performed in the paper demonstrate that travelers as a whole may benefit from congestion charging even before the use of revenues to compensate the users.

Keywords: Congestion charging; Efficiency; Equity; Welfare effects; Regressive; Progressive; Mesoscopic simulation;

JEL Codes: R41, R48
1 INTRODUCTION

Urban congestion charging has for a long time been advocated by economists because of its possible efficiency gains of internalizing the external cost of congestion (Vickrey 1963). Furthermore, a major advantage of congestion charging is that the decision of which trips to cancel or change, so that they are no longer subject to pricing, stays with the user (Fosgerau and Van Dender 2010). In the last decade, congestion charging has emerged as a viable travel demand management tool for cities that face increasing levels of congestion and want to deal with its negative consequences such as time losses, uncertain travel times and high emission levels. The experiences of real-world implementations or feature studies of congestion charging systems in sixteen European cities has been summarized and compared in the CURACAO project (May et al. 2009).

The implementation of congestion charging in any city is preceded by a design process in order to decide the location, timing and level of charges. Experts designing the scheme will try to charge for congestion, i.e. place the charging stations at bottlenecks. Experts will in general also move towards a more complex scheme design, with charges differentiated depending on time of day, time of year, location and direction of traffic, in order to increase the efficiency of the charging scheme, i.e. to achieve the highest socio-economic benefit. Policymakers on the other hand often seek a simple design which is easily communicated to the public, e.g. driving out of the city centre is as expensive as driving into the city centre. On this topic Eliasson (2010) notes that: There is a conflict between “effective” and “easily communicated” design, but erring towards the too simple seems more common (p. 6). Not only seek policymakers a simple design, they seek also a design in which the charge one has to pay is similar for different user groups. At the same time, some user groups are often exempted from the charges or get a discount motivated by equity considerations.

Equity issues are often recognized as an important part of public resistance to implementation of congestion charging (Viegas 2001; Oberholzer-Gee and Weck-Hannemann 2002). The issue of equity can also be seen as a constraint on to what extent efficiency and environmental objectives of congestion charging can be fulfilled (Sumalee 2003). Opponents of congestion charging argue that low-income groups are priced off the road and that those, who can afford to pay the charge, benefit from reduced travel times in the road network. This argument that low-income groups lose and high-income groups gain supports the notion that congestion charging is regressive. However, the opposite opinion – that congestion charging is progressive – also exists. The argument is then that high-income groups pay most charges and that low-income groups are not affected much if a good public transport alternative exists and has capacity enough to handle the increased demand.
There are (at least) two traditions in evaluating equity of a transport policy: a welfare-based approach that stems from economic research and a transportation access approach that stems from traffic planning research (Ecola and Light 2009). The transportation access approach concentrates on individuals that are for some reason disadvantaged in the transportation system, may it be because of income, gender, age, ethnicity or another reason (Rajé 2003). Common methods used in the transportation access approach are surveys and focus groups with representatives of the transportation disadvantaged. The welfare-based approach is based on evaluation of benefits and losses induced by the policy under consideration. There are several welfare-based methods, differing in their degree of complexity. When evaluating equity effects of congestion charging, the simplest welfare-based equity analysis compares how much individuals or groups pay in charges. More advanced analyses take into account also travel time gains, behavioural responses to congestion charging, such as changing mode or departure time, and how revenues are spent. The Gini-coefficient is commonly used to assess the distribution of incomes in the population. A Gini-coefficient equal to 1 means total inequality where all incomes belong to one person in the population, whereas a Gini-coefficient equal to 0 means total equality with incomes spread equally in the population. A few studies have evaluated the impacts of congestion charging on the Gini-coefficient (Karlström and Franklin 2009; Maruyama and Sumalee 2007).

In Stockholm we have the advantage of a real-world implementation, with substantial amount of data collected before and after the introduction of congestion charging in 2006. The introduced charging scheme was during its first ten years of operation a simple cordon around the inner city. On the cordon surrounding the inner city the same amount is charged at all points of entry and in both driving directions. There is however a differentiation by time of day, where the charge varies in steps before and after the morning and afternoon peak hours. The same charge for inbound and outbound driving direction was introduced for simplicity and because politicians were worried that unequal charges depending on driving direction could be seen as favoring residents in the inner city (Eliasson 2008).

