The Stockholm congestion charges: an overview

Jonas Eliasson, KTH Royal Institute of Technology


Abstract
Congestion charges were introduced in Stockholm in 2006 as a seven-month trial, followed by a referendum where a majority voted in favour of the charges. This led to the reintroduction of congestion charges in August 2007, and they have been operational since then. The system has attracted worldwide attention worldwide, both because it achieved substantial congestion reductions, and because the system overcame fierce initial hostility, surviving a heated and complicated political and legal process, and eventually gaining support by more than 2/3 of the population and all political parties. This report summarises the story of the Stockholm congestion charges, pointing out experiences and lessons learnt.

Note: Parts of the material in this report have been published before in other publications by the author and various coauthors. The intention of this report is to make the material accessible in one single, coherent source.

Keywords: Congestion charges, Stockholm, sustainable transport.

JEL Codes: H23, H54, R41, R48.
1 INTRODUCTION

Congestion charges were introduced in Stockholm in 2006 as a seven-month trial, followed by a referendum where a majority voted in favour of the charges. This led to the reintroduction of congestion charges in August 2007, and they have been operational since then. The charging system consists of a cordon around the inner city, with a time-differentiated toll being charged in each direction. Traffic across the cordon was reduced by around 20%, leading to substantial congestion reductions in and around the city.

Congestion pricing has been advocated by transport economists and traffic planners for a long time as an efficient means to reduce road congestion. Despite growing problems with urban congestion and urban air quality, and despite a consensus that investments in roads or public transit will not be sufficient to tackle these problems, cities have been reluctant to introduce congestion pricing. In recent years, however, it seems that this is changing. London (2003), Stockholm (2006), Durham (2002), Milano (2008), Rome (2001), Gothenburg (2013) and Valletta (2007) have all introduced different forms of charging or permit systems to combat congestion and/or environmental problems, and many other cities are considering it. New York, Manchester, Copenhagen and Edinburgh have all recently tried to introduce congestion charges, and even if these attempts have been unsuccessful, it is a sign that congestion charges are being seriously considered to a greater extent than a decade ago. The soon ubiquitous “value pricing” roads in the US are also examples of how congestion problems are now being tackled through pricing measures.

The congestion charges in Stockholm have attracted enormous attention worldwide. Obviously, the opportunity to gauge the effects of congestion charges on traffic, congestion levels and travel behaviour has attracted great interest. But perhaps even more interesting is that the congestion charges overcame fierce initial hostility, survived a heated and complicated political and legal process, including a referendum initially forced through by opponents to the charges, and has eventually gained support by more than 2/3 of the population. The Stockholm charges went from “the most expensive way ever devised to commit political suicide” (to quote the then-secret feelings expressed by the Head of the Congestion Charging Office) to something that the initially hostile media eventually declared to be a “success story” (e.g. Dagens Nyheter, June 22, 2006).

This report tries to summarise the lessons that can be learnt from the Stockholm experiences.

---

1 Legally, the congestion charge is a tax, according to Swedish law, and the official Swedish term is hence “congestion tax”. We have chosen to use the international standard term “congestion charge”.

2 Quote Gunnar Söderholm, social-democratic head of the Congestion Charging Office during the trial, when (after the trial) describing the local Social Democrats’ feelings when the national Social Democratic government more or less forced the congestion charges onto the local Stockholm party district.

3 The report draws heavily on previously published papers by the author and coauthors. Some parts are taken verbatim, in particular from (Börjesson, Eliasson, Hugosson, & Brundell-Freij, 2012; Eliasson, 2008, 2010, 2014). Further references are given in the report. The intention is not to present new material, but merely to collect previously published works in a single easily accessible place.
1.1 Background: some basic insights about congestion pricing

It is well established, both theoretically and empirically, that infrastructure investments are not sufficient to eliminate road congestion in the cores of large cities. There are several reasons for this: two of the most important are the eventually inevitable scarcity of urban land and public resources.

Congestion charging will of course not solve everything. Introducing congestion charges will usually reduce the need for transport investments, but generally speaking not eliminate it. Normally, a growing urban region will need both congestion charging and transport investments, perhaps both roads and public transport. Obviously, cities are different as to what investments are the most cost-efficient and the most needed. Generally speaking again, a sustainable urban transport system must incorporate four strategies: attractive public transport, walkability, compact spatial planning, and restraints on car traffic. All these four will strengthen each other, and without one of them, the remaining three will lose effectiveness. The purpose of this paper is to present evidence and lessons about congestion charging – which is one way to restrain car traffic; parking pricing is another example – but the other three need to be included in a comprehensive sustainable urban transport strategy.

It is probably worth to point out that it is also well established that car drivers are in fact sensitive to costs. It is a common misconception among laymen, and sometimes among decision makers, that drivers “have to” drive, and do not react to changes in the driving costs. This notion has been refuted numerous times in many kinds of contexts. Increasing the cost to drive at certain places at certain times will decrease the number of drivers choosing to drive there and then. How large the decrease becomes depends on the ease of adaptation, among other things – in other words, how good alternatives there are. Alternatives may be other time periods, modes, routes, destinations etc. It is imperative to keep as many options open as possible to achieve good traffic reduction effects – but it is up to the drivers themselves to choose how to adapt.

2 THE STOCKHOLM STORY IN BRIEF

2.1 Basic facts about Stockholm

The City of Stockholm has around 0.9 million inhabitants, and is the central part of the Stockholm county, with a total of 2 million inhabitants. Around 2/3 of the City inhabitants live in the inner city – that is, within the toll cordon - and the rest outside. The area of the toll zone is around 35 km². The zone has around 330 000 inhabitants, of which approximately 60 000 commute to workplaces outside of the zone. The zone has close to 23 000 workplaces, employing approximately 318 000 persons, of which more than two thirds are commuting from outside the zone.

The population of Stockholm has been growing rapidly for many years, and together with increasing trip lengths and car ownership, this has led to steadily increasing traffic volumes. Traffic volumes across the cordon used to increase at the same pace as the traffic in the county as a whole from the early 1970's (when regular measurements started) up until the early 1990's, when traffic across the cordon stopped growing. Traffic in the rest of the county, however, continued growing at the same pace, as did the number of transit trips across the cordon. The most likely explanation of this sudden end to traffic growth is simply a lack of road capacity. Traffic across the cordon

---

4 As a comparison, the congestion charge scheme of London introduced in 2003 encompassed a 21 km² zone, this was almost doubled by the western extension in 2007.
then remained surprisingly stable for the next 15 years or so, despite significant changes in employment levels, fuel prices etc., with the only appreciable effect being a traffic decrease of around 5% when the Southern Bypass opened in 2004.

Because of its topology, with lots of water and well-preserved green wedges, road congestion levels in Stockholm are high compared to the city’s moderate size. Before the introduction of the congestion charges, the main roads arterials leading to, from and within the city centre had congestion indices typically averaging around 200%, i.e. three times the free-flow travel time. Partly because of this, and partly because of good public transport supply, the transit share is high: 60-65% of all motorized person trips to and from the city centre are made by transit. During rush hours, the share increases to 80%. The public transport system in the county of Stockholm consists of a subway network with 100 stations and over a million trips per day, a commuter rail network with 51 stations and nearly a quarter of a million trips per day, five light rail lines with 98 stations with a bit more than 100 000 trips per day, and an extensive bus network with nearly a million trips per day. Public transport fares are subsidized at a rate of around 50% of actual costs.

2.2 The story of the charges – an overview

Just like in many other cities, transport planners and economists had suggested that Stockholm should introduce congestion pricing for a long time, without getting either public or political support. In the early 1990’s, road tolls were proposed as a way to partially finance a large infrastructure package for Stockholm. This ignited the interest from environmentalists, who appreciated the traffic management potential of the tolls, even if they didn’t approve of that the revenues were planned to be partially used for new motorways. The infrastructure package agreement broke down in the late 1990’s, and the tolls were never introduced – but the ball had been set rolling. Several stakeholders carried out analyses of congestion charging schemes, and perhaps more important, the issue had entered the agenda of the environmental movement, in particular the Green party.

In 2002, the social-democratic national government set up a commission to negotiate a new infrastructure agreement for Stockholm. The idea was floated to use road pricing as a funding source. When the Conservative party accused the social-democrats of having secret plans to introduce “road tolls” after the election 2002, the social-democratic mayor in Stockholm promised very clearly and publicly that there would be no road tolls in Stockholm during the next election cycle (although she wanted to prepare a suggestion in time for the next election). The social-democrats went on to win both the national and the Stockholm election, provided that they could ensure support from the Green party. In return for support for a social-democratic national government, the Green party demanded that a “several-year, full-scale congestion charging trial” should be carried out in Stockholm. The social-democrats obliged.

This led to an extremely heated debate. Congestion pricing was an unpopular measure from the outset, and the broken election promise made matters worse. The opposition raged, while silently celebrating what they anticipated to be a landslide victory in the next election. Both proponents and opponents of the charges used dramatic rhetoric to describe what would happen with or without congestion charges, respectively. Even many of those in favour of congestion charges were sceptical: the way the trial was introduced, and the short time available for preparation, made them fear that a failed attempt at introducing congestion charges would block the question for many years ahead. The media picture was overwhelmingly negative: 39% of all newspaper articles on the topic were negative, compared to 3% positive (the rest were neutral) (Winslott-
The Stockholm congestion charges

Hiselius, Brundell-Freij, Vagland, & Byström, 2009). Opponents to the charges suggested a referendum about the charges, confident that they would win. The idea was silently welcomed by the social-democrats, who saw it as a way to put some distance between them and the charges: with a separate referendum, it would be possible to vote for the social-democrats and still vote no to the charges. However, it was decided that the referendum should not be held until after the trial, in conjunction with the next regular election in September 2006. This turned out to be of crucial importance.

The trial had the purpose to “test whether the efficiency of the traffic system could be enhanced by congestion charges”. The toll was expected to “reduce congestion, increase accessibility and improve the environment”, both in terms of emissions from car traffic and the perceived urban environment. The toll rate was set so as to reach the target of reducing car traffic across the cordon with 10-15 percent, a target loosely based on previously suggested road pricing schemes for Stockholm.

The congestion charging trial started in January 2006, when a time-differentiated cordon toll around the inner city was introduced. Traffic across the cordon dropped immediately, leading to dramatic congestion reductions all over the city. After a few weeks, the decrease in traffic volumes across the cordon during the charged period stabilized around 22% compared to 2005 levels, resulting in congestion reductions around 30-50% (Eliasson, Hultkrantz, Nerhagen, & Rosqvist, 2009; Eliasson, 2008). Public attitudes gradually became more positive, while the media picture changed completely: the share of positive newspaper articles increased from 3% to 42% while the share of negative articles fell from 39% to 22% (Winslott-Hiselius et al., 2009). In the referendum in September, 53% of valid votes were in favour of keeping the charges.

Representatives for all political parties in Stockholm had promised to follow the outcome of the Stockholm referendum. The election ended up with liberal/conservative majorities both nationally and in Stockholm, and the new majority in Stockholm obligingly asked the new national majority to reintroduce the congestion charges, which had been turned off before the referendum. The crux was the negotiation about the revenues. Legally, the charge was a national tax that ended up in the national government’s coffers, but the Stockholm region understandably argued that it was really their money. Eventually, the regional and national politicians brokered a huge, ten-year infrastructure package worth around 10 billion euros, where one part of the deal was that the charge revenues were earmarked for a new bypass around Stockholm.

The trial hence turned out to be a milestone in the development of urban road pricing. First, to the surprise of all but a few hard-headed road-toll enthusiasts, it finally tipped the balance of a forty-year political consideration of road-tolls in Stockholm by invoking more or less a land-slide change of the opinion of the general public in favour of tolls. Second, it was the third full-scale demonstration of an urban congestion charge,

---

5 Some surrounding municipalities also arranged referenda, although they had no legal influence, since the congestion charges were entirely within the border of the city of Stockholm. (Around 2/3 of the city’s population live inside the cordon. It should be noted, by the way, that residents within the cordon pay on average much more in charges than residents outside the cordon, while they get less of the travel time benefits, since congestion problems mainly exist into the city in the morning and out from the city in the afternoon). Counting the votes of all referenda that were held, there was a majority against the charges, but the selection was heavily skewed: referenda were only held in municipalities with liberal/conservative majorities and where opinion polls showed that there was a majority against the charges.
after Singapore and London, and the second to be based on a time-differentiated scheme, after Singapore.

