

Accuracy of congestion pricing forecasts

Jonas Eliasson, KTH Royal Institute of Technology
Dirk van Amelsfort, WSP Analysis & Strategy
Maria Börjesson, KTH Royal Institute of Technology
Karin Brundell-Freij, WSP Analysis & Strategy
Leonid Engelson, WSP Analysis & Strategy *and* Royal Institute of Technology

CTS Working Paper 2012:31

Abstract

This paper compares forecasted effects of the Stockholm congestion charges with actual outcomes. The most important concerns during the design of the congestion charging scheme were the traffic reduction in bottlenecks, the increase in public transport ridership, the decrease of vehicle kilometres in the city centre, and potential traffic effects on circumferential roads. Comparisons of forecasts and actual outcomes show that the transport model predicted all of these factors well enough to allow planners to draw correct conclusions regarding the design and preparations for the scheme. The one major shortcoming was that the static assignment network model was unable to predict the substantial reductions of queuing times. We conclude that the transport model worked well enough to be useful as decision support, performing considerably better than unaided “experts’ judgments”, but that results must be interpreted taking the model’s limitations into account. The positive experiences from the Stockholm congestion charges hence seem to be transferable to other cities in the sense that if a charging system is forecasted to have beneficial effects on congestion, then this is most likely true.

Keywords: Forecast accuracy, congestion pricing, model validation, policy transfer.

JEL Codes: R41, R48.

1 INTRODUCTION

Congestion charging is often advocated as an efficient way to combat road congestion. Many cities around the world have considered it and quite a few have introduced it, either as systems covering the city centre (e.g. London, Stockholm, Singapore) or as systems covering single links or lanes (e.g. Melbourne, Toronto, the “value pricing” systems in the United States).

Even if the general principle of marginal cost pricing is simple enough, designing a real-world system with all its necessary simplifications is a difficult task. Since investment costs and political stakes are usually high, there may only be one chance to get the design “right”, in the sense that the system delivers tangible benefits without creating new problems such as substantial traffic rerouting or transit crowding. The complexity of the transport system and the large number of possible design variables make the use of transport models essential. This raises the question: are transport models sufficiently reliable to be used as decision support when designing congestion charging systems and deciding whether to implement such a system?

We explore this question by comparing predicted and actual effects of the Stockholm congestion charging system. The central question of the paper is whether the transport model allowed correct conclusions to be drawn regarding the design and preparation of the charging scheme. Hence, the comparison is structured according to the issues that were the main questions during the design and preparation process, such as whether the reduction of car traffic would meet the target, whether there would be capacity problems in the public transport system, whether traffic would decrease within the inner city, and whether congestion on circumferential roads would increase. The question whether the model was “good enough” can be interpreted as “if planners had had access to a perfect model, able to perfectly predict the effects of the charges – had the system been designed differently in any respect, or had different preparations been made?”.

However, the purpose of this paper goes beyond mere model validation and development. From a policy perspective, the crucial question is whether the positive Stockholm experiences – where the charges resulted in substantial congestion reductions – are transferable to other cities. Obviously, introducing a copy of the Stockholm system in another city would not give precisely the same effects. But if transport models can be trusted to predict the effects of congestion charging systems reliably enough, then a best-practice transport model for another city should be able to predict what effects congestion charges would have in *that* city. There are hosts of model-based analyses of congestion charging schemes for cities around the world (e.g. (de Palma, Lindsey, & Niskanen, 2006), (Santos, Newbery, & Rojey, 2001), (Rich & Nielsen, 2007), (Eliasson & Mattsson, 2006), (Fridstrøm, Minken, Moilanen, Shepherd, & Vold, 2000), (Kickhöfer, Zilske, & Nagel, 2010)), but predictions are often met with scepticism from decision-makers and the public. Before the charges, Stockholm was no exception: predictions that relatively low charges would lead to substantial reductions of congestion were widely disregarded. The ulterior motive of this model validation exercise is to contribute to this debate, and enable cities to make informed assessments of the potential benefits of congestion charges and how a system should be designed.

Despite transport models’ importance as decision support, there are not many published studies providing detailed analyses of the validity of transport model predictions. Moreover, virtually all such studies deal with physical investments such as new roads or railways. An exception is (Leape, 2006), who in a review of the London experiences notes that the forecast of a 10-15% reduction of vehicle miles in the

charged area turned out to be accurate. The present paper hence seems to be the first detailed study of a transport model's ability to forecast the effects of an urban congestion charging system. (Bain, 2009), (Bain, 2011) and (Li & Hensher, 2010) compare forecasts and outcomes for a large number of toll roads, finding that forecasts tend to overpredict actual patronage. This seems to be less due to shortcomings of models than to various forms of optimism bias, including psychological effects and winner's curse (the most overoptimistic bidder wins a toll road concession).