In this paper, alternative congestion charging schemes for Stockholm are studied using a dynamic transport model called SILVESTER (Kristoffersson and Engelson 2009). Choice of departure time is modelled in SILVESTER based on a distribution of preferred departure times between 6:30 and 9:30 AM. Car route choice, travel times and time spent in queues are calculated using mesoscopic traffic simulation. SILVESTER models car users, but they can switch to public transport, which makes it possible to evaluate the mode switch effect induced by a specific congestion charging scheme. The alternative scenarios, compared in this paper to the charging scheme introduced in Stockholm in 2006, emanate from an objective to improve efficiency. From the efficiency point of view it is preferable to relate the charged amount as much as possible to the level of congestion. The charging locations and amounts in the alternative scenarios of this paper are based on the results of
Ekström et al. (2009), in which heuristic algorithms are used to find second best pricing schemes for an aggregated Stockholm network using static traffic assignment. In a first step Ekström et al. (2009) relax the amount charged at different stations and in a second step also the locations where charges are levied, such that the charge is more closely related to congestion level. Results of Ekström et al. (2009) show that differentiating the levied amounts by charging station and driving direction can improve social surplus (consumer surplus + revenues) by about 32%. Typically the authors find that from an efficiency point of view the charge is too low on inbound links and too high on outbound links in the morning. The reason is that during this part of the day there is more traffic going into the central business district than going out. Relaxing also the location of the charge, Ekström et al. (2009) find that social surplus can be improved by approximately 225% compared to the cordon implemented in Stockholm in 2006.

It is however important to also investigate equity effects of differentiated charging levels and/or new locations where the charge is levied. One needs to consider the socio-economic differences in geographical living areas, for example differences between the inner city (inside the cordon) and the suburbs. Given that households in the inner city of Stockholm have a relatively high income, raising the inbound and lowering the outbound charge in the morning could put a high burden on car-users who are already worse off. Such considerations were not included in the calculations of optimal second-best charges in Ekström et al. (2009).

The research literature on urban congestion charging is extensive. A review that describes state-of-the-art in design and evaluation of congestion charging is Tsekeris and Voss (2009), which includes more than 400 references. However, as Tsekeris and Voss (2009) point out, most of this research has dealt with idealistic pricing regimes such as marginal social cost pricing (MSCP) or second-best pricing in small laboratory networks. Out of these, studies that include welfare analysis are for MSCP e.g. Vickrey (1969), Small (1983) and Arnott et al. (1994), and for second-best pricing in small networks e.g. Liu and McDonald (1999), De Palma et al. (2005) and Rouwendal and Verhoef (2006). However, only a few studies investigate both efficiency and equity issues, e.g. for Washington DC (Safirova et al. 2004), Utsunomiya area in Japan (Maruyama and Sumalee 2007), Stockholm (Eliasson and Mattsson 2006; Franklin 2005), and a comparison for Cambridge, Northampton and Bedford (Santos and Rojey 2004).

The discussion about the equity impacts of congestion charging continues. Using data and models developed in the wake of the Stockholm congestion charging system, we in this paper further investigate the question: Is congestion charging regressive or progressive?

---

1 Because of long simulation run-times it was not possible to use SILVESTER to find optimal second-best charging amounts and locations.
Indeed, the results of this paper show that moving towards a more efficient scheme design, where amounts are more closely related to congestion level, the charging system turns from progressive to regressive. The reason is the uneven distribution of workplaces and residential areas in Stockholm. The density of dwellings is considerably lower on the north side (more one-family-houses) whereas the density of workplaces is considerably higher, which results in more traffic and thereby congestion in the northbound direction in the morning. Combined with richer socio-economic groups to a larger extent living on the north side this leads to a trade-off between charging for congestion and designing an equitable system. In general, results show that congestion charging is not regressive or progressive per se, it depends on the structure of the city in question and on the design of the charging scheme.

The paper continues in the next section with a short description of the Stockholm congestion charging scheme. Section 3 then describes the transport model set up for Stockholm, the basic features of the analysed scenarios and the method for calculating welfare changes for each socio-economic group. Section 4 reports and analyses the results of each scenario and Section 5 concludes.

2 THE STOCKHOLM CHARGING SCHEME

The Stockholm congestion charging scheme was introduced as a trial in 2006 (Eliasson et al. 2009) and made permanent from autumn 2007 (Börjesson et al. 2012). By the time of introduction\(^2\), the scheme consisted of a simple cordon around the inner city (Figure 1). The peak hour charge was 20 SEK per crossing, the shoulder charge 15 SEK and the off-peak charge 10 SEK (see Figure 1 for the exact charging schedule). The same amount was charged at all points and for both inbound and outbound driving directions. The motorway Essingeleden (solid line in Figure 1) was free of charge.