As time went on, all political parties accepted and, eventually, even embraced the congestion charges. The reasons for this included the congestion reduction, the means to finance infrastructure, the possibility to get leveraged funds from the government, and the steadily increasing public support for the charges. The media interest for the charges faded, after having been in the headlines almost daily for four years. Rather than discussing the existence of the charges, the political parties and other stakeholders gradually moved on to discussing how the charges could be redesigned and how the revenues should be used. The traffic reduction has remained remarkably stable over time (Börjesson, Eliasson, Hugosson, & Brundell-Freij, 2012). At the time of writing (February 2014), the liberal/conservative majorities in the national government, the city of Stockholm and the county of Stockholm – the former opponents of the charges – have agreed to substantially increase the level of the charges and introduce a new toll on the western bypass, with the dual purpose to finance a metro extension and reduce congestion even further. The only objection from the left/green opposition is that it is too little, too slow and too late.

3 THE CHARGING SYSTEM AND ITS EFFECTS

The effects of the Stockholm system were studied in an extensive evaluation program. This evaluation program was particularly important since the fate of the charges would be decided by a popular referendum. The evaluation covered not only effects on traffic volumes and travel times, but also effects on emissions, perceived urban environment, traffic safety, delivery traffic, public transport, taxis etc., generating almost 30 different sub-reports.

3.1 Description of the system

The system consists of 18 charging points located at the main bottlenecks on the arterials leading into and out from the inner city. These 18 points form a cordon around the inner city. The cordon shape is just a consequence of the topology of the city, though, not a design constraint; it is a common misconception that congestion charging systems have to be shaped as cordons (system design is discussed in section 5).

Vehicles are registered automatically by cameras that photograph the number plates; there is no opportunity to pay at the control points. The owner of the car is then sent a monthly invoice for the total charge incurred during a month. During the trial, the main means of identification was transponders (“tag-and-beacons” or DSRC, dedicated short-range communication). When the charges were reintroduced, the automatic camera identification, originally intended only as a secondary means of vehicle identification, worked so well that it was decided to abolish the transponders.
The cost for passing a control point in any direction is 1-2€ (using 10 SEK = 1€) depending on the time of day (see Figure 1), with a maximum amount per vehicle and day of 6€. The cost is the same in both directions, and each passage is charged. No congestion charge is levied during nights, weekends, holidays on in July. Various exemptions (e.g. buses, foreign cars and for traffic between the island of Lidingö and the rest of the county) mean that about 15% of the passages are free of charge.

Originally, there was an exemption for alternative-fuel cars, intended to stimulate the market introduction on such cars. This proved to be an effective measure: the share of alternative-fuel cars increased from 3% in 2006 to 15% in 2009. The sales of alternative-fuel cars was also stimulated in other ways, but several studies have concluded that the congestion charge exemption played an important role. The exemption was abolished for vehicles sold from 2009 and later, with the argument that it had filled its role as a facilitator for market introduction. In 2012, the exemption was abolished for all vehicles. The intention of the exemption was to stimulate more environmentally friendly vehicles, but the decision to equate “more environmentally friendly” with “alternative fuels” received criticism after a while, especially when the positive environmental effects of ethanol cars were questioned. The exemption applied to any car that could be propelled by other means than gasoline or diesel, such as ethanol, biogas or hybrids, regardless of which fuel was actually used (biogas and ethanol can alternatively be driven on ordinary gasoline) and regardless of how much emissions they generated. Subsequent Swedish policies intended to stimulate “green cars” have used other definitions, usually using constraints on fuel consumption per kilometer. The important policy conclusion from the “green car” exemption used in the congestion charging system is not the particular definition of what constituted a “green car” – many would argue that this definition in hindsight proved not well suited to the intention – but that it showed that this kind of exemptions can be a very effective stimulus on the vehicle market.

There is no congestion tax levied on vehicles driving on the Essinge bypass past Stockholm. This is the only free-of-charge passage between the north and south part of the county. The Essinge bypass was heavily congested even before the charges, so from a pure traffic perspective, there was a strong argument for also charging vehicles on the
The Stockholm congestion charges

bypass. The opposition from the surrounding municipalities was so strong, however, that the politicians of the City of Stockholm decided that the bypass should be free of charge.

3.2 Traffic effects

Traffic across the Stockholm cordon had remained largely constant since the early 1990’s, despite growing population, and car ownership. Road traffic had grown in the rest of the region, so the most likely explanation for the lack of traffic growth in the inner city was simply lack of capacity. Figure 2 shows average daily traffic volumes across the cordon (weekdays 6:00-19:00) since 2000. A slight drop is seen in 2005 when the Southern Bypass was opened.

When the charges were introduced in January 2006, they had a substantial effect on car traffic from day one. After a few weeks with an even larger initial effect, the traffic reduction stabilized around 22% across the cordon during the charged period (comparing month-by-month to account for seasonal variation). Before the start of the trial, there was some doubt as to whether the traffic reduction would actually take place, especially since the trial was only seven months long, but effects turned out to be immediate and persistent.

When the trial ended July 31 2006 and the charges were abolished, traffic volumes immediately rebounded to almost the same level as before the charges – but not quite. A residual effect remained even after the charges had been abolished. From August 2006 to August 2007, i.e. between the end of the trial and the reintroduction of the charges, traffic volumes remained 5-10% lower than in 2005. The most likely hypothesis is that some car users developed new travel habits during the trial that persisted even after the charges were abolished.

![Figure 2. Average traffic volumes across the cordon, weekdays 6:00-19:00 excl. July. Blue: no charges. Red: charges. “2006a” is the trial period January-July 2006, and “2006b” is the remainder of 2006.](image)

6The exact size of the residual effect is uncertain, since data from this period are less reliable due to roadworks and technical problems with the measurement equipment.
With the reintroduction in August 2007, traffic levels were again reduced to the same level as during the trial period in 2006. Since then, they have remained roughly constant, despite inflation, economic growth, growing population and an increasing car fleet. Rather than attenuating over time, the effect of the charges seems to have grown over time – otherwise, traffic would have been expected to increase due to various external factors (population growth, inflation etc.) – see Table 1. Controlling for such external factors, Börjesson et al (2012) show that traffic’s elasticity with respect to the charges has apparently increased from -0.70 in 2006 to -0.85 in 2009 and onwards. It may be too early to tell if this is a stable, long-term value, although it seems likely. But the most important conclusion is that there are no signs that the effect of the charges is wearing off, but instead increasing somewhat over time. This is consistent with the observation that there are more adaptation mechanisms available in the long term than in the short term. This result is in correspondence with that of Goodwin et al. (2004), who note that price impacts tend to increase over time as consumers have more options. In September 2012, the exemption for alternative-fuel vehicles was abolished. This led to a further reduction of traffic; before that, 6-8% of traffic across the cordon had been subject to that exemption. The observed reduction indicates that the exempted vehicles had the same price sensitivity as the rest of the traffic (a 20% reduction of these 6-8 percentage points translates to a 1.2-1.6% reduction of total traffic, which is very close to what is observed).

Table 1. Traffic reduction across cordon compared to 2005 traffic levels (charged weekdays 6:00-19:00). Second row: reduction compared to a theoretical counterfactual where external factors are kept constant. (no such calculations are available for 2012-2013).

<table>
<thead>
<tr>
<th>Year</th>
<th>2006a</th>
<th>2007b</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic reduction from charges, compared to 2005</td>
<td>-21.0%</td>
<td>-18.7%</td>
<td>-18.1%</td>
<td>-18.2%</td>
<td>-18.7%</td>
<td>-20.5%</td>
<td>-21.4%</td>
<td>-22.1%</td>
</tr>
<tr>
<td>Traffic reduction adjusted for changes in external factors(^7)</td>
<td>-21.4%</td>
<td>-20.9%</td>
<td>-20.7%</td>
<td>-21.9%</td>
<td>-21.7%</td>
<td>-22.3%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Detailed studies of the charges’ effects get less and less meaningful over time, since the counterfactual becomes increasingly difficult to predict. Most of the detailed analyses of effects presented below are hence based on comparing 2005 with 2006, i.e. the last year without charges with the first year with the charges. Since traffic levels had been so stable for several years leading up to 2005, and have remained so stable from 2006 onwards (with a gap for the period August 2006-August 2007), the essential insights gained from these comparisons still seem relevant and applicable.

In relative terms, the decrease was largest in the afternoon peak period (-23 percent between 16:00-18:00), and somewhat lower in the morning peak period (-18 percent between 7:00-9:00). This indicates that a larger share of discretionary trips is made during the afternoon peak than in the morning and/or that departure times from work

\(^7\) Note that these elasticities are neither comparable with the usual cost elasticity of car traffic, nor with the fuel price elasticity of car traffic. Since fuel costs make up around half the marginal cost of driving, the cost elasticity of car traffic is around twice the fuel price elasticity. The elasticity of traffic across the cordon with respect to the charge is higher, since there are more adaptation mechanisms available, such as changing route, destination or time of travel.

\(^8\) Fuel price, total employment, car ownership, inflation, exempted share of traffic.
are less fixed than arrival times to work. Traffic declined in evenings as well. Hence, the reduction of outbound traffic during evening because of fewer incoming vehicles in the morning outnumbered the increase of evening traffic because of within-mode substitution from travel during day-time to free-of-charge evening. Despite the lower charge during mid-day (9:00-15:30), traffic decreased almost as much during this period (-22 percent). The seemingly high cost sensitivity during this time period is partially explained by the fact that many of the trips crossing the cordon during mid-day pay the higher charge when going in the other direction.

In addition, effects on traffic were seen further out from the toll zone than initially was expected. The number of vehicle kilometers driven in the inner city decreased by around 16 per cent. Outside the inner city, on the outlying approach roads and outlying streets, traffic volumes fell by just over 5 per cent. These effects have also remained roughly constant; traffic volumes have either remained constant in the following years, or followed their long-run trend. Queues were reduced also far from the cordon due to reduced spillback congestion. Hence, the charges do not seem to have generated any severe second-best problems neither in the short nor in the long term perspective. Consequently, unwanted side effects that were anticipated outside the cordon, such as an increase of traffic on circumferential roads at the city’s outskirts, were not found.

3.3 Travel times

Even more dramatic than the reduction in the number of vehicles was the reduced congestion. The newspaper headlines in Figure 3 provide an illustration. Improvements in travel times were tangible and easily perceived by the general public. Travel time improvements also occurred far from the inner city, when spillback queues were substantially reduced.

![Figure 3. The front pages of Metro and Dagens Nyheter after the first day of congestion charging. (The headline in both newspapers reads: “Every fourth car disappeared.”)](image)

Figure 4 show average congestion indices for various types of roads, comparing May 2005 and May 2006. Congestion is measured by a congestion index where 0 percent
The Stockholm congestion charges

corresponds to free-flow travel time, while 100 percent corresponds to twice the free-flow travel time. As can be seen, travel times for vehicle traffic declined substantially inside and close to the inner city. Particularly large declines were seen on arterials, on which delay times (time in excess of free-flow travel time) fell by one-third during the morning peak period and by one-half during the afternoon/evening peak period. This considerably improved reliability of travel times, i.e., travellers could be more certain about the duration of a car trip. The figure also shows the range of travel time variability, measured as the distance between the highest and lowest deciles of travel times (during a given 15-minute period). The highest decile of the travel-time distribution fell to a third or less compared to the pre-trial state for some categories of roads (such as arterials during PM peak).

