SO YOU’RE CONSIDERING INTRODUCING CONGESTION CHARGING?
HERE’S WHAT YOU NEED TO KNOW

Jonas ELIASSON, Royal Institute of Technology, Stockholm
So you’re considering introducing congestion charging?  
Here’s what you need to know

An FAQ based on Stockholm’s experiences

Jonas Eliasson
Centre for Transport Studies
Royal Institute of Technology
Stockholm
Sweden

January 2010

The views expressed in this paper are the author’s, and do not necessarily represent the opinions of the Royal Institute of Technology, Stockholm, the International Transport Forum or the OECD.
So you’re considering introducing congestion charging?  
Here’s what you need to know

*An FAQ based on Stockholm experiences.*

Jonas Eliasson  
Centre for Transport Studies, Royal Institute of Technology,  
100 44 Stockholm, Sweden  
jonas.eliasson@infra.kth.se

This paper is a draft prepared for the OECD Round Table on congestion charging,  
February 4-5. It is accompanied by two other papers, covering the experiences from  
Stockholm in considerably more detail. These two papers are

Carl Hamilton: Revisiting the cost of the Stockholm congestion charging system  
(*Discussion Paper 2010-5*);  
Jonas Eliasson: Lessons from the Stockholm congestion charging trial (published in  
*Transport Policy, 2008*).

**Introduction**

Congestion pricing has been advocated for a long time as an efficient means to reduce road  
congestion, without much success in practice. In the last few years, however, congestion  
pricing has been introduced in various forms, with London and Stockholm attracting the most  
international attention, starting in 2003 and 2006, respectively.

The Stockholm experiences have attracted widespread attention from many cities that  
consider introducing similar schemes, and the people involved in the development and  
evaluation of the Stockholm system (including the author, who was responsible for the  
design of the charging system and forecasting its effects, and subsequently chaired the  
Expert Evaluation Panel) have acted as advisors to many policymakers from cities and  
countries all around the world. This paper is based on these experiences. A number of  
questions repeatedly surface in discussions with policymakers, politicians and planners. A  
number of facts and insights repeatedly turn out to be especially relevant and interesting, and  
to a certain extent surprising. The purpose of this paper is to summarise these questions,  
facts and insights, for the benefit of cities considering congestion charging schemes.

The paper draws from already published material (see the Reference section – section 6). In  
fact, a reader already familiar with the congestion charging literature will find few completely  
new findings or insights. The contribution of the paper is rather the selection of the most  
relevant, interesting, important and sometimes surprising facts, insights, findings, advice and  
conclusions for policymakers, out of a vast literature on congestion charging in general and  
Stockholm in particular.
It should be stressed that the paper is not meant to be a comprehensive survey of theory or experience. In fact, the aim has been to keep the paper as short as possible, and to cut as many things as possible, in order to focus on the most relevant ones: the key facts and advice that should be emphasized to policymakers as early as possible.

There are three areas that are necessary to consider for a successful congestion charging implementation. First, the charging structure needs to be well designed, in order to achieve large social benefits. This is the topic of section 3. Second, investment and operational costs must be kept low. This is the topic of section 4. Third, one must get public acceptance for the system. This is the topic of section 5. The paper starts (section 2) with some fundamental facts about congestion charges – why they are needed, how they work and their general effects.

**Basic Facts: Why and how it works**

*This section covers issues related to why congestion charges are needed, what they can and cannot do, and what effects charges have on traffic.*

**Congestion cannot be eliminated by investments alone.**

It is a well established fact, both theoretically and empirically, that investments in road and/or public transit is not sufficient to eliminate road congestion in the cores of large cities. There are several reasons for this: two of the most important from a practical point of view are the inevitable eventual scarcity of urban land and public resources.

**Congestion charging will not solve everything.**

Introducing congestion charging will 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 (both roads and public transit). Obviously, cities are different as to what investments are the most cost-efficient and the most needed.

**Only introduce congestion charges when they are needed.**

This may seem obvious – but in fact, one is sometimes confronted with cities with virtually no road congestion that are nevertheless considering “congestion charges” (there are several current Swedish examples of this). The purpose may instead be to raise revenues or reduce traffic emissions. Generally speaking, there are usually more cost-efficient ways to do such things than introducing congestion charges – not least because a charging system is typically a fairly expensive investment.

1. It should be pointed out, however, that road user charging may be a cost-efficient way to improve air quality in urban centres.
Car drivers are cost sensitive.

At least sufficiently many of them. This means that increasing the cost to drive a car 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 – 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.

There are many ways to adapt.

Not all car drivers that change due to the charges will switch to transit, nor to other routes or time periods. Route and mode changes are far from the only adaptation strategies. Trips, especially discretionary trips (shopping, leisure etc.) 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 change”. Especially for discretionary trips, 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 oversimplistic: there are many more ways to adapt than changing mode or route.

Traffic isn’t just work trips.

Work trips only make up a fraction of car traffic – a typical figure could be 40%, with the rest being discretionary trips and professional traffic (where a typical figure could be 15%). 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.

Many car drivers won’t even know that they have adapted.

In fact, many car drivers will not even know if or in what way they adapted. 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. 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 – they often won’t know themselves. Moreover, there are many other changing processes going on. People move and change jobs, for example. After just a few years, it may pointless to ask how a given person has “adapted” – because the entire situation where travel choices are made has changed.
Retail effects are generally small

Fear for adverse impact on retail inside a cordon is common. Large efforts have been made to track such effects, only to conclude that they are very small or non-existent. There may be effects on particular stores, especially if they lie close 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.