\(^2\) From January 2016 a charge is also levied on the Bypass Essingeleden and the charge on the cordon around the inner city has been increased, especially during peak hour with a current charge of 35 SEK.
Figure 1: The Stockholm congestion charging scheme with the cordon around the inner city (dashed line), the Essingeleden Bypass Motorway (solid line) and the charging stations (red dots). Source: (Eliasson et al. 2009). The charging schedule is given in the table next to the figure.

3 METHOD

3.1 The SILVESTER Transport Model

The SILVESTER transport model (Kristoffersson and Engelson 2009) simulates car traffic in Stockholm from 6.30 to 9.30 am. SILVESTER consists of two parts: 1) a mixed logit departure time and mode switch model in which car users choose which fifteen minute time interval to depart in, alternatively choose to switch to public transport, and 2) the mesoscopic traffic simulation model CONTRAM (Taylor 2003) which calculates travel costs for each fifteen minute interval including travel times, distance costs and charging costs. Iteration is performed between the two parts to reach a general equilibrium. The combined model has been calibrated using reverse engineering, such that it produces travel demand for the No-toll scenario that are in accordance with traffic flow measurements made before congestion charging was introduced in Stockholm (Kristoffersson and Engelson 2008).

On the demand side, the SILVESTER model is divided into three trip purposes: 1) commuting trips with fixed working hours and school trips (short: fixed), 2) business trips (business) and 3) commuting trips with flexible working hours and other trips, where “other trips” includes e.g. shopping and leisure trips (flexible). Trips with these different trip purposes are likely to respond in different ways to congestion charging. The evaluations of the Stockholm Trial showed that commuting trips to work or school mainly changed mode to public transport and (to a small extent) route, whereas trips with other purposes showed a variety of different ways to adapt including changing destination, trip chaining and cancelling...
the trip (Eliasson et al. 2009). During the simulated period (6.30-9.30 am) the dominating trip purpose is commuting to work or school. The most important user responses are therefore included in SILVESTER, which models mode, route and departure time choice. Saleh and Farrell (2005) stress the importance of work schedule flexibility for citizen’s possibility to respond to congestion charging through retiming of their trips. This supports the segmentation made in SILVESTER where commuting trips are divided into different trip purposes depending on work schedule flexibility.

The main idea behind the departure time and mode switch model builds on the tradition of Small (1982) that it calculates the cost to the user for changing departure time as a schedule delay cost which increases with deviation from the preferred departure time. The choice of departure time then becomes a trade-off between the schedule delay cost and the generalized travel costs in the different time intervals.

The departure time and mode switch model used in SILVESTER was estimated based on SP and RP data for Stockholm (Börjesson 2008). SP data was collected using road side number plate registration followed by telephone interviews. RP data regarding travel times for the same individuals as in the SP study was simulated using CONTRAM calibrated for Stockholm. There is heterogeneity in the value of time (VOT) and the value of schedule delay (VSD) in the demand model population. Van den Berg and Verhoef (2011) stress the importance of including heterogeneity in drivers VOT and VSD when modelling effects of congestion charging. Correlation between departure time intervals is taken into account by allowing for correlation between the schedule delay early and schedule delay late parameter distributions. The schedule delay parameters have been adjusted in calibration since validation of SILVESTER showed that the model overestimated number of trips changing to a time interval before 6:30 am, which is not subject to congestion charging (Kristoffersson and Engelson 2011; Kristoffersson 2013). In general, the departure time choice effects of congestion charging have been surprisingly small in Stockholm (Eliasson et al. 2009).

Route choice is in SILVESTER performed by the mesoscopic model CONTRAM. Adding congestion charging to the network makes route choice highly dependent on car users’ VOT. In order to save calculation time, route choice is not performed for each simulated VOT, but for four user classes in CONTRAM. Demand is thus divided into four user classes depending on VOT: 1) low VOT (<43 SEK/h), 2) medium VOT (43-200 SEK/h), 3) high VOT (>200 SEK/h) and 4) vehicles that are exempted from charging. For each trip purpose the percent of vehicles in each class depends on the VOT distribution of that trip purpose, except that a fixed percentage (28%) of the vehicles are modelled to be exempted independent of trip purpose.

CONTRAM is a semi-dynamic model, i.e. time is divided into time periods of, in our case, fifteen minutes. Network effects are modelled in the sense that upstream links and crossings are blocked by spill-back from downstream links. Queues are
however built vertical at the downstream crossing and block the upstream link only when queue length exceed link length. Shockwaves are thus not modelled.