![Figure 4. Relative increase of travel times for various categories of links. 0 percent corresponds to free-flow travel time. The coloured bars show average travel times while the “error bars” indicate the worst decile and the best decile of the travel time distribution. Measurements were taken from all weekdays for six weeks in April-May. “AM peak” refers to 7.30-9.00, “PM peak” refers to 16.00-18.00.](image)

Figure 5 shows a similar picture, but showing more years and another categorization of streets and roads. The figure shows three periods without charges (April 2005, October 2005 and October 2006) and two periods with charges (April 2005 and October 2007). Generally, congestion is worse in April than in October, due to slightly larger traffic volumes and more pedestrians and bicyclists interacting with the traffic flow. It is apparent that the drop in traffic volumes across the cordon translated to congestion reduction in large area both inside and outside the cordon. In particular, the congestion reduction on the outer arterials should be noted, since these measurement points are situated rather far away from the actual cordon. Moreover, it can be seen that the general congestion level was around the same during the trial (April 2006) as after the reintroduction (Oct. 2007). Congestion levels during the non-charged October periods 2005 and 2006 were about equal, while the congestion level during the non-charged April 2005 was a bit higher.
The reduced congestion also meant that travel time reliability increased. For example, the worst travel time decile was reduced by a factor of 3 or more for arterials during the PM peak. The scattergram below (Figure 6) illustrates the increase in travel time reliability, showing standard deviation for link travel times for 2005 (on the x-axis) and 2006 (on the y-axis). Next to it is a similar scattergram showing the congestion indices for several links. Inspection of the diagrams reveals that most points lie below the 45 degree line, i.e. the situation has improved (if it has changed at all).

3.4 Environmental effects

The reduction in vehicle kilometers travelled meant that emissions from traffic were reduced. The reduction was largest in the inner city, between 10 and 15 per cent (the reduction differed across different types of emissions). Since this is a most densely populated area, this is an important effect from a health point of view. In the inner city, of air-borne pollutants were reduced between 10 and 14 percent. For nitrogen oxides (NOx) the reduction was smaller (8.5 percent), since the extended bus traffic used older buses with higher emission factors. Overall, the results reveal that air quality was improved in many
streets in the inner city. Carbon dioxide emissions from traffic in the whole metropolitan area (the county of Stockholm) decreased by 2-3 per cent.

The use of air pollution modelling allowed for a detailed assessment of the changes in population exposure due to the change in traffic. Since the main effects were seen inside the cordon where the daytime population density is high, this also results in important changes of average population exposure. The estimated reductions of NOx and exhaust particles emissions are refer to the change in concentration at rooftop level in the inner city (so-called urban background). This provides an indication of the average load of the population in this part of Stockholm. International research ascribes reduced mortality due to for example fewer cases of cardiovascular diseases and lung cancer as the most important health benefit. Forsberg et al. (2006) estimate that there will be 20-25 fewer premature deaths per year in Stockholm’s inner city and a total of 25 - 30 less premature deaths annually in the Stockholm metropolitan area. These are approximately three times larger effects than what would be found if a more general policy measure, such as a fuel tax increase, was used to obtain a decrease of emissions of an equal magnitude, since these reductions were concentrated to the most densely populated areas.

One of the goals for the trial was to "improve the perceived urban environment". It is difficult to draw definitive conclusions about this, not least because of confounding with weather effects. However, studies indicate certain improvements for traffic-related indicators such as perceived traffic congestion, air quality and accessibility – not only for car drivers but also for cyclists and pedestrians. The result points to perceived improvements of exactly those factors for which measured changes can be demonstrated, i.e. those connected to traffic reductions. In the city environment study, citizens feel there is an improvement in traffic tempo, air quality and vehicle accessibility. The same tendency is seen in interviews with cyclists in the inner city and children living in the inner city. Inner-city children's perception of the city environment has very clearly improved and many cyclists think there are fewer cars in the inner city and that the traffic environment has got better.

3.5 Retail

There had been fears that retail inside the cordon would be adversely affected, but studies of the retail markets were not able to show any effects of the congestion charges (Daunfeldt, Rudholm, & Rämme, 2009). For example, the durables survey in shopping centres, malls and department stores during the Stockholm Trial period showed that these developed at the same rate as the rest of the country. The same held for other retail sectors.

Fear for adverse impact on retail inside a cordon is common in many cities. Large efforts were made in Stockholm to track such effects, only to conclude that they were very small or non-existent. Similar conclusions have been reached in other cities with congestion charges. There may be effects on particular stores, especially if they lie close to a cordon, but the average effect in an urban centre is usually small. This should be evident, especially in the long term: if the retail market inside the cordon gets less attractive, then floorspace rents will, in equilibrium, decrease to counteract this, making the effect on the number of stores even smaller.

---

9 The standard cost-benefit model used for infrastructure planning in Sweden has a dose-response relationship based on older studies. This indicates that the reduction in traffic due to the trial saved about five life years on an annual basis. This is the number used in the cost-benefit analysis by Eliasson (this issue).
3.6 What did the disappearing drivers do instead?

Traffic measurements will only give changes in aggregate terms, and does not reveal how car drivers adapted – how many switch to other modes, departure times or destinations, for example. A panel travel survey with two waves was used by Franklin, Karlström and Eliasson (2010) to investigate how private trips were affected. Results (in 1000's car trips across the cordon during charged hours) are shown in Figure 7. These results should be treated with caution, since it is difficult to separate seasonal and trend effects from the effects of the charges.

Approximately 25% person trips across the cordon disappeared. Out of these, around 10 percentage points out were work trips switching to transit, while one percentage point was work trips switching to the Essinge Bypass (hence, route switching was only a minor adaptation strategy). Six percentage points were discretionary trips switching to other destinations or reducing trip frequency, possibly by trip chaining or combining trip purposes, and under one percentage point switching to the Essinge Bypass. The remaining five percentage points are disappearing professional traffic—deliveries, taxi, craftsmen etc. Since we do not have travel surveys for this type of traffic, we cannot decompose it further. That professional traffic is affected at all by the relatively low charges may come as a surprise, but there is significant evidence in the interview studies carried out with professional drivers that they in fact tried to plan their routes and trip chains in order not to cross the cordon unnecessarily often, and moreover to decrease the number of trips altogether.

![Figure 7. Estimated changes in car trips across the cordon during charged hours (1000's trips).](https://example.com/figure7)

*When comparing to measurements of vehicles across the cordon, note that “through” trips are counted as two cordon crossings.*

It is worth emphasizing that there are many ways to adapt. Route and mode changes are far from the only adaptation strategies. Trips, especially discretionary trips, are not “replaced” in a simple one-to-one fashion. Many people, traffic experts not least, seem to be unconsciously stuck with the assumption that there is a more or less fixed number of trips to be made, and that the effect of the charges should be possible to sort neatly into categories like “mode change”, “destination change” and “departure time...
In reality, adaptations are much more multi-faceted. This means that commonly encountered statements such as "congestion charging won't work in our city because our transit system is too bad" or "...because we have no ring road" are miss an essential point: there are many more ways to adapt than changing mode or route. Which adaptation strategy will dominate depends on the characteristics of the city and the design of the system, and which travel alternatives it leaves open.

Another important insight is that traffic isn't just work trips. Work trips only make up a fraction of car traffic - a typical figure could be 30-40%, with the rest being discretionary trips and professional traffic (where a typical figure could be 15-25%). Discretionary trips are easier to affect, because there are more ways to adapt in the short run, and represent a significant fraction of traffic, especially during afternoon peak hours, when congestion is often just as severe as during the morning peak. Professional trips are very heterogeneous: some types are very difficult to change, while some are not. Typically, values of time are very high, which means that time savings for professional traffic will constitute a significant part of the travel time benefits. Despite all this, it's common that the discussion focuses exclusively on work trips, both among planners, policy makers and the general public. This is a mistake that surprisingly often confuses the discussion about what congestion charging can do and how they may work.

Over time, the question of how drivers adapted becomes increasingly pointless. This is because travel patterns are much less repetitive and stable than many people think. Many of the affected drivers are occasional car drivers, who drive on the charged road perhaps a couple of times each month. Around two thirds of the drivers crossing the Stockholm congestion charging cordon on any given day are "occasional" drivers, who drive across the cordon three days a week or less. Other days, they use other modes, times or routes. These drivers will change on the margin, and it will often be impossible to tell if and how they changed. In fact, many car drivers will not even know if or in what way they adapted. A study in the spring of 2006 showed most drivers were unaware that they had reduced their trips across the cordon. A comparison of drivers’ stated change in behaviour and objective traffic measurements showed that around ¾ of the decrease in trips had apparently gone unnoticed by drivers.

Moreover, there are many other changing processes going on. People move and change jobs, for example: between any two years, 20-25% of the workforce will have changed jobs (or started working), and 15-20% of the population will have moved. After just a few years, it is pointless to ask how a given person has “adapted” – because the entire situation where travel choices are made has changed.

3.7 Public transport

The Stockholm trial consisted not only by congestion charges but also of an extension of public transit services. The extended services were motivated partly to meet increased demand for public transport, and partly by a political will to show “carrots” and not just “sticks” (Kottenhoff & Brundell Freij, 2009).

Drivers switching from car to public transport meant that the number of passengers in the transit system increased by around 4-5%. Crowding in the public transport system, measured by the number of standing passengers, increased somewhat in the metro but decreased on the commuter trains, most likely thanks to expanded public transport capacity.
Reduced road congestion in and around the inner city led to increased speeds and punctuality for bus services. During the trial period in 2006, bus timetables were not adjusted so the improved accessibility did not significantly shorten travel times for inner-city buses, but there were signs of improved punctuality. Bus traffic across the charge cordon – which do not have fixed time tables once they have passed the cordon - experienced considerably shorter travel times. After the trial period, no dedicated studies of this are available, since the counterfactual becomes more and more hypothetical.

Travel surveys showed that few car drivers were enticed to switch to public transport by the service extension in itself, even if the data are inconclusive. The second purpose was to provide additional capacity in the transit system. The added capacity was relatively minor: the increase in passenger capacity by the buses meant that around 14 000 trips were made each day by the buses, compared to well over one million public transport trips across the cordon each day. But in the specific corridors served by the buses, they most likely contributed to keeping crowding on commuter trains from increasing. The third purpose of the transit extension was to increase the effect of the congestion charge by making the switch from car to public transport easier. However, onboard surveys on the new buses found very few former car drivers on the buses. Of the vehicle-traffic reduction of 22% over the charge cordon, at most 0.1% can be ascribed to the extended bus services (Eliasson et al., 2009).

4 PUBLIC AND POLITICAL ATTITUDES

The main obstacle for congestion charging is often the lack of public acceptability. Hence, the most remarkable and interesting development in Stockholm may be the change in attitudes, from fiercely hostile to overwhelmingly positive. This section presents some evidence about this development, and discusses what factors influenced it.

4.1 The change in attitude – “familiarity breeds acceptability”

When the decision was made to carry out a congestion charging trial in Stockholm, it was met with vocal resistance, although polls showed a more mixed picture. Figure 8 shows how the support for congestion charges has evolved over time. In the spring of 2004, 43% of Stockholm citizens stated that they would “probably” or “most likely” vote yes to permanent congestion charges. Support fell, however, once the start of the trial approached. Right before the start of the trial, support had fallen to 34%, with the “most likely yes” group falling the most. Once the trial started, however, support increased to 53%. The media image also changed once charges were in place, from intensely critical to, in many cases, very positive. The percentage of trial-related newspaper articles with a positive angle increased from 3 % in the autumn of 2005 to 42 % in the spring of 2006, while the share of negative newspaper articles was almost halved from 39 % to 22 % (Winslott-Hiselius et al., 2009). The trial ended on 31 July, 2006, and was followed by a referendum in September at the same time as general and local-government elections were held. Excluding blank votes, 53% of Stockholm citizens voted to keep the charges. After the election, the centre/right coalition gained power both at the national level and in the city of Stockholm. The centre/right coalition in Stockholm had opposed the congestion charges, but had promised to follow the outcome of the referendum, so they had to ask the national Government to reintroduce the charges permanently. After a few weeks of consideration, the new centre/right Government said it would do so, but as part of a broader package of transport services.

10 This section draws heavily from Eliasson (2014).
investments in Stockholm, to be negotiated. The revenues from the congestion charges were earmarked for road investments. On the other hand, the investment package also contained major rail investments, but these were claimed to be financed by other sources of funding. After the decision to include the charges in an investment package, no political parties proposed abolishing them anymore. The charges were reintroduced permanently in August 2007, although the negotiation over revenue use was not settled until late 2007. A poll in December 2007 showed a 65% support for the charges, and several polls since then have shown similar or higher support. The most recent, in late 2013 (not shown in the diagram), showed a support of 72% (excluding "don't know").