(Flyvbjerg, Holm, & Buhl, 2005) analyse a large collection of forecasts and outcomes for rail and road investments, finding systematic overpredictions of demand for new infrastructure projects, in particular for rail projects. They conclude that the main culprit seems to be incentives to strategically misrepresent demand forecasts caused by institutional structures, such as competition for public funds, rather than poor prediction techniques. Also (Van Wee, 2007) finds poor quality of transport demand forecasts, especially for rail projects, and that it is not so much shortcomings of forecasting models that is the problem but rather the opportunities and incentives for strategic behaviour for some actors offered by the formal and informal structures.

An informative case study of such strategic misrepresentation is provided by (Kain, 1990), relating the story of how demand forecasts for a railway were exaggerated by tinkering with the land use scenarios used in the transport model. However, the fact that these studies point to strategic mishandling of models as the main source of errors mean that they do not provide much guidance as to models' actual forecasting capabilities, although (Bain, 2009) and (Bain, 2011) also quote overestimation of drivers' willingness-to-pay as one reason for overpredicting traffic on toll roads. Contrary to the studies above, (Parthasarathi & Levinson, 2010) find that forecasts tend to underpredict actual traffic, in a study of a large number of Minnesota roads. The reasons for underprediction are unclear, but erroneous scenario assumptions seem to be a major source of error. Turning to case studies of specific investments, (Pedersen, Denver, Daly, & Rohr, 2001) compare forecasts and outcomes for the Danish Great Belt link, noting a slight underprediction of trips. (Anguera, 2005) study the Channel Tunnel, showing that the substantial overprediction of traffic and revenues was due to neglecting competitive price response of Channel ferries combined with massive overprediction of total market growth. (Petersen, 2010) compares forecasts for the Danish-Swedish Öresund bridge with actual outcome, noting that political visions seem to have affected what forecasts to use.

Model validation often becomes difficult due to long time lags between forecast and start of operation. Input parameters such as population, land use, fuel prices, fares, and levels of service may be different in reality than was assumed in the forecast. Moreover, it is often difficult to get detailed information of outcomes, forecasts and underlying scenario assumptions (Parthasarathi & Levinson, 2010). As a contrast, the introduction of the Stockholm congestion charging system provides an unusually good opportunity for model validation. First, the effects of the charges were measured in an extensive evaluation program, gathering all sorts of data on traffic flows, travel times, travel patterns etc. Secondly, the forecast was made shortly before the introduction of the charges, eliminating the problem of errors in input parameters or scenario assumptions. Thirdly, we know precisely how the model was used as decision support: what questions it was used to analyse, how results were interpreted and what conclusions were drawn. This makes it possible to focus on those model aspects and outputs that were most important from the point of view of the users. For the reasons stated above this paper evaluates the model forecast against the short-term effect of the congestion charges. The actual effects of the charges are taken to be the differences between measurements from the spring of 2005 (without charges) and measurements

from the spring of 2006 (with charges). Previous analyses have shown that other factors than the charges made very little difference to the change in traffic levels between 2005 and 2006 (Eliasson, 2009a). Moreover, the effects of the charges seem to have remained virtually unchanged over time (Börjesson, Eliasson, Hugosson, & Brundell-Freij, 2012). Hence, no rebound effect is observed in the sense that new car traffic has been attracted in the longer run.

The paper is organized as follows. Section 2 describes the Stockholm congestion charging system and the transport model used for the forecasts. Section 3 compares forecasts with outcomes. Section 4 summarizes the findings and concludes.

2 BACKGROUND

2.1 The Stockholm congestion charging system

The City of Stockholm has around 0.8 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, which has an area of around 30 km². Approximately 320,000 persons are employed in the inner city, of which more than two-thirds are commuting from outside the cordon. Before the congestion charges (2005), the cordon around the inner city was crossed by around 530,000 vehicles and 800,000 transit passengers each day.

The Stockholm congestion charging system consists of a toll cordon around the inner city, thereby reducing traffic through the bottlenecks located at the arterials leading into the inner city. The cost¹ of passing the cordon (in any direction) on weekdays is € 2 during peak periods (7:30-8:30, 16:00-17:30), € 1.5 during the shoulders of the peaks (30 minutes before and after the peak periods) and € 1 during the rest of the period 6.30-18.30. The total charge per day is capped at € 6. Various exemptions (e.g. buses, alternative-fuel cars and traffic between the island of Lidingö and the rest of the county) mean that about 30% of the passages are free of charge. There is no congestion tax levied on vehicles driving on the E4/E20 (the Essinge bypass) that goes through the west part of the toll zone past the inner city. 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 bypass. The opposition from the surrounding municipalities was so strong, however, that the politicians of the City of Stockholm decided that the Essinge bypass should be free of charge.