### 3.2 Congestion Charging Scenarios

Four scenarios are compared in this paper:

1. **No charge** – The scenario without congestion charging.
2. **Simple cordon** – The cordon introduced in Stockholm in 2006, which constitutes of charging stations on bridges around the inner city (Figure 1). The peak hour charge is 20 SEK, both in inbound and outbound direction. The charges build up in steps before the peak and decrease in steps afterwards. The bypass motorway Essingeleden is free of charge.
3. **Differentiated cordon** – In this scenario the simple cordon locations are fixed, but amounts have been differentiated depending on charging station. The amounts have been optimized in Ekström et al. (2009) for the morning peak hour using static assignment on a crude network model of Stockholm and a simplistic mode choice model. Compared to the simple cordon, charges are generally higher inbound and lower outbound in this differentiated scenario. The inbound peak hour charge vary from 18 to 43 SEK per crossing and the outbound peak hour charge vary from 8 to 22 SEK per crossing, depending on charging station. The bypass motorway Essingeleden is still free of charge.
4. **Four cordons** – In this scenario both locations and amounts of the charge are relaxed. This charging scheme was found through optimization of charging locations and amounts with the objective to maximize net social surplus (social surplus minus collection costs) as described in Ekström et al. (2009). In the original solution to the location problem 69 links in the static network where charged. In order to translate this pricing scheme to the CONTRAM network four cordon lines where identified (Figure XX). For the cordon line north of the city centre the morning peak hour charge is 25 SEK per crossing and only southbound traffic is charged (towards CBD). The morning peak hour charge on the other three cordon lines is 15 SEK per crossing charged only for northbound traffic. Note that the bypass motorway Essingeleden is subject to charging in this scenario.
Figure 2: The CONTRAM Stockholm network and charging locations for the simple and differentiated cordon (violet o-marks) and for the four cordons scenario (blue cordon lines and red x-marks).

The scenarios compared in this paper all have the same time profile: the complete charge at peak hour 7.30-8.30 am, 75% of peak hour charge at 7.00-7.30 and 8.30-9.00 am, and 50% of peak hour charge at 6.30-7.00 and 9.00-9.30 am. Results of SILVESTER simulations for congestion charging schemes that have different time profiles are given in Kristoffersson (2013).

3.3 Equity Analysis across Socio-Economic Groups

In this paper, equity effects of congestion charging are evaluated using a welfare-based approach. In welfare-based analysis the benefits of travel time gains, losses through behavioural adjustments and paid charges etc. need to be translated into the same unit for comparison. This unit is usually taken to be money. The question then arise how conversion into monetary units should be made in evaluation. The conventional welfare economic approach is to apply the Kaldor-Hicks criterion of efficiency (Kaldor 1939). The Kaldor-Hicks criterion is a requirement that benefits should be enough to be able to compensate those who have lost. An actual compensation is however not required. The underlying assumption is that all individuals have equal utility of money. In transport economics this translates into conversion of travel time gains into monetary units via multiplication by individual VOT’s. Harberger (1971) argues for the use of this method, as does Sugden (1999) when he suggests that individual willingness to pay estimated in behavioural
studies should be carried on unchanged to evaluation. The main argument here is that the economic analysis should consider only whether the policy increases aggregate welfare, and that compensation to some groups for their losses is a political question.

The method has however been criticized for example by Galvez and Jara Diaz (1998) and by Nash et al (1975). The critics argue that individual VOT's should not be used in evaluation since correlation exists between VOT and income. Variability in VOT can arise either because time is more important for some trips than others - e.g. for the same individual peak hour work trips usually have a higher VOT than weekend shopping trips - or because income and expense burden differ among individuals. Critics argue that the first source of variability in VOT should be included in the evaluation but not the second. Conversion into monetary units should therefore be made after summation of individual costs and benefits, using an average marginal utility of cost (Mackie, Jara-Diaz, and Fowkes 2001).

The equity analysis in this paper uses a combination of geographic distribution of socio-economic groups and geographic distribution of gains and losses due to introduction of congestion charging. Basis for calculations is the average consumer surplus per trip in each origin zone. The change in consumer surplus ($\Delta CS$) due to introduction of congestion charging is calculated as the difference between the the logsums ($L$) computed with and without the charging scheme. The logsum for trip purpose $k$ and origin-destination pair $\omega$ is given by:

$$L_{k\omega} = \frac{1}{50} \sum_{d=1}^{50} \left[ \ln \left( \sum_{t} e^{V_{c\omega}^{tdk\omega} + V_{p\omega}^{k\omega}} \right) \right],$$  

where $V_{c\omega}^{tdk\omega}$ is the observed part of the utility for traveler (draw) $d$ travelling by car at departure time $t$ with trip purpose $k$ and origin-destination pair $\omega$, $V_{p\omega}^{k\omega}$ is the observed part of the utility function for public transport for trip purpose $k$ and origin-destination pair $\omega$, and $\beta^{kd}$ is the cost parameter for draw $d$ and trip purpose $k$.