Figure 8. Would vote "yes" in referendum about congestion pricing (excl. "Don't know").

Figure 9 shows support for the charges in four groups: people without car in the household, car owners who never or very seldom cross the charge cordon, car owners who sometimes pay the charge, and car owners who often pay the charge. (2007 and 2010 data are missing, since these surveys did not ask about driving across the cordon.) The support shows the same U-shape in each group. In fact, the dip from 2004 to 2005 is most pronounced for the unaffected groups, i.e. people not expecting to pay the charge.
Figure 9. Support for congestion charges depending on car ownership and paid charges.

Evidently, the amount of tolls paid makes a large difference for the support, looking at each cross-section. But the figure also indicates that the change in attitudes is at least partly driven by other factors than self-interest variables such as tolls paid and time gains. The changes in attitude over time for each group look very similar, regardless of how much they are affected by the changes in travel costs and travel times. Even the group who do not own a car shows the same U-shaped attitude pattern. And even in the most affected group, support for the charges has more than tripled from a low point of 15% to 53%. In other words, there is a majority in favour of the charges in all groups by 2011.

Just as respondents seem to be unaware of their behavioural changes, they seem to be remarkably unable to remember their past attitudes. This is actually well known from voting research: people tend to forget that they have ever had another opinion than their current one, or have ever voted differently than they would vote now. Surveys in 2006 and 2007 asked respondents whether they had changed their attitude to the charges. From 2005 and 2006, voters intending to vote “yes” in the referendum increased by 19 percentage points from 30% to 49% (including undecided voters in the base). 29 percentage points of the 49% yes-voters stated that they had “become more positive” during 2006 – that is, some of the positive voters must have become even more positive. But when the same question is repeated in 2007, only 13% of voters state that they had “become more positive” during 2006. Not only is this less than half the 29% from 2006, it is lower than the 19 percentage points that became so much more positive that they changed from “no” to “yes”.

The development of attitudes is remarkably close to the general pattern described in Goodwin (2006), reproduced in Figure 10.
Once the idea has been introduced and explained, a fairly large fraction of the population is generally willing to support the idea of congestion pricing. How large this fraction is depends on how the question is formulated and framed – for example, revenue use, the purpose of the charges and what policy alternatives it is contrasted against all matter. But once a detailed proposal is worked out, support generally decreases. There may be several reasons for this – for example, that the disadvantages suddenly become more evident than the potential advantages, or fears that the technical system will not work or become very expensive. This is sometimes summarised in the formula “acceptability decreases with detail”. But once the system is in place, support will generally increase, which is often summarised as “familiarity breeds acceptability”. There are probably several reasons for acceptability to increase once a system is in place.

1. “Better than you thought”: the benefits may turn out to be larger than anticipated. A major reason for public resistance to congestion charges is a belief that they simply will not work (Bartley, 1995; Jones, 2003). In Stockholm, the positive effects on road congestion and urban environment were much larger than most people expected.

2. “Not as bad as you thought”: the downsides of charges – increased travel costs and/or changes in travel behaviour – may prove to be not as bad as expected. Once the charges are in place, many people may discover that the charges do not in fact affect them negatively as much as they had feared. There is a evidence that this phenomenon was important in Stockholm (Henriksson, 2009). Transit crowding didn’t increase significantly, congestion didn’t appear in new places, the technical system worked smoothly, and it may have been easier to adapt to the charges than anticipated (for example, drivers have reduced their car trips across the cordon around four times more than they are aware of, according to a study which asked drivers how they had changed their behaviour and compared this to actual traffic measurements).
3. Familiarity with road user charging may reduce the general reluctance towards pricing a previously unpriced good. There is evidence that people in many cases do not like prices as an allocation mechanism (Frey, 2003; Jones, 2003). But once familiar with the concept that road space is in principle a scarce good that can be priced – much like parking space or telecommunication capacity – this reluctance may tend to decrease.

4. Finally, it is a general psychological phenomenon to resist change. In the psychological literature, different variants of this general phenomenon are referred to as status quo bias, loss aversion, endowment effect, or cognitive dissonance. This means that once people have adapted to the charges and the short travel times, they value the time gains higher and the increased travel costs lower than they did before the change. Some people may react with “accept the unavoidable”: once the charges are in place, it is less worthwhile to spend energy on opposing them.

Further analysis of Stockholm attitudes show that even if self-interest and belief in effects strongly affect attitudes, these are not the sole determining factor, especially when looking at how attitudes change over time. All groups, regardless of travel patterns, car ownership and belief in the charges’ effectiveness (ex ante and ex post) show the same U-shaped change in attitudes – more negative attitudes before the introduction of the charges, and increasingly more positive attitudes after the introduction. In fact, this pattern is more pronounced for unaffected groups.

This can be interpreted using results from social psychology explaining how attitudes are formed. First of all, attitudes are more or less well developed. They tend to be more developed in issues where an individual for example has a lot of direct experience, has encountered the issue many times, know a lot about the issue and towards which they have strong emotions. When people are faced with a new issue where attitudes are not well developed, new attitudes are often formed by associating the new issue to some similar but familiar issue, where the individual already has a well-developed attitude.

There have been suggestions that the introduction of congestion pricing established a new social norm, where driving in rush hours was less socially accepted. There is scant evidence of this, however. In fact, it is easier to interpret the introduction, polarization and subsequent lack of controversy in the opposite way – that the charges could gain broad acceptance once it was shown that it could in fact be reconciled with existing social norms: in particular, that the charges were not an attack on mobility or car use as such. This interpretation simultaneously explains why it was politically rational to first ignore congestion charges, then to advocate them in spite of public resistance from large groups, and why this resistance then died down. It also illustrates the importance of legitimacy. The story has four phases.

1. For a long time, planners and economists in Stockholm had limited success in advocating congestion pricing. Their argument was that it would increases efficiency in the transport system. But very few people have a pre-existing attitude to this concept, and even fewer have an emotional engagement in it. Instead, when faced with the question, people associated to superficially similar issues such as mobility restrictions or taxation, where they have an existing attitude, most likely a negative one. For an issue to be politically interesting, it must generate enthusiasm among a sufficiently large group of voters. But since transport efficiency is simply not an issue that many people get enthusiastic about, the issue had virtually no political upside.

2. This changed when congestion pricing was reinterpreted as an environmental policy, which happened in Stockholm during the mid 1990’s. While allocation
efficiency in the transport sector could not arouse enthusiasm or engagement among the general public, environmental concerns definitely could. This was what was needed to get congestion pricing on the political agenda – a link to an area with where strong and emotional attitudes existed.

3. When the decision to carry out the congestion charging trial was made after the election in 2002, a fierce debate broke out. Consistent with what was said above about the necessity of emotions in politics, the arguments soon turned principal, moral and emotional, leaving little room for compromise. This might have been an inevitable development: if congestion pricing had not been elevated to a moral-emotional question, it hadn’t entered the political stage in the first place. But just as inevitable, the morally supercharged arguments for congestion pricing implied (or could be perceived to imply) that all car traffic was evil and unnecessary, and should be banished. This might be one reason for the decreasing support also among car drivers that were actually unaffected: they might simply have been alienated by the anti-car rhetoric. There are also other reasons that may have caused non-affected groups to develop more negative attitudes: the charges were claimed to have adverse equity effects, be unfair, and a waste of taxpayers’ money. The most recurring argument was the lack of democratic legitimacy. The social democrats had made a very clear promise not to introduce road pricing during the election period – and here they were doing it anyway.

4. When the referendum ended in a yes to the charges, the new government decided to earmark the revenues for a motorway tunnel west of Stockholm. From an attitude point of view, this was probably important for several reasons. First, the charges now had democratic legitimacy. In addition to the referendum result, there was now a political agreement about the charges and the revenues – made by the liberal/conservative alliance no less, which meant that all political parties had now sanctioned the charges in some way. Second, the revenues were earmarked for roads. As it was really part of a multimodal package, the revenues could just as well have been earmarked for the railway investments that were also part of the package. But earmarking the revenues for roads not only spoke to motorists’ self-interest. It sent a moral signal: it’s OK to be a car driver. It indicated a reinterpretation or reassociation of the congestion pricing issue from a morally charged anti-car measure to a technical-rational measure that was effective – it “worked” in the sense that it generated revenues and reduced congestion. And technical measures arouse much less emotions: people usually do not love them, but they do not hate them either. The most important function of the earmarking may hence have been to discharge some of the sentiments around the charges, moving the debate from the moral domain to the technical-rational domain. Thirdly, it calmed the fears of Stockholm politicians (from all parties) that revenues would end up in the national coffers, either directly or indirectly, by subtracting the revenues from Stockholm’s “fair share” of national infrastructure grants.

The introduction of a certain measure always has the potential of becoming politicized. Congestion pricing is an evident example: pricing as such may meet resistance even among people who agree in principle that road traffic should be reduced. In fact, regulation (bans on car driving in certain areas, odd/even number plate restrictions, slowing down traffic through physical measures) are often more popular than pricing measures, despite being inefficient in the standard economic sense. Much of the debate before the charges were introduced centred around whether pricing was, in principle, a fair, legitimate and democratic allocation mechanism. After the charges were in place, this debate faded into the background – pricing seemed to become a reasonable
The Stockholm congestion charges

accepted way to allocate road space, just as with many other scarce goods (including transport-related things as train tickets, airport slots or car ownership).

One way to view the congestion charging debate is that the issue seems to have moved from the technical-rational domain to the moral domain and back again. When presented as a purely technical-rational suggestion, it failed to gain political interest because the type of benefits it could potentially bring (increased transport efficiency) could not generate sufficient enthusiasm. It seems that too few people had strong attitudes regarding this type of benefits, meaning that it could not generate the necessary emotion to gain political traction.

By reinterpreting it as an environmental measure, the issue moved gradually to the moral domain. This connected to strong attitudes regarding local and global environment, and maybe also general anti-car sentiments in some groups, and hence the necessary political engagement emerged. But the other side of the coin is that it made the issue divisive – even unaffected groups became more negative. These negative sentiments were probably bolstered by other moral arguments, e.g. about lack of democratic legitimacy, waste of public funds, over-taxation and freedom of mobility. These attitudes were also strong and well-developed, making the debate very heated.

But after the referendum, the infrastructure agreement and the establishment of democratic legitimacy moved the issue back into the technical-rational domain. This discharged some of the negative moral-based attitudes. Moreover, this connected the issue to concepts such as “rationality” and “efficiency”, which in Sweden have very positive associations.

4.2 Other factors influencing public attitudes

Several factors influencing attitudes to congestion charges have already been mentioned above, but it is worth summarizing what kinds of factors have repeatedly been found to influence acceptability in several countries.

First, self-interest variables are obviously important. All else equal, individuals get more positive the less charges they pay (or expect to pay), the more time gains they get, the higher they value travel time savings, and the more satisfied they are with public transport. Individuals also become more positive if revenues are used in a way they appreciate, which can be viewed as a form of self-interest (Eliasson & Jonsson, 2011; Hamilton & Eliasson, 2012; Hårsman & Quigley, 2010; Schade & Schlag, 2003a). Among other things, this means that the system has got to deliver benefits. In Stockholm, the perceived effect of the charges was the most important factor explaining attitudes to the charges. Even if one should not confuse “perceived” effects with “objective” effects – since attitudes influences what effects are actually perceived – it seems clear that achieving objective effects is necessary to reach acceptance. This underscores the importance of designing the system carefully, and only using congestion charges when congestion really is a problem. Moreover, it seems likely that measuring effects and communicating the results through, for example, the kind of scientific evaluation carried out in Stockholm will increase the awareness of positive effects – provided, of course, that there are in fact positive effects.