Getting large benefits: Efficient charge design

This section covers issues related to the design of the charging system – decisions such as where and when to levy charges and how to forecast effects.

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 topography 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 interactive 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 “sub-optimal” – it may likely cause more problems than before.

You need a good transport model.

(This section is important for experts designing a charging system, but policymakers and general planners can skip this.)

Most transport models are constructed for other purposes than modeling 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 topography, 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 is 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 involves 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 fairly often. Despite this, 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 (we return to this below).

Getting low costs: Efficient procurement and technology

This section covers issues related to investment and operations costs of the technical system, in particular achieving a procurement process that keeps these costs down. Designing a cost-efficient technical system amounts to more than just getting a system that identifies cars: the difficult part may be to construct cost-efficient payment channels, sorting out the legal status of the charges etc.

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.

Consider the cost-efficiency of 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 were 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 decisions as to whether to drive or not based on the fact that they are highly likely to have to pay if they do so. From this perspective, availability requirements could have been relaxed substantially, and thereby also system build costs, 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 offers 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 call centre, answering questions about the charges, 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 parties2, 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.

---

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

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 center, 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.

**Getting public acceptance**

*This section covers issues related to public and political acceptability. It draws from the conclusions in Eliasson and Jonsson (2009).*

**Acceptability decreases with detail but increases with familiarity**

Support for congestion charging often follows a typical pattern. The figure below shows the principle.

![Graph showing acceptability decreasing with detail and increasing with familiarity.](image)

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 with familiarity with a real system. One
oft-quoted reason in Stockholm was that the positive effects on road congestion and urban environment were much larger than most people expected. A second, also oft-quoted reason is that the public fear that travel costs will increase more and/or their travel patterns will have to change more than is actually the case. Once the charges are in place, many people may realise that the charges do not in fact affect them as much as they had thought. A third reason is the psychological effect known as cognitive dissonance, a phenomenon that can be simply summarised as “accept the unavoidable”. In other words, once the charges are in place, it is less worthwhile to spend energy on opposing them. A fourth reason may be decreased reluctance towards pricing a previously unpriced good. There is evidence that people in many cases do not like prices as an allocation mechanism. But once familiar with the thought 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.

**Plan the political process accordingly.**

The general acceptability curve above has implications for the political process. For a successful implementation, one needs to avoid having elections when support for the charges can be expected to be at its lowest. In London, the election was held before the details were completely worked out; in Stockholm, a referendum was held once the charges had been in place for a little more than half a year (in fact, support then continued to increase, so a later referendum would probably have been preferable from the point of view of charging supporters). This can be contrasted with Manchester and Edinburgh, where detailed proposals were worked out (support fell, just as could have been expected) but charges were not implemented before a referendum was held. In both cases, the referenda rejected the charges.

**The system has got to deliver benefits.**

In Stockholm, the most important factor explaining attitudes to the charges turns out to be the perceived effect of the charges – in particular less congestion. People agreeing that the charges have had positive effects are much more likely to support them – which is of course expected. 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 use 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.

**“Branding” matters.**

In Stockholm, the charges were to a certain extent marketed as “environmental” charges. The charges did certainly have environmental effects (in particular on emissions in the inner city), but this effect was dwarfed by the very large effects in terms of congestion reduction. It turns out, however, that voters’ environmental concerns was an important factor for the acceptability of the charges. (Interestingly, data shows that it seems as if it is not environmental behaviour per se that is important, but the self-image of being an environmentally concerned person.) 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 costs and benefits can also strongly affect acceptability. Hence, the “branding” of the charges matters — how they are marketed, explained and perceived. In Stockholm, “re-labeling” 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 branded as “congestion charges” system should for example not levy charges where or when there is no congestion.

Car dependence and transit satisfaction seem to matter less than many believe.

Analysis of attitude data from Stockholm shows that, as expected, car dependence decreases support for the charges while transit satisfaction increases it. Somewhat surprisingly, it is a less important factor than environmental attitudes and perceived effects of the charges. Econometric simulations on this data shows that even with a fairly high car dependency, support stays relatively high. Moreover, residential location zone hardly matter once the other factors are controlled for. This is surprising since the consequences of the charges in terms of e.g. tolls paid differ quite a lot depending on residential area. Other analyses, though, have shown significant impacts on voting behaviour from the changes in travel times and travel costs of voters’ residential zones. The evidence is not conclusive, but it seems that “self-interest” variables such as car and transit use matter less than one might think initially.

“Fairness” can mean many things.

The “fairness” or “equity” issue is always important. Initially, the dominating perspective is often “before-after”. In other words, how do travel costs (and perhaps times) change for different groups – high- vs. low-income, 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.

“Winners” and “losers” gets increasingly difficult to identify over time.

As discussed above, travel patterns are not static. Not even the context - workplace and residence location, for example – is constant over the span of a few years. This means that over time, identifying “winners” and “losers” will become increasingly meaningless. The charges will have changed from being an “external shock” to being a factor considered when making choices of workplaces, travel modes etc. Hence, the question of who “wins” and “loses” is only relevant in the short term.

Power issues may be decisive for political acceptability.

From politicians’ 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, the state “matches” income from regional charges with national funding. A recent trend in Sweden is that regional funding is “leveraged” with national funding. The regional funding often comes from charges (called “congestion” charges, which in most cases is a terrible misnomer). This has made “congestion” charges much more popular among politicians, which shows the importance of institutional context and incentives.

References

This paper draws heavily on already published material. In this draft version, references have not been cited in the text. Instead, we will list the most important published sources of the statements made in the paper.