The logsum measure includes the positive effects of congestion charging, i.e. the time gains, as well as the cost of the charge itself and adjustment costs for the users that change departure time, mode or route when congestion charging is introduced. Since the logsum is used, this measure of consumer surplus is consistent with the demand model and no exogenous VOT is introduced, as is mostly the case when using the "rule-of-a-half" (De Jong et al. 2007). One limitation of our measure is that consumer surplus should ideally be calculated on a 24 hour basis, whereas our simulation model only allows us to calculate consumer surplus for the morning from 6.30 to 9.30.
As can be seen from Equation (1), conversion to monetary units is performed before averaging over draws in the mixed logit model, which means that the assumption on equal utility of money underlying the Kaldor-Hicks criterion is applied in this paper to the equity analysis. In other words, the trips have different VOT not due to income differences but rather due to different utility associated with travel time savings. Income did not differ a lot in the population on which the demand model was estimated (car users driving from western Stockholm to the inner city on a weekday morning) and the income parameter was not significant in estimation. We thus assume that most of the variability in VOT in our model is not because of an income effect.

The socio-economic groups studied in this paper are income groups and household types. Welfare effects are also studied across trip purposes and geographical areas. Figure 3 shows the division of Stockholm into north, south and inner city, which are the aggregate areas studied and which will also play a role in explaining the patterns of congestion in the city.

![Figure 3: Geographical areas of Stockholm: North (N), Inner city (I) and South (S).](image)

It was not possible to consider gender since welfare effects are evaluated on a geographical basis. The income grouping has three categories: 1) low income (less than 195 kSEK/year), 2) medium income (195-375 kSEK/year) and 3) high income (more than 375 kSEK/year). Income values correspond to income before taxes. There are four household types: 1) single adult, 2) two or more adults, 3) one adult with child/ren and 4) two or more adults with child/ren. Data on population per household type was not directly available, but had to be inferred from data on population per age group in each geographical zone. Equation (1) shows how the change in consumer surplus ($\Delta CS$) due to introduction of congestion charging is calculated for user group $g$ and trip purpose $k$. 

\[ \Delta CS_{gk} = \frac{\sum_z \Delta CS_{zk} \Omega_{zg} N_g}{\sum_z \Omega_{zg}} \]  

Here \( \Delta CS_{zk} \) is average change in consumer surplus per trip for zone \( z \) and trip purpose \( k \) when comparing the situations with and without charge, \( \Omega_{zg} \) is number of inhabitants in zone \( z \) that belong to group \( g \) and \( N_g \) is number of car trips produced during an average workday morning by an individual in group \( g \) in the scenario without charges. The population in each zone segmented on income group is obtained from SCB (Statistics Sweden).

Car trip production per person (\( N_g \)) is calculated from the latest travel behaviour survey for Stockholm (Trivector Traffic AB 2006). As one would expect, the production of car trips per person is largest in the high income group and lowest in the low income group. When comparing household types, households with two or more adults and child/ren make most and single adult households least car trips per person and morning.

The CONTRAM Stockholm network covers more than the municipality of Stockholm but somewhat less than the whole county (Figure 2). Traffic flows originating outside of the network area are modelled to start at so called gate zones at the border of the network. Trips starting at these gate zones are not included in calculations of total consumer surplus or revenues. This is because it is not possible to match a distribution of inhabitants on income groups and household types for these zones, since traffic entering at the gate zones come from several zones with different characteristics. Only trips originating inside the network area are thus analyzed regarding their consumer surplus, revenue accumulation and distributional impact. Traffic entering the modelled area from the gate zones serve the purpose of contributing to congestion and have the same possibilities to react to the charges.