Second, positive attitudes to congestion charges are strongly influenced by concerns about and engagement in environmental issues (Eliasson & Jonsson, 2011; Hamilton & Eliasson, 2012). In Stockholm, the charges were to a certain extent marketed as “environmental” charges, and voters’ environmental concern was an important factor explaining the acceptability of the charges. This is in line with findings in the literature
that social norms of this type influence acceptability in general, and that support depends not only on the ‘objective’ characteristics of the measure itself, but also on the defined objective of congestion charges. Moreover, several authors have found that it is not just perceived individual benefits that determine acceptability: perceived social or collective costs and benefits can also affect acceptability strongly. Hence, the “branding” of the charges matters – how they are marketed, explained and perceived. In Stockholm, re-labelling congestion charges to “environmental charges” and emphasizing their positive effects on air quality may very well have had an impact on acceptability. Other cities may employ different strategies, but the general conclusion remains: it is important how the charges are “branded”. A condition for this to be possible is that the system design is well aligned with the stated purpose of the charges. A system marketed as “congestion charges” system should for example not levy charges where or when there is no congestion.

As was emphasized above, the most important factor seems to be own experience of congestion pricing. The same pattern has been observed in e.g. Oslo and London. One study (Hamilton & Eliasson, 2012) compared attitudes in Stockholm, Helsinki and Lyon, concluding that the only variable that could explain the much higher support for congestion charges in Stockholm was simply that the Stockholm population had experienced the introduction of congestion pricing, while the others had not.

Congestion is viewed as one of the most important urban problems (65-80% of the three populations agreed). But Hamilton et al. found no significant correlation between attitudes to congestion pricing and concerns about road congestion. On the other hand, they found a strong correlation between concerns about road congestion and being in favour of expanding road capacity. Apparently, it is not mainly concerns about congestion that is driving support for congestion pricing, despite the fact that such concerns are widespread. This is consistent with earlier several studies finding that one of the most common arguments against congestion pricing is a distrust in congestion pricing’s ability to reduce congestion (Jones, 2003; Schade & Schlag, 2003b). On the other hand, this distrust may partially be a reflection of self-interest: Schade and Baum (2007) find that respondents who expect congestion pricing to be disadvantageous to themselves not only have more negative attitudes to it, but also perceive them as less effective and more unfair than other respondents.

Congestion pricing attitudes are related to attitudes to public interventions in general. Hamilton et al. show that negative attitudes to congestion pricing are strongly correlated with negative attitudes to taxation in general, speed enforcement cameras, and belief in a public administration’s ability to distribute a scarce resource fairly. This finding may partly explain the apparent paradox that left-wing parties are often more in favour of congestion pricing than liberal/conservative parties.

Equity effects are often cited as one of the main reasons for opposition to congestion pricing. Whether congestion pricing has progressive or regressive effects depend on the design of the system and on initial travel patterns. As to Stockholm, there are several studies (Eliasson & Levander, 2006; Eliasson & Mattsson, 2006; Franklin et al, 2010; Karlström & Franklin, 2009). These find no regressive effects; some indicate progressive effects, while some indicate neutral effects. It should be noted that the equity argument may be used for other reasons than honesty equity concerns: it may simply be perceived as a more legitimate argument than self-interest (Schade & Baum,

---

11 Other studies have got mixed results on this issue. While Rienstra, Rietveld and Verhoef (1999) found that respondents with more concerns about congestion were more positive towards congestion charges, Härman et al. (2000) and Schade et al. (1999) found the opposite.
The Stockholm congestion charges

2007). This is supported by the finding that Hamilton et al. found weak or no correlation between the attitude to congestion pricing and agreeing with the statement “More should be done to reduce the difference between rich and poor in society”.

Equity and “fairness” can be interpreted in different ways. Initially, the dominating perspective is often “before-after” – how travel costs and travel times change for different groups, such as rich vs. poor, men vs. women, inner city vs. suburb residents. At least in cities with decent transit shares, it is often the case that “rich” will pay more than “poor”, with middle-income groups “suffering” the most, relatively speaking. But once the charges are in place, another perspective becomes more important – “fair” pricing. In other words, what price is “fair” to charge? From this perspective, it is “fair” that one pays more to drive on a congested road or to cause emissions in densely populated areas – irrespective of income or place of residence, for example.

4.3 Political acceptability

Political acceptability is different from public acceptability. Obviously, political acceptability is influenced by the level of public acceptability – but public acceptability is neither a necessary nor a sufficient condition for political acceptability. Crucial for the analysis and understanding of political acceptability are power issues: the power over the design of the charging scheme, the power over the revenues, and how the charges and their revenue stream will affect decisions and funding of transport investments in general. The fact that congestion charges are now politically accepted in Sweden is not only, or perhaps not even primarily, due to the higher public support. It is also because the charges have been integrated in the general transport investment planning process, and this has – at least partly – solved the power and negotiation issues above.

To understand the political and institutional drivers behind this development, one must start with the legal context. Swedish congestion charges are not “charges” but national “taxes” from a legal point of view. Existing infrastructure cannot be “charged”, only “taxed”, according to the constitution’s definition of a charge, and Swedish municipalities cannot levy taxes on other than their own citizens. Hence, although it was the city of Stockholm that was responsible for designing the charging system and carrying out the congestion charging trial, the responsibility for actually levying and administering the charges had to be assumed by the national government through a parliamentary decision. More important, this meant that it is the national government that has the formal power over both scheme design and revenues. Although the Government promised to refund the revenues to the Stockholm region, disagreements quickly emerged regarding how revenues should be calculated, how revenues should be used and which vehicles should be exempt. Further disagreements, such as whether and how charge levels should change along with inflation and economic growth, can be expected. Many politicians have stated that their main argument against introducing the congestion charge was the uncertainty about the political power over scheme design and revenues.

Adding to these uncertainties was the uncertainty about how the existence of the new revenue stream would affect the complicated negotiation between national and regional levels about national infrastructure grants. Most of the major transport

---

12 A “municipality” (“kommun” in Swedish) is the smallest geographical administrative unit in Sweden, roughly corresponding to a city. Most of the spatial planning responsibility, including infrastructure planning, lies at the municipal level.

13 This task was given to the National Road Administration, and later moved to the National Transport Agency.
investments in Sweden are paid for by the national government, whereas municipalities and regions are responsible for local streets and transit operation. As expected, there is often disagreement on where the border between different responsibilities should lie. The politicians in Stockholm, regardless of political colour, had long argued that they were not receiving their fair share of national infrastructure grants. Whether this claim was founded or not, it meant that the arrival of a new revenue stream in the form of congestion charges was not necessarily welcomed. Several politicians feared this would mean that Stockholm would have to pay an even larger share of transport investments with their own money. The government, they argued, would point to the revenues from the congestion charges and claim that Stockholm obviously needed even less national infrastructure grants than before.

The solution to this dilemma was an agreement where the charge revenues were funding parts of a major transport investment package, where the national government also made a major funding commitment – much larger than had been the case for a long time. The charge revenues were earmarked for the road investments in the agreement, while the substantial rail investments were claimed to be paid for with money from other sources. With this, support for the charges had been secured from regional politicians of all parties.

From a political perspective, it is often a decisive question who has the power over revenues and charge levels. If it is the national level, then regions and cities will obviously be much more reluctant to introduce charges. But even if the region keeps the revenues, another issue is important: how the existence of this new revenue stream affects the complicated negotiation between national and regional levels about national infrastructure grants. In Norway, this problem is handled by a decision that the state "matches" income from regional charges with equal national funding. A recent trend in Sweden is that national funding is often leveraged with regional funding, sometimes from congestion pricing or toll schemes. This has made congestion charges more popular among politicians, which shows the importance of institutional context and incentives.

Giving regions the incentive and opportunity to introduce road user charges to obtain transport investments, where regional funds are leveraged by national funds, may fundamentally change the transport investment planning process. There are several advantages: regions are given an incentive to prioritize between transport investments and other responsibilities, as they are forced to "put their money where their mouth is". When there is congestion, regions are more likely to introduce congestion charges, which is obviously a potent and efficient policy measure. On the other hand, there are several disadvantages: since regional funding is leveraged, regions will be tempted to overinvest in transport infrastructure relative to other types of (non-leveraged) spending.

5 THE SYSTEM DESIGN PROCESS

5.1 Designing the charges

The goal of the system, as formulated by the politicians, was to "reduce congestion on the most congested roads and improve speed through the bottlenecks". This goal was then quantified, somewhat illogically, as a decrease in traffic volumes: the number of vehicles crossing the cordon should be reduced by 10-15%, a number loosely based on previous proposals for congestion charging systems.
Based on transport model simulations made before the trial a 10-15 percent reduction of the number of vehicles crossing the cordon was expected during charging hours. In fact, the traffic forecasts predicted a larger and, as it turned out, more correct predicted magnitude (Eliasson et al. 2003a, 2003b, 2004) – but such a large decrease seemed unreasonable at the time, even to the modellers themselves. No forecasts were made of the effects on accessibility, as the static network equilibrium models available for forecasting were considered as not being reliable for such purposes.

During the process of designing the charges using traffic models, it quickly became apparent that a target was needed that was directly related to travel speeds. The traffic volume reduction target, although easy to communicate and measure, simply revealed too little information about the effect on travel speeds. An obvious candidate for such an alternative target was to try to maximize social surplus. However, this was problematic because the travel times of the static network equilibrium model (which was the only available traffic model) were judged to be too unreliable at high congestion levels. Instead, the most important target used during the design process was the congestion levels at major bottlenecks.

Several variants of the charging scheme were tested, of which the most important variations were higher or lower charge levels and charges on the some of the bridges inside the cordon. There is a potential conflict between making congestion charging “efficient” in the theoretical sense, and making it easily understood (Bonsall, Shires, Maule, Matthews, & Beale, 2007). In the design of the Stockholm system, ease of understanding was a major consideration. This explains the symmetries of the design: one single charging cordon with the same charge at all points of entry, the same charge in both directions, and the same for both the morning and afternoon/evening peak periods. The system could most likely have been made more efficient if these symmetries had been abandoned – but at the time, ease of understanding was judged to be more important than improving efficiency, and it is possible that this might have contributed both to acceptability and effectiveness.

A lesson from trying out different designs in the forecasting models was that it was easy to design a charging system that created more problems than it solved, by “moving around” congestion. Hence, spending a lot of effort on the design of the charging system is imperative for any city wanting to introduce congestion charges.

The Stockholm Trial provides interesting insights into what a road-toll system should look like – something which is also useful for other cities. Transport planners and economists have long discussed to what extent a charge-zone toll of the kind used in Stockholm is sufficient for controlling traffic in an entire city. Traffic relations change from street to street and from minute to minute. When the charge zone is as large as it is in Stockholm, there was concern that even if it had a big effect on travel over the charge cordon, streets inside the zone would soon be full of motorists already in the zone increasing travel as they realized the streets were less congested. Alternative solutions were discussed for several years prior to the Stockholm Trial, involving several sub-zones with varying rates of the congestion tax. None of the existing road-toll systems threw much light on this question. In London, it is a question of a small area in the city centre, in Singapore access to cars is also regulated and in Oslo and Bergen the system is designed to affect traffic as little as possible. The Stockholm Trial confirms that a simple charge-zone toll creates significant effects within a large area.
5.2 Can transport models be trusted?

Eliasson et al. (2013) provide a detailed comparison between forecast and outcome, concluding that the main predictions about behavioural responses were sufficiently accurate to draw correct conclusions. For example, traffic across the cordon was predicted to decrease 17% during peak hours and 16% during the entire charged period (6:30-18:30); the actual figures were 19% and 20%. The transport model predicted that around half of the disappearing trips would switch to public transport, which would lead to a 6% increase in passenger volumes; in reality, around half of the drivers did in fact switch to public transport, leading to an increase of passenger volumes of 4-5%.

In fact, the model seems to be much better at predicting changes in behaviour than the travellers themselves, both *ex ante* and *ex post*. Surveys in the fall of 2004, the fall of 2005 and the spring of 2006 asked respondents about changes in their travel patterns in response to the charges. Respondents gave reasonably consistent answers in the three surveys. The answers can be transformed to an equivalent aggregate traffic reduction, yielding an equivalent aggregate traffic reduction of 5-10%. This can be compared to an observed reduction of private trips\textsuperscript{14} around 30%. In other words, around 3/4 of the reduction in car trips across the cordon seems to have gone unnoticed by the travellers themselves.