¹ Assuming 10 SEK = 1 €.

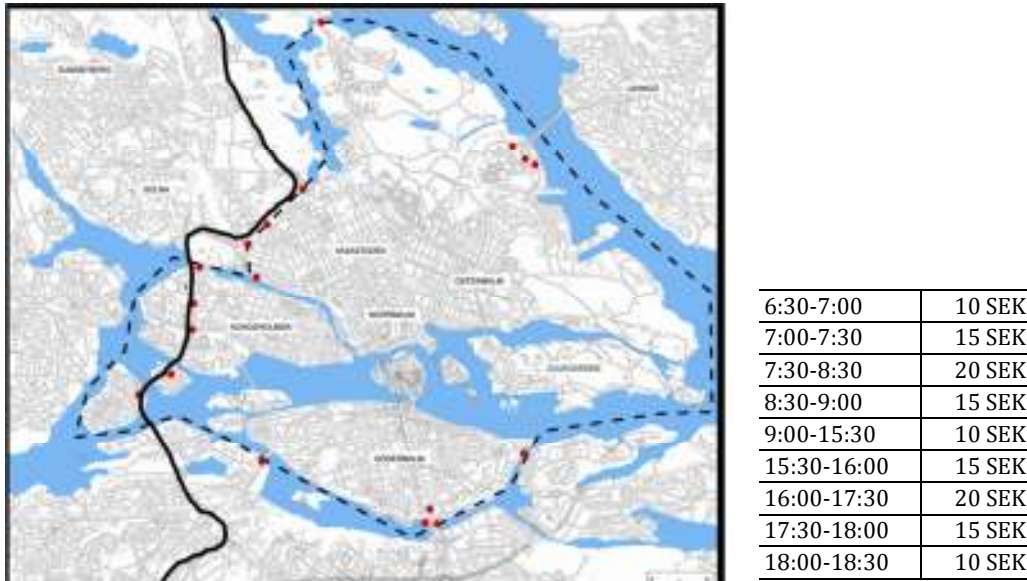


Figure 1. The charged area. The dashed line is the charging cordon, the dots are charging points and the solid line is the non-charged Essinge bypass.

The system was introduced on a trial basis during the period January 3 – July 31 2006. The trial period was followed by a referendum resulting in a majority for keeping the charges. The congestion charges were reintroduced in August 2007.

An overview of the effects of the charges can be found in (Eliasson, Hultkrantz, Nerhagen, & Rosqvist, 2009), (Börjesson et al., 2012) and (Eliasson, 2008a), where the latter also discusses the main lessons from the trial in terms of design, effects, acceptability and political process. (Eliasson, 2009a) provides a cost-benefit analysis of the congestion charging system. Experiences from the design and evaluation processes are described in (Eliasson, 2009b). (Karlström & Franklin, 2009) and (Franklin, Eliasson, & Karlström, 2010) analyse behavioural responses and equity effects. A detailed account of the political process can be found in (Gullberg & Isaksson, 2009). (Brundell-Freij, Jonsson, & Källström, 2009) and (Eliasson & Jonsson, 2011) provide analyses of the developments of the public attitudes up to late 2007, focusing on what factors explain differences and changes in public acceptability.

2.2 A description of the transport model

The model used for forecasting the effects during the scheme design process was the national transport model *Sampers*. *Sampers* consists of a nested logit demand model linked to the network assignment model EMME/2. The demand model was estimated on data from the Swedish national travel survey Riks-RVU 1994-2001.

The demand model is tour-based, comprising choices of trip frequency, destination and mode, with separate sub-models for seven trip purposes (commute, service, shopping, leisure, visit, kiss-and-ride and employer's business), comprising five modes (car as driver, car as passenger, public transport, walk, and cycle). Demand matrices are split into four time-of-day matrices (morning, mid-day, afternoon and night) using constant time-of-day shares (i.e. departure time choices are not modelled). Demand matrices are then assigned to road and transit networks using the network assignment model. Road assignment is carried out using multi-class assignment with five value-of-time classes. Professional traffic (freight, distribution, garbage collection etc.) is represented by a fixed matrix, meaning only route choices change in the analyses. A few other fixed matrices are also added before the assignment step, e.g. long-distance traffic to and from the Stockholm ports. The travel times and costs from the assignment model are

then fed back into the travel demand models in an iterative feedback loop until convergence is reached.