4 RESULTS

4.1 Traffic Effects

Table 1 shows a summary of the traffic effects for the simulated scenarios. Traffic flow over the cordon decreases in the simple cordon scenario with 19% in the morning (6.30-9.30 am). This is similar to measurements from the Stockholm Trial that showed a reduction in traffic flow over the cordon of approximately -18% during the morning (Eliasson et al. 2009).
Table 1: Overall Network Results for the Different Scenarios

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>1) No charge</th>
<th>2) Simple cordon</th>
<th>3) Differentiated cordon</th>
<th>4) Four cordons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compared to No charge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average traffic flow on cordon (veh/h)</td>
<td>38224</td>
<td>-19%</td>
<td>-24%</td>
<td>-12%</td>
</tr>
<tr>
<td>Average network speed (km/h)</td>
<td>39.1</td>
<td>+5%</td>
<td>+7%</td>
<td>+10%</td>
</tr>
<tr>
<td>Average speed on cordon (km/h)</td>
<td>40.9</td>
<td>+11%</td>
<td>+16%</td>
<td>+10%</td>
</tr>
<tr>
<td>Total distance travelled in network (veh-km)</td>
<td>3744584</td>
<td>-4%</td>
<td>-6%</td>
<td>-7%</td>
</tr>
<tr>
<td>Total queuing time in network (veh-h)</td>
<td>18970</td>
<td>-21%</td>
<td>-29%</td>
<td>-36%</td>
</tr>
</tbody>
</table>

Average network speed and especially the speed on cordon links increase when the network is subject to congestion charging, as expected. The total distance travelled in the network decreases, since some trips switch to public transport and some trips change departure time to a starting time outside of the modelled morning period (6.30-9.30 am). The route choice counteracts this decrease in total distance travelled somewhat, since the simple and differentiated cordon scenarios encourages car users to travel a longer route around the inner city on the motorway Essingeleden which is free of charge in these scenarios. Total queuing time decreases a lot in the simulations - from -21% for the simple cordon to -36% for four cordons.

4.2 Welfare Effects and Equity Analysis
Table 2 shows that both alternative charging scenarios increase total welfare compared to the *simple cordon* scenario. The aim of the alternative scenarios to relate the congestion charge more to congestion level have worked and efficiency (socio-economic benefit) has increased. Social surplus is 20% larger in the *differentiated cordon* scenario compared to the *simple cordon* scenario. In the *four cordons* scenario, the increase in social surplus is even greater with 82% compared to the *simple cordon* scenario.
Table 2: Overall Welfare Effects

<table>
<thead>
<tr>
<th>MSEK/year</th>
<th>Simple cordon</th>
<th>Differentiated cordon</th>
<th>Four cordons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Δ CS</td>
<td>-1</td>
<td>24</td>
<td>263</td>
</tr>
<tr>
<td>Total revenues paid</td>
<td>428</td>
<td>487</td>
<td>514</td>
</tr>
<tr>
<td>Total Δ SS</td>
<td>427 (ref)</td>
<td>511 (+20%)</td>
<td>777 (+82%)</td>
</tr>
</tbody>
</table>

Note that the results of this paper shows welfare effects of introducing congestion charging that are considerably larger than previous analyses performed with static assignment and a single value of time for each trip purpose. In two of the scenarios of this paper – the differentiated cordon ring and the four cordons – drivers benefit even without a refund of revenues. The large benefits of congestion charging in the dynamic setting with heterogeneous users is investigated in Börjesson and Kristoffersson (2014), where it is found that queue spill-back and heterogeneity in value of time increase substantially the predicted benefits of congestion charging.

From Table 3 one can see that the change in consumer surplus for the simple cordon is least negative for low income car users and most negative for high income users. For the scenario with a differentiated cordon the benefits are highest for low income users and smallest for high income users. A sensitivity analysis showed that the distribution of consumer surplus across income groups does not change if car trip production rates are set to equal for all income groups. That high income users make more car trips thus amplifies a pattern caused by differences in where trips are undertaken and how much users in different groups pay in charges. The results thus suggest that both pricing scenarios simple and differentiated cordon are progressive.

Table 3: Welfare Effects across Income Groups

<table>
<thead>
<tr>
<th>SEK/person/year</th>
<th>Simple cordon</th>
<th>Differentiated cordon</th>
<th>Four cordons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Δ CS Low income users</td>
<td>-4</td>
<td>60</td>
<td>460</td>
</tr>
<tr>
<td>Average Δ CS Medium income users</td>
<td>-40</td>
<td>50</td>
<td>560</td>
</tr>
<tr>
<td>Average Δ CS High income users</td>
<td>-110</td>
<td>20</td>
<td>690</td>
</tr>
</tbody>
</table>

In the fourth congestion charging scenario four cordons, results are however the opposite: The change in consumer surplus is most positive for high income users and least positive for low income users. The simulation results of this scheme are thus regressive. However, from an efficiency point of view this scenario is the best, as can be seen from
Table 2. The trade-off between efficiency and equity thus shows up here: When relating the charge more to congestion level we end up with undesired distributional patterns. In this case study the regressivity arises because of the high peak hour charge of 45 SEK for driving from the very south of Stockholm to CBD (passing three of the four cordons), in combination with incomes in general being higher in the inner city than in the outer suburbs, and incomes in general also being higher north compared to south of Stockholm.