The one major shortcoming of the model system was the prediction of travel times. Some travel times were predicted fairly accurately, but some effects were wide of the mark. This was the case especially for links subject to spillback congestion (queues propagating back from a bottleneck, blocking other junctions and links). For links crossing the cordon, model-predicted travel time reductions were close to observations. For links within the cordon, model predictions were less accurate: delay reductions were underpredicted by 34 per cent, since the network does not model congestion at junctions, traffic lights etc. For links outside the cordon, predictions were completely off, since a static model does not capture spillback congestion. Hence, it was impossible for the network model to foresee that congestion outside the cordon will decrease when the queues propagating upstream from bottlenecks located at the cordon are reduced. This was a known issue during the design of the charging system. Rather than using travel time-based evaluation measures (such as consumer surplus or congestion indices), design targets were formulated in terms of volume/capacity ratios in the most important bottlenecks.

The results indicate that best-practice transport models seem to be reliable enough to be used as decision support and design tools even for substantial changes of the transport system – provided that the analysts are aware of inherent limitations of the model and interpret results accordingly. It should be stressed that the predicted effects of the charges were so large that several experts considered the forecasts unrealistic. As it turned out, however, the model gave much more accurate predictions than experts’ judgments, in addition to providing more detail and consistency. The scepticism towards the predictions was understandable: the introduction of the congestion charges affected the whole Stockholm transport system in a completely unprecedented way. Traffic across the cordon decreased over 20%, meaning that traffic was down to levels not seen the 1970’s; reducing queuing times by 30-50%. On some links and routes, the effects were even larger. Despite this, circumferential traffic did not increase, and public transport ridership increased by just a few per cent.

\textsuperscript{14} Total traffic reduction is less because professional traffic decreased much less than private traffic.
One way to judge whether the transport model was "good enough" is to consider whether the system had been designed differently, different preparatory measures had been undertaken, or the scheme been abolished altogether if the forecast had been perfect in all respects. Generally speaking, the answer is a qualified "no". The qualification is that the model's deficiencies, in particular the lack of dynamic congestion representation and departure time modelling, may have had become more of a problem if the system design had been more complex, with more charging points and more fine-tuned time/place differentiation. But given that the system design was constrained to be relatively simple, the model was good enough to answer the most important design questions: what traffic reduction was needed to reduce queues significantly, what charge levels were needed to accomplish this reduction, and what the secondary consequences in terms of possible traffic rerouting and transit crowding would be. One of the authors headed the design process, and can confirm the wide-spread observation that the effects of a significant change in a complex transport system are usually too complex and multi-dimensional to foresee without the help of a model. The most important advantage of using a transport model may not be that it gives exact answers, but that it gives coherent answers. During the design process, the model repeatedly gave results that were surprising at first, but were self-evident and easy to explain intuitively after some thinking. The point is that these "intuitive" explanations and conclusions had not been realized before the model results had been produced. In this sense, the model turned out to be an indispensable tool for system design and evaluation, and several design suggestions were discarded after model results had shown that they would not work satisfactorily.

5.3 General advice on designing congestion charges

The goals have to be explicit and relevant
First, the system needs a goal. The goal may be to reduce congestion reduction, improve air quality, yield revenues, or a combination of such goals. Whatever the goals are, they need to be explicit. Moreover, they should be quantified, at least to some extent. This quantification usually has to be done in cooperation between policymakers and traffic experts: setting up relevant goals and targets is harder than most policymakers realise. Goals must above all be relevant and consistent. Specifically, one should at this stage not specifically strive to make them easy to communicate to the public. Communication is important, but comes later. The goals set at this stage are the ones that will be used during the design process, and they need to be consistent and relevant, not necessarily easy to explain or sound good. An example of a consistent and relevant goal that happens to be rather difficult to communicate is "achieve maximal social benefit from congestion reduction" (perhaps given some restriction on charge levels). A common example of a goal popular among policymakers and communicators is "getting more people to choose public transit". This is not a relevant goal for congestion charges, which should be obvious. (The relevant goal is to make less people choose car during congested hours. If they instead choose transit, that's fine; if they prefer to adapt by cancelling trips or changing departure times or destinations, then this is just as fine.) Choosing ill-formulated goals and targets will very likely cause problems during the design process, at the very least causing confused discussions.

Designing the charges is a job for experts.
Designing a charging system is, as a rule, a very difficult task – how difficult depends on the topology of the city. (For example, Stockholm is reasonably easy, with the worst congestion problems located along a natural cordon, while Gothenburg is difficult, with congestion problems spreading out from a complicated multiple-arterial junction.) It is absolutely necessary to have sufficient time, and access to a reasonably good transport
model. If one has access to design optimisation tools, this can come in very handy. Even given this, it will be difficult. In particular, intuition and prior knowledge will in general not be sufficient, even for experienced traffic planners: transport systems are simply too complex. There will almost certainly be surprises, and the first attempt at a charging system design will most likely not be optimal or even good – it may even make congestion worse overall by “moving congestion around”. The system design is an iterative process, where involvement of politicians does not help. This is why it is so important that goals are stated clearly at the outset. Design and goal-formulating is, ideally, an iterative process as well: it is likely that some goals or (more likely) some design restrictions were forgotten at the outset. But this does not change the basic premise that while formulating goals and restrictions is a job for policymakers, designing the details of the system – locations and levels of the charges – is a job for experts. An ill-designed system may not only be suboptimal – it may likely cause more problems than before.

**You need a good transport model.**

Most transport models are constructed for other purposes than modelling the impact of congestion charging. Certain shortcomings of most current models become especially important in the context of congestion charges, and one needs to be aware of them. First, the value of time differs between vehicles, but this is often neglected in the assignment step. Technically, one must use multi-class assignment, i.e. divide traffic into several “classes”, each with a value of time of their own. The value of time of each class will decide whether it is worth taking a detour to avoid a charge. Depending on the network topology, the value of time distribution over classes may affect results strongly. Often, there is little evidence on the value of time distribution, so sensitivity analyses will play an important role. Second, departure time choices and scheduling considerations are often sketchily implemented, if at all. Obviously, this will underestimate the impact of a time-differentiated charging system, since the opportunity to adapt by changing departure time is not reflected in the model. Less obvious, one may underestimate traffic decreases during non-charged hours – since those trips is partly made up of “return trips”, i.e. the second leg of a trip whose first leg was during the charged time period. Third, static assignment models will in general underestimate travel times in the presence of severe congestion. They will, among other things, by definition neglect the effect of spillback congestion. This means that during the design process, it may be better to focus on traffic decreases in known bottlenecks, rather than to focus on actual travel times from a static traffic model (although travel times need to be used as well).

**Try to get political and legal possibilities to adjust the system once it is in place.**

Even with careful planning, surprises are likely to appear when the system starts. In the best case, surprises are positive (in Stockholm, travel time improvements were larger than anticipated, for example). But there may be negative surprises as well: unexpected “rat-running”, for example. Because of this, it is good if one can get political and legal leeway to make minor adjustments to the system with a minimum of delay and hassle. Politically, this will be easier if goals are clearly formulated: if so, then it will be easier to see if they are met or not, and if not, the system can be changed. The legal problem may be harder to solve. In Sweden, for example, the charges (which are formally a state tax) have to be decided by parliamentary decisions, which requires a lot of time and political effort.

**There is a conflict between “effective” and “easily communicated” design, but erring towards too simple seems more common.**

Policymakers often stress that they want a design that is “simple to understand”. While it is an important consideration that the system must be sufficiently simple for the presumptive users to understand, policymakers often seem to underestimate people's
cognitive ability. The Singapore system and the US “value pricing” roads, for example, appear complex at first glance. The charge is finely differentiated by time and location, and on top of that may change several times each year. Despite its apparent complexity, it turns out that users are able to grasp and adapt to the system. Forcing the system to be too simple too early in the design process is likely to cause design restrictions that are difficult to solve. The reluctance of many politicians and planners to consider “too complicated systems” can lead to the point where the system becomes so simplified that it will not deliver the promised congestion reduction. This will not only be a waste of resources – it will also lead to low acceptability of the charges.

5.4 Information and technical design

In Stockholm, vehicles are registered automatically by cameras that photograph the number plates. During the trial, the main means of identification was transponders (“tag-and-beacons” or DSRC, dedicated short-range communication). When the charges were reintroduced, the automatic camera identification worked so well that it was decided to abolish the transponders. From the beginning, each daily charge was handled as one payment (one invoice per day), since it was believed that this was important to emphasize that each passage did in fact cause a payment. After some time, however, payments were aggregated to monthly invoices, without any consequences for traffic volumes. Obviously, this has reduced both operations costs and the hassle for users. In particular, the initial design with a very large number of invoices and very short time allowed for payment (initially five days) caused major problems for businesses.

Camera identification is used in London as well, although automatic identification rates are reported to be considerably lower in London. The main difference is that in London, it is the responsibility of the driver to pay the correct sum, and the system merely checks that vehicles passing through the charging zone have paid. In Stockholm, the system automatically issues the charge once it has registered a vehicle; if the vehicle is not identified, there is no need (or even possibility) to pay. The key argument in favour of “issuing/registration” rather than simply “checking” was that this made it easier to apply different charges according to the time of day.

The technical system was procured by the National Road Administration in a complicated tendering process eventually won by a consortium led by IBM. Information to the public was handled by the Road Administration. The system turned out to work very well from the start, both from a purely technical point of view and from an informational point of view; people knew what to do, how to pay, etc. Payment compliance was high, and the number of complaints was much lower than expected. On an average day in May 2006, 371,300 journeys took place over the charge cordon. 19 percent of passages were made with exempted vehicles (buses, taxis etc.), and an additional 9 percent were exempted due to the “Lidingö exception”. 267 500 passages were hence charged, resulting in 115 100 tax decisions (one tax decision was made per day and vehicle) and yielding toll revenues of SEK 3.2 million. Of the 115,100 daily tax decisions, only 100 were investigated by the Swedish Tax Agency and five were appealed. The Swedish Road Administration customer-service unit received on an average day in May 2,200 calls, in contrast to an expected number of 30,000 calls. Based on this, our assessment is that the system and the information generally worked well from a user’s perspective.

5.5 General advice on the technical system

The procurement and the design of the technical system in Stockholm is analysed Hamilton (2011) in offer some important lessons. Some of the general conclusions are that the political risk when deciding to implement congestion charging affects the cost.
of the system, that an agency spending to reduce that risk can be seen as rationally paying an insurance premium, and that public acceptance is affected both by the cost of the system and its functionality at launch. In addition, there are several design considerations, summarized below.

*Get the legal conditions clear early*

Early in the technical design process, one must know the legal conditions. For example, what is an acceptable proof that a vehicle has passed a gantry? What possibilities to appeal must exist? The answers to such questions will have important repercussions on the technical design, for example whether transponders can be the sole means of identification or not.

In Stockholm, a problem occurred that hopefully should be rare: midway in the procurement process, the legal status of the congestion charge changed from a “municipal environmental charge” to a state tax (a legal investigation concluded that it was illegal for a city to charge moving vehicles on existing roads). This had many effects, including that the responsibility of the procurement had to be changed from the City of Stockholm to the national government. This increased the cost for establishing the system considerably.

*Choose cost-efficient service level targets.*

Consider what the cost-efficient targets of service levels are, given what the goals of the system are and how different service levels affect the intended function of the system. Going from, say, 95% to 99% or from 99% to 99.9% on any given service level may be a significant cost driver. In Stockholm, the “uptime” of the system (measured as the share of “lane-minutes” the system was actually registering passages) was required to exceed 99.9%. To meet this high requirement, the prime contractor designed a system where (almost) every component was duplicated, spare parts were obtained in large quantities, trained staff was made available to do on-site service with short notice, and technical IT support was initially on standby 24/7. Obviously, this increased investment and operations costs. Moreover, it should be obvious that lowering the uptime requirement to, say, 95%, would not affect the traffic-reducing effect of the charges. After all, the travellers are making their travel decisions based on the fact that they are highly likely to have to pay if they go by car across the charging cordon. From this perspective, uptime requirements could have been relaxed substantially without losing any of the ultimate effect on the traffic situation. This illustrates the principle of having cost-efficiency in mind when formulating technical system requirements.

*Choose cost-efficient payment channels*

Each payment transaction comes at a cost, both in terms of convenience for the user and as a fee from the financial service provider. Hence, allowing for aggregated monthly payments rather than paying each passage individually will reduce operating costs. Cash over counter (in shops, for example) might be necessary for user acceptance, but it is probably the most expensive form of payment.

*Handling transponders is expensive*

Transponder (or “tag-and-beacon”) technology is efficient in many ways, not least because it allows complex charging structures and makes it easy for the driver. The production of many transponders may be a significant cost driver, though, but less well known is that it is often a major cost to administrate transponders. New cars need new transponders, cars change owners, and transponders are lost, stolen, and broken. In Norway, where over 40 different road toll schemes are in operation, transponders are used in some, while others are managed by manned tollbooths. And even there, where the comparison technology is highly manual, there is a slight productivity advantage for those not using transponders (Odeck, 2008). With today’s technology, cameras and automatic number plate recognition (ANPR) can potentially reach a very high
identification ratio, which offer ample competition for any transponder-based solution. The Stockholm system started out as a transponder-based system, with ANPR as an add-on for legal reasons, but has relied on ANPR exclusively since a few years.

*When doing a functional procurement, make sure to align cost and risk responsibilities*

In Stockholm, the congestion charging call centre was initially vastly oversized, which was a major cost driver initially. Part of the reason for this was that if the call centre would not meet its service quality targets (e.g. maximal answering times), then the prime contractor would be financially penalized, while it was buyer that carried the cost of call centre staff. Hence, risk and cost were borne by different parties15, and the contractor had no incentive to increase its own risk by cutting down on resources. If procuring a system as a function, one should make sure that the party carrying the risk is also the one taking the cost for risk mitigation, in all areas of the operation.

*High political risks will weaken the public negotiation position, and will increase costs by having the contractor require a risk premium*

In Stockholm, the stakes were high for almost all actors involved. Individual careers as well as the prosperity of private firms and political coalitions was at risk, or at least perceived as being so. This dominated the context in which the project was carried out, and it was under the influence of this risk environment that decisions were made. There were many unknown factors that were thought to kill the project on their own. Above all, if the system did not work, or was perceived not to work, right from the start, it would almost certainly be abolished immediately. This is at least a partial explanation of cost drivers such as the oversized call centre, the excessive service level requirements etc. It all goes back to the intense political pressure and high political stakes: the outcome of the next election would depend on the outcome of the trial, perhaps not only in the city but also on the national level. This meant that the public negotiation position was weak – the system had to work, and it had to be finished on time. Obviously, such a situation creates opportunities for a contractor to charge more money. For the contractors, a failure – even if it was not due to mistakes of their own – could be potentially disastrous for future business. This means that contractors will require a risk premium to even engage in the work of constructing the system. Hence, the lesson is that a stable political environment and ample time to plan and implement the system will keep costs down. Conversely, the extremely tense and uncertain political situation in Stockholm at the time, and the legal uncertainties that required a large number of changes and redesigns during the development process, all contributed to a substantially higher investment cost than would otherwise have been necessary. The main contractor argues that a similar system could now be built for half the cost or less.

6 **COST-BENEFIT ANALYSIS**

Even if it is well established that congestion pricing will yield a social surplus, it is not evident either that it will be enough to cover investment and operational costs, or that a real congestion pricing system, with all its practical and political limitations, will be socially profitable. Eliasson (2009) provides a cost-benefit analysis of the charges.

6.1 **Technology costs**

The total cost for the charging system was approximately SEK 1.9 billion, including operations costs during the first year of operation. 1.05 billion was incurred prior to

15 To be fair, it should be pointed out that this misalignment of costs and risks was an exception in the Stockholm procurement.
The Stockholm congestion charges

the start of operations. A significant part of this startup cost was costs for extensive testing. The system would only be operational for 7 months, making it absolutely necessary that everything worked right from the start. The startup cost also included, in addition to purely technical investments, system development in a wide sense, educating and training staff, testing, public information, etc., and certain other additional minor costs, such as those for traffic signals, and the services of the Swedish Enforcement Agency and the Swedish Tax Agency.

The rest of the costs (850 million SEK) were running costs and additional development costs during 2006. Far from all costs incurred during 2006 were pure running costs: the system was improved in several ways during the spring of 2006. Also included are the Swedish Road Administration's costs for closing down the system and evaluating the results during the second half of 2006. The investment and startup costs were considerably increased by the uncertain and heated political situation (as discussed above). The main contractor argues that a similar system could now be built for half the cost or less.

Actual running costs decreased significantly by each month of 2006, when it quickly became obvious that things in fact went better than planned: the number of complaints and legal actions were for example considerably lower than what had been anticipated, reducing costs for legal and tax administration. Further, the number of calls to the call center (the single biggest item in running costs) turned out to be around 1/20 than what had been anticipated – around 1500 calls per day instead of 30 000 per day. This meant that the call center was very much oversized, and during the spring, it was downsized – a considerable reduction of running costs. This means that investments costs could probably have been reduced quite substantially if the conditions (and not least the time constraints) had been different. This point may be especially important to note for other cities considering similar schemes. In 2006, the National Road Administration estimated future running costs to around 220 million SEK per year (around 25 million Euros). Since then, the responsibility for the system has been moved to the National Transport Agency system has been developed in various ways, in particular to incorporate the congestion charging system in Gothenburg which started in January 2013, and running costs have decreased further. At the time of writing (February 2014), the Transport Agency estimates yearly running costs to around 250 MSEK for both systems. The cost-benefit analysis presented here, however, uses the old estimate for running costs.

Yearly revenues for the first years of operation were a little more than 800 MSEK. Since a few exemptions have been abolished, revenues have increased somewhat to around 850 MSEK in 2013. Compared to the forecast before the start, revenues turned out to be around 14% less than predicted. 5% of this was due to the traffic reduction being larger than predicted. The major part of the revenue shortfall, though, was due to an underestimation of how many vehicles that would be exempted. This was purely due to a lack of data on the traffic composition, which should have been avoidable. Instead, neglecting to conduct reasonably simple traffic surveys beforehand meant that the exemptions became more costly than expected, and also created problems for the predicted cash flows.

6.2 Benefits

Most of the benefits are accessibility benefits, i.e. travel time savings and reduced travel time variability. From an economic point of view, it should be noted that it is these benefits that are translated into economic productivity and growth. For example, travel
time savings are partly converted into more working hours (increasing economic production) and partly to better matching on the labour market. These effects fall partly outside standard cost-benefit analysis, since they usually do not account for tax wedges and agglomeration effects (Anderstig, Berglund, Eliasson, Andersson, & Pyddoke, 2012).

Secondary benefits include reduced emissions and accidents. In addition, the revenue itself may generate benefits if they are used to reduce distortionary taxes or investments with positive benefit-cost ratios.

The CBA in Table 2 (Eliasson, 2009) shows that the Stockholm system yields a large social surplus\(^\text{16}\), well enough to cover both investment and operational costs. The annual social surplus is around 650 MSEK (after deducting operating costs). All major effects are primarily based on measurements, the most important sources being travel time and travel flow measurements.

<table>
<thead>
<tr>
<th>Loss/gain</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consumer surplus</strong></td>
<td></td>
</tr>
<tr>
<td>Shorter travel times</td>
<td>536</td>
</tr>
<tr>
<td>More reliable travel times</td>
<td>78</td>
</tr>
<tr>
<td>Loss for evicted car drivers, gain for new car drivers</td>
<td>-74</td>
</tr>
<tr>
<td>Paid congestion charges</td>
<td>-804</td>
</tr>
<tr>
<td>Increased transit crowding</td>
<td>-15</td>
</tr>
<tr>
<td><strong>Consumer surplus, total</strong></td>
<td>-279</td>
</tr>
<tr>
<td><strong>Externalities</strong></td>
<td></td>
</tr>
<tr>
<td>Reduced greenhouse gas emissions</td>
<td>64</td>
</tr>
<tr>
<td>Health and environmental effects</td>
<td>22</td>
</tr>
<tr>
<td>Increased traffic safety</td>
<td>125</td>
</tr>
<tr>
<td><strong>Externalities, total</strong></td>
<td>211</td>
</tr>
<tr>
<td><strong>Government costs and revenues</strong></td>
<td></td>
</tr>
<tr>
<td>Paid congestion charges</td>
<td>804</td>
</tr>
<tr>
<td>Increased public transit revenues</td>
<td>138</td>
</tr>
<tr>
<td>Decreased revenues from fuel taxes</td>
<td>-53</td>
</tr>
<tr>
<td>Increased public transport capacity</td>
<td>-64</td>
</tr>
<tr>
<td>Operational costs for charging system (incl. reinvestment and maintenance)</td>
<td>-220</td>
</tr>
<tr>
<td><strong>Government costs and revenues, total</strong></td>
<td>606</td>
</tr>
<tr>
<td><strong>Tax effects etc.</strong></td>
<td></td>
</tr>
<tr>
<td>Marginal cost of public funds</td>
<td>182</td>
</tr>
<tr>
<td>Correction for indirect taxes</td>
<td>-65</td>
</tr>
<tr>
<td><strong>Net social benefit, excl. investment costs</strong></td>
<td>654</td>
</tr>
</tbody>
</table>

\textit{Table 2. Costs and benefits of the charges, MSEK per year.}

\(^{16}\)The value of travel time was assumed to be 122 SEK/h per vehicle (65 SEK/h per person for private trips, 1.26 persons per private car and 190 SEK/h per person for business trips and distribution traffic). The value of reliability was assumed to be 98 SEK per hour of standard deviation (i.e. 80% of the value of travel time). CO\(_2\) emissions were valued at 1.50 SEK/kg. The Swedish value of a statistical life is 17.5 MSEK, the value of a severe injury 3.1 MSEK and a light injury 0.18 MSEK. The standard Swedish estimates of marginal cost of public funds and correction for indirect taxes are 1.3 and 1.23 respectively. The CBA is described in detail in Eliasson (2006, 2008).
The total start-up cost of the system was 1 900 MSEK, including information campaigns, extensive system tests and so on. Together with marginal cost of public funds and correction for indirect taxes, this gives a total social start-up cost of 2 900 MSEK. Hence, the start-up cost is “recouped” in terms of social benefits in about 4 years. The estimated yearly operational cost of the system (220 MSEK) includes not only running costs but also necessary reinvestments and maintenance such as replacement of cameras and other hardware.

Consumer surplus is negative, as expected, but the value of the time gains is high compared to the paid charges – time gains amount to almost 70% of the paid charges, which is very high compared to most theoretical or model-based studies. This is mainly due to "network effects", i.e. significant amounts of traffic that do not cross the cordon and hence do not pay any charge but still gain from the congestion reduction.

7 EQUITY EFFECTS

One of the major criticisms of congestion pricing is that it is regressive, since those with the least income have the hardest time affording to pay the toll. At the same time, there is a strong counter-argument to be made that congestion pricing can be regressive if the drivers are, by and large, those with the highest incomes, and those with lower incomes take transit. An advantage of the approach here is that we can examine the average effect either on income categories as a whole, or on income categories among drivers, transit riders, or even those who switch. There are a number of studies on the equity effects of the Stockholm charges (Eliasson & Levander, 2006; Eliasson & Mattsson, 2006; Franklin et al., 2010; Karlström & Franklin, 2009), yielding similar conclusions. The analysis and results presented here is taken from Franklin et al. (2010).

The average effects for income and toll effect sub-groups are shown in Table 3. Asterisks mark the income categories whose effects are significantly different from the pooled income categories. The most notable result here is actually the absence of significant differences. The only significant trend here is among the un-tolled, where the income categories are significantly different from the pooled income categories. There is also a trend of increasing benefit with income, but an important note here is that all five income categories did indeed see an average benefit. In other words, for the only subgroup where income played a significant role, everyone was better off anyway.