Results for different household types are rather similar for the different scenarios (Table 4). Households with two adults and child/ren benefit most in both the differentiated cordon and the four cordons scenario. In the simple cordon scenario households with one adult and child/ren benefit most.

Table 4: Welfare Effects across Household Types

<table>
<thead>
<tr>
<th>SEK/person/year</th>
<th>Simple cordon</th>
<th>Differentiated cordon</th>
<th>Four cordons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Δ CS Single adult</td>
<td>-20</td>
<td>30</td>
<td>360</td>
</tr>
<tr>
<td>Average Δ CS 2+ adults</td>
<td>-20</td>
<td>40</td>
<td>500</td>
</tr>
<tr>
<td>Average Δ CS Single adult + child/ren</td>
<td>40</td>
<td>80</td>
<td>400</td>
</tr>
<tr>
<td>Average Δ CS 2+ adults + child/ren</td>
<td>20</td>
<td>100</td>
<td>690</td>
</tr>
</tbody>
</table>

When looking at simulation results for different geographical areas, large differences become apparent (Tables 6-8). These geographical differences are input to the effects for different socio-economic groups shown in Table 4-5, since these results are calculated as change in consumer surplus per geographic zone combined with share of population in the zone that belong to each socio-economic group. The simple cordon simulation is least beneficial for trips originating in the inner city and ending north or south of Stockholm (Table 5). The same relations suffer in the differentiated cordon together with trips originating in the northern suburbs and ending in the inner city (Table 6). Both cordon scenarios are most beneficial for trips inside the cordon and within the northern suburbs, which is natural considering that these trips are not charged but can still gain from less congestion in the network. Overall results for the two cordon ring scenarios have the same sign for all origin-destination-relations. The geographic result pattern is different in the four cordons scenario (Table 7). Here all trips with origin in the inner city gain (out of which two were the big losers in the cordon scenarios). Furthermore trips within the southern suburbs loose. These geographic differences are the likely explanations for effects on different income groups.
An apprehension was raised in the introduction that the differentiated cordon would be regressive since trips from suburbs to CBD would pay more because of more congestion in this driving direction in the morning. The result does however not support this apprehension. The geographic pattern of gains and losses does not change sign between the simple and differentiated cordon ring scenarios. There seem to be less socio-economic differences between drivers going to and from CBD, compared to the differences between drivers going to/from CBD and drivers travelling within the suburbs. This analysis is also supported by the regressive impacts of the four cordons scenario which is negative for trips within the southern suburbs. Thus, in our case the location of the cordoning stations seems to be more decisive for equity effects than the level of the charge in different driving directions.

Considering different trip purposes, simulation results indicate that business trips benefits most from congestion charging whereas fixed trips benefit least, with flexible trips in between (Table 8). This is what one would assume considering that business trips usually has a high VOT and flexible trips often have the possibility to adjust to avoid charges.
4.3 Equity Analysis with Revenue Strategies

Figure 4 and Figure 5 show the net benefits of introducing congestion charging when return of revenues are included. The net benefit is expressed as the change in SEK per person and year in each socio-economic group between the situations with and without congestion charging. Four strategies for return of revenues are compared: To give an equal amount back to all adults (lump sum), to give an amount back which is proportional to number of car trips made (reduced costs for car trips), to give an amount back which is proportional to number of public transport trips made (reduced costs for public transport trips) and to give an amount back which is proportional to income (tax cut).

Figure 4: Revenue strategies and effects for different income groups.

Figure 4 shows that the tax cut refund strategy benefits the high income group much more than the other refund strategies. The other three strategies show only minor differences in the outcome for both income groups and household types. One exemption is households with two adult and child/ren that benefit less from the reduced costs for public transport trips than from the other strategies (Figure 5). This could be expected because these households have relatively high car usage.
A study on equity effects of congestion charging was made before charges were actually implemented in Stockholm (Eliasson and Mattsson 2006). The charging system analyzed in that paper is the cordon ring of this paper combined with a cordon line on the central bridges in Stockholm (similar to the central cordon line in the four cordons scenario). Just as for the two cordon ring scenarios of this paper, Eliasson and Mattsson (2006) also find that the charging system is progressive. A difference compared to this paper is that Eliasson and Mattsson (2006) find a greater impact of refund strategy choice. It is possible that this could depend on the more detailed model of socio-economic characteristics of individuals in their paper. In this paper we had to be satisfied with return of revenues proportional to average travel characteristics in the group.