Table 3. Average Welfare Effects by Refund Scenario

<table>
<thead>
<tr>
<th>Group</th>
<th>Average Welfare Change, SEK/year/person</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Refunds</td>
</tr>
<tr>
<td>All</td>
<td>–78</td>
</tr>
<tr>
<td>By Income Category:</td>
<td></td>
</tr>
<tr>
<td>&lt; 25 000</td>
<td>–50*</td>
</tr>
<tr>
<td>25-40 000</td>
<td>–138</td>
</tr>
<tr>
<td>40-55 000</td>
<td>–39</td>
</tr>
<tr>
<td>55-70 000</td>
<td>–12</td>
</tr>
<tr>
<td>&gt; 70 000</td>
<td>–152</td>
</tr>
</tbody>
</table>

Note: * = significantly different (>95%) from the average welfare change for pooled groups.
Also, although there are not significant differences between income levels, there are some differences in the average numbers, but these do not follow a clear trend from low to high incomes. The reason for this is likely that other factors, such as distance travelled and amount of toll paid (based on time of day), and number of times crossing the toll cordon, varied more widely within each income categories than between them. It appears that individual circumstances have a greater effect on welfare than income level.

The use of revenues can be extremely important to the equity effects of a toll system, as evidenced by prior studies (Eliasson & Mattsson, 2006). Therefore, while we cannot be definitive about the redistributive effects of toll revenue disbursements, we can at least identify some bounds, by testing two theoretical scenarios for refunding toll revenues. The first is deliberately progressive, a lump sum refund to all in our study population; the second, deliberately regressive, with an across-the-board reduction in the income tax rate.

Adding the effects of each of these refund schemes to the total welfare effects found above, we arrive at the average welfare effects shown in Table 3. Importantly, all of the effects are now positive, and this holds true for both a lump sum and a tax reduction. Thus, we see an affirmation of one of the core arguments for congestion pricing: that by reducing externalities, we can see a net positive effect. Moreover, the empirical evidence supports a theoretical result from Small (1983) suggesting that all income levels could see a benefit on average, as long as revenues were appropriately returned.

In comparing the two refund scenarios, we see the expected result that the lump sum scenario is progressive, with the lowest income categories receiving the greatest benefit, while the tax reduction scenario is regressive, with the highest income categories gaining the most. Certainly, the refunds themselves for the two scenarios should be progressive and regressive, respectively, but what this tells us is that these original tendencies are not overwhelmed by the pattern of costs due to the congestion charging system itself, as represented by the "No Refund" scenario. Treating the two refund scenarios as bounds, the conclusion we can reach is that a wide range of uses for the toll revenues could maintain a positive average effect for all income categories, even if some uses would be more progressive than others.

7.1 Equity, fairness and winners/losers

The equity analysis above concentrates on a short-run winner/loser perspective, just as most equity analyses do. This perspective lacks two considerations, however: the relationship between objective equity and the perception of fairness, and the relationship between objective (monetary) effects and support for the charges, which is not a simple linear relationship.

First, a problem with the “winners/losers” perspective concerns the way this translates to the question of “fairness”. Often, if a system affects high-income groups more than low-income groups, it is claimed to be a “fair” system. Hence, “fairness” considerations – which are known to affect acceptability – are interpreted as a question of identifying “winners and losers”. In Stockholm, the equity effects were generally speaking progressive: high-income groups paid more than low-income groups, men paid more than women, employed more than unemployed etc. But once the charges are in place, and the short-term winner/loser perspective fades, another perspective becomes more important: what price is actually “fair” to charge for a car trip? From this perspective, it is “fair” that one pays more to drive on a congested road or to cause emissions in densely populated areas – irrespective of income or place of residence, or what a
The hypothetical travel pattern would have been without the charges. This means that a system needs to be perceived as “fair” in this sense: it needs to be consistent with its stated objective. In Stockholm, one of the most common objections to the system nowadays is that traffic within the cordon is not charged. Although there are two good answers to this (the congestion is mainly located on the arterials along the cordon; most of the traffic inside the cordon crosses the cordon at some point on the trip), this shows how the debate has moved from “who wins/loses” to “what’s fair relative to the objectives”.

Second, according to standard transport-economic theory underlying classical equity analysis, most motorists would not think that the time saved was worth the charges they had to pay. Theoretically, the income from the charges is sufficient to compensate the losers, so the standard recommendation in the acceptance literature is that congestion charges must be part of a “package”, within which it is clear how the income is going to be spent for the advantage of the general public, if it is going to have any chance of being accepted. But the standard theory neglects three important considerations. First: network effects. Since queues propagate “upstream”, even those not going through the actual bottleneck will suffer from queues. Pricing traffic in the bottleneck to reduce queues, all upstream traffic will benefit – not only drivers actually paying the charge. Second: the effect on the urban environment. Typically, standard analysis of congestion charges takes no account of effects for pedestrians or cyclists, or the effect on the perceived urban environment. Third: the self-selection effect on trips and on the value of time. Congestion charges will tend to “sort” trips such that trips with high value will stay on the road (and enjoy time benefits), while low-valued ones will be priced off. Not taking this phenomenon into account will underestimate the value of the time benefits. From an acceptance perspective, the important point is that individuals can belong to different valuation “groups” on different days, or different journeys.

Another potentially important observation is that car drivers apparently changed behaviour without even noticing it. When motorists were asked if the congestion charging had made them change their travelling habits, there were too few answering “yes” to correspond with the actual reduction in measured traffic volumes.

One important reason is that travel patterns are much more variable than most people are aware of. It is simply not the same drivers paying the charge each day. On the contrary: a large majority of the drivers are “occasional” drivers, who pay the charge a one or a few times a week or even more seldom. Figure 11 shows the distribution of passage frequency groups during a two week period.
The Stockholm congestion charges

Figure 11. Drivers passing the cordon a given day, grouped by how often the pass cordon during a two-week period.

This means that most car owners are affected somewhat, but very few are affected a lot. Figure 12 illustrates this. Any given day, around 5% of private trips in the county are affected by the charges. Over two weeks\textsuperscript{17}, however, 43% of private cars in the county will pay at least once. However, over this two week period, only around 2% of car owners will pay an amount equivalent to the average return fare each day (270 SEK for 2 weeks), and only about 0.4% will pay an amount equivalent to rush-hour return fare each day (400 SEK per 2 weeks).

Figure 12. Share of private cars who pay various total charge amounts during a two-week period.

8 ARE THE RESULTS TRANSFERABLE?

In many respects, each city is unique. Introducing a charging scheme similar to the Stockholm system is by no means a guarantee for achieving the same congestion reductions as in Stockholm. But in a more fundamental way, it can be argued that the experiences are indeed transferable to other cities, in the sense that the charges actually affected car drivers in the way that had been predicted by transport models.

\textsuperscript{17} Due to integrity legislation, statistics can only be aggregated over individuals over two-week periods, not longer.
During the design process, three independent transport models yielded very similar results regarding the anticipated effect of the charges. This was despite the fact the three models were estimated and calibrated on different data sets and used different methodologies (although all of them were nested logit models linked to a static equilibrium model). Despite this, the forecasts were deemed to be unrealistic – the charges could not possibly affect traffic that much, it was believed. The lesson here is twofold: a good transport model is an invaluable tool for designing an efficient charging scheme, and the result is likely to be correct (after accounting for known model limitations such as underestimating travel times in severe congestion and many models’ inability to account for changes in departure times). The “transferable lesson” is hence that a well-designed charging scheme that seems to work in a traffic model is actually also likely to work in reality. A very common reaction from other cities is that “charges may have worked in Stockholm and London, but the particular situation in our city means that charges won’t work here, despite our transport model saying otherwise”. The same claim used to be common in Stockholm, in fact. The Stockholm experience is that transport models can actually be trusted.

The success of the Stockholm charges in reducing congestion and achieving public and political support has attracted great interest from cities around the world. A natural question is whether the positive results are transferable – if congestion charges would work just as well in other cities. Judging from the authors’ experience as advisors to cities around the world, a common reaction is “it would not work in our city”. Of course, all cities have their particular characteristics and local conditions, so a copy of the Stockholm system would not give exactly the same effect in another city.

But in a more fundamental way, it can be argued that the experiences are indeed transferable to other cities, in the sense that the charges actually affected car drivers in the way that had been predicted by transport models. The conclusion that a transport model was able to predict demand responses with good-enough accuracy leads to a more qualified answer to the question of transferability: if a congestion charging system is predicted to “work” in a given city – that is, reduce peak traffic in bottlenecks without unacceptably adverse side-effects or having to use unacceptably high charge levels – then that is likely to be true in reality as well, not just in the model. It should be noted, however, that the beneficial effects on congestion and travel times are likely to be underestimated by static network models. During the Stockholm design process, three independent transport models yielded very similar results regarding the anticipated effect of the charges. Despite this, the forecasts were deemed to be unrealistic – the charges could not possibly affect traffic that much, it was believed.

A related question is whether the effectiveness of congestion charges are highly dependent on the specific features of the land use and transport system, implying that the good experiences of the congestion charging systems implemented in Stockholm or London would not be transferable to cities with topological and demographical conditions, availability and attractiveness of non-charged routes and public transport provision or sizes.

This question is explored in detail in a study by Börjesson, Brundell-Freij and Eliasson (2014). The main conclusion is that although the social benefit of a given charging system (including the size of the charge) is considerably and non-linearly dependent on initial congestion levels, traffic effects and adaptations are surprisingly stable across different transport systems. Specifically, the level of public transport provision has only small effects on baseline congestion, and therefore on the total benefit of the charges. Interestingly, adaptation cost, traffic reduction across the cordon and the share of drivers priced off the road diverting to public transport are also surprisingly insensitive
to the level of public transport provision, contrary to the common argument that public transport provision is crucial for effectiveness and efficiency. Drivers can adapt in many different ways except for switching to public transport, which explains the robustness of the results.

8.1 Why a “success”?

The newspaper quote in the introduction calling the trial a “success” was fairly typical. Even if certainly not everybody was in favour of the charges, the change in general opinion as reflected in polls and media seemed to justify such statements. But what were the key factors behind this? There is no conclusive answer to this question, but five main reasons were often mentioned by four key people involved in the trial – Gunnar Söderholm, head of the Congestion Charging Office responsible for (among other things) evaluation and information; Birger Höök, head of the Congestion Charging Unit at the Road Administration, responsible for the technical system and information regarding payment and technology; Gunnar Johansson, head of the IBM-led consortium that developed and operated the technical system; and the author, who was responsible for the early system design and chairman of the expert panel summarising and scrutinising the large evaluation package.

- **The technical system worked.** That the system worked from the start was of course a key factor. The number of misidentifications was extremely low, and from users’ perspectives, everything worked seamlessly. Further, the Road Administration made great efforts to develop a customer-friendly system.

- **The information campaign had worked.** Apparently, people knew what to do. Anticipated problems with people who did not know that they should pay, or did not know how to pay, did not materialise. Moreover, the anticipated problem of protests in the form of large number of court appeals or refusals to pay was never a problem. in spite of a lot of talk before the trial about “civil disobedience” in the form of refusal to pay or appealing to court.

- **Visible congestion reduction.** The improvements in travel times and the urban environment were visible right from the start. The astonishment of seeing almost empty streets during rush hours, in particular during the first months, cannot be stressed enough. After that, the potency of road pricing had been overwhelmingly proved, and the negative arguments shifted from “it won’t work” to other, often more constructive, arguments.

- **Extensive and scientific evaluation.** Even if effects were visible, one should stress the importance of being able to supply media with hard figures about the reduction of traffic volumes and congestion. Especially when the debate recovered somewhat from the initial shock caused by the enormous initial effects, it was extremely important to have professional, independent researchers and experts, coming from different backgrounds and organisations, being able to explain and evaluate what was happening. The size of the evaluation was itself an important factor: so many experts and researchers were involved in one way or another that it was impossible to wave it away.

- **Clear objectives.** The system had clear and measurable objectives – reducing congestion and improving the environment in the inner city – and the system was visibly designed with these objectives in mind. Moreover, the objectives were fulfilled.

9 REFERENCES

The Stockholm congestion charges