5 CONCLUSIONS

The aim of this paper is to investigate further a question related to equity of urban congestion charging through a case study for Stockholm:

Is congestion charging progressive or regressive?

We find that whether congestion charging is progressive or regressive depends on the structure of the city in question, where charging stations are placed and how revenues are spent.

For Stockholm we find that the most efficient charging scheme (four cordons) changes the charging system from progressive to regressive compared to a cordon around the inner city. The four cordons scenario is the most efficient scenario in the paper, but the least equitable. As measure of efficiency we use revenues plus change in consumer surplus and both the increase in consumer surplus and revenues are larger in the four cordons scenario than in the other
two scenarios. The four cordons scenario thus best manages to apply a charge to trips that contribute to congestion. However, high-income travelers gains most from this scenario and low-income travelers least. The four cordons scenario is thus regressive.

The simple and differentiated cordon scenarios are both progressive, whereas the four cordons scenario is regressive. The difference between the two toll ring scenarios on one hand and the four cordons scenario on the other hand is the location of the charge. Whether congestion charging in Stockholm is progressive or regressive thus depends on where charges are levied. This is because of socio-economic differences in where people live and work, and geographic differences in accessibility to public transport.

It is not so surprising that the simple cordon scenario turns out to be progressive, given that accessibility to public transport is very high for trips to and from Stockholm CBD during peak hours and that income is high among people living inside the cordon ring who have to pay for driving out of the inner city. Somewhat more surprising is that also the differentiated cordon shows progressive effects. In this scenario it is more expensive to drive into the CBD in the morning than to drive out of the CBD, because the charge is related to congestion level. On beforehand an apprehension was raised that this could lead to a regressive scheme, since it would benefit high-income travelers living in the inner city. However, the results do not support this apprehension. The reasons for the progressive effects of the differentiated cordon ring are that users travelling by car into Stockholm CBD during charging hours were already before introduction of charges users with an income above average and, again, the high accessibility to public transport for trips to and from Stockholm CBD during peak hours. The Stockholm congestion charge has a very simple design (the main variation is by time-of-day). The scheme could be more efficient if the charged amount was adjusted more to congestion level. The results of this paper support a scheme design in Stockholm where charges are allowed to vary by driving direction.

As noted above, the four cordons is the most efficient scenario in this paper, but it is regressive since high income groups gain most from this scenario and low income groups least. The reason is the structure of Stockholm with an overweight of residential areas on the south side and most workplaces in the richer north side of the city (including the northern parts of the inner city). This creates congestion mainly in the northbound direction in the morning and in the southbound direction in the afternoon. Cities with similar structures with relatively poor residential areas on one side of the city and workplaces on another side will most likely face a similar situation where the most efficient charging system is regressive. This is because most congestion occurs from the poorer side to the richer side in the morning and in the other direction in the afternoon. The four cordons scenario implies that a higher charge is payed the further south from the city centre the trip starts, with a maximum peak hour charge of 45 SEK for drivers starting in the outer southern suburbs. Southern suburbs close to the inner city in Stockholm are in general rich areas, but also with good access to public transport. Outer southern suburbs are poorer and
have to pay the most. Furthermore, to avoid adverse route choice effects, charges are in the *four cordons* scenario not only levied on the congested arterials leading into the inner city, but also on smaller roads within residential areas in the southern suburbs (see Figure 2). This leads to large negative impacts for traffic travelling within the southern suburbs.

Ideally, in order to accurately investigate correlation between benefit from congestion charging and the socioeconomic variables the change in consumer surplus should be calculated per individual. However the model we use is based on traffic analysis zones and the income and other socioeconomic variables are only available as zone averages. Still it is possible to observe correlation between the income and the change in consumer surplus and make reasonable conclusions consistent with economic intuition and local knowledge. We believe that the conclusions would not be considerably different if the analysis would be carried out on the individual level.

To summarize, the paper shows that also in a real-world network welfare rises when charging for congestion, but that a separate equity analysis should be performed, since the most efficient design may be undesirable from an equity point of view. Furthermore, congestion charging cannot be said to be progressive or regressive per se, rather it varies between cities and even between different scheme designs for the same city.

REFERENCES