The work trip demand model is doubly constrained, meaning that the total number of work trip ends per zone is constrained to be equal to the number of workplaces in the zone. The model predicted that work trip destinations would adapt as a response to the charges, despite that adaptation by switching workplace is unlikely to happen in the short run. The best solution would have been to constrain the work trip O-D matrix to be constant in the forecast of the charges, but this turned out to be unfeasibly complicated due to the structure of the nested logit model (with destination choice below mode choice). Instead, model results had to be reinterpreted and adjusted to account for this.

Travel times in static assignment models do not reflect effects of severe congestion, such as spillback queues, dynamic congestion and blocking of intersections and ramps. 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. It was unclear, however, exactly how the anticipated traffic reduction would affect travel times, and how this would affect the accuracy of the demand forecasts.

The Sampers model can be said to constitute “best practice” for large-scale transport models. Its structure – logit demand models linked to a network assignment model – is similar to most operational, large-scale multi-modal transport models. It has been carefully estimated and calibrated using state-of-the-art methods, and validated against in-sample data. However, it does not model changes in departure times, nor does it include novel features such as activity-based travel demand or simulation-based network assignment. Since the model was not specifically constructed to model congestion charges, some of the system features were difficult or impossible to capture properly in the model, in particular the time-differentiation of the charges and the cap on total accumulated charges per day.

3 COMPARING PREDICTIONS WITH OUTCOMES

The actual effects of the charges are, as mentioned, taken to be the differences between measurements from the spring of 2005 (without charges) and measurements from the spring of 2006 (with charges). Since there are significant seasonal variations in traffic flows, it is essential that the comparison refers to the same time-of-year both years.

Transport models produce huge amounts of output data, which means that a strategy is needed to structure the comparison between forecasts and outcomes. We have chosen to structure the comparison according to the main questions asked during the design and preparation of the congestion charging system:

- Would the charges reduce traffic across the charging cordon (where the main congestion problems are located)? How large, if any, would the traffic reduction be?
- How would the “disappearing” car drivers adapt?
- How much would public transport trips increase? Capacity increases may be needed to reduce potential crowding problems.

- Would congestion problems increase on links outside the charged area? Re-routing traffic may be a substantial problem. Even small increases in congestion on alternative routes could cause considerable political problems.
- Would there be any positive local environmental effects by reduced emissions? The charges were to a large extent marketed as “environmental charges”, and hence environmental effects needed to be quantified.
- How should the charges be differentiated across time, and how would the charges affect the timing of trips? In particular, would congestion problems appear before and after the charged period, or before and after the highest charges?
- Would congestion be reduced, and if so, by how much? Would the traffic reduction be enough to make a substantial impact on congestion?

Each of these questions will be analysed in separate sections below.

3.1 Traffic across the cordon

The fundamental question when the charges were planned and discussed initially was of course “How much will the charges affect traffic – if at all?”. There were many concerns that charges would not affect traffic at all, the most common arguments being “car drivers choosing to drive in spite of the heavy congestion do so because they have no alternatives” and “it is already so expensive to drive and park that an extra charge will not make any difference”.

The target for the traffic reduction was a 10-15% decrease of traffic volumes during rush hours on the most congested links. This target was chosen rather heuristically, and tentatively decided by politicians already before more detailed traffic analyses were made. It was based partly on earlier model-based analyses, and partly on the observation that the 10-15% drop in traffic during the winter and summer holidays resulted in substantial congestion reductions. Subsequent analyses studying bottleneck capacities confirmed that this seemed to be a sensible target (see (Eliasson, 2008b) and (Eliasson, 2009b) for a description of the scheme design process).

After testing several different designs, the final version of the scheme consisted of a single cordon with a time-varying charge of 1-2€ (see section 2). The charge was predicted to decrease traffic across the cordon by 16% during the charged time period, with some variation across the day.

Table 1. Predicted and actual effects on traffic across the cordon

| Time period | Direction | Without charges (veh/hr, 2005) | Predicted effect | Actual effect |
|---------------------------------------|------------------------------------|-----------------------------------|---------------------|------------------|
| AM peak ² | <i>Towards city centre</i> | 22 480 | -16% | -19% |
| | <i>Out from city centre</i> | 16 373 | -17% | -15% |
| PM peak | <i>Towards city centre</i> | 19 310 | -18% | -22% |
| | <i>Out from city centre</i> | 20 893 | -17% | -18% |
| Charged period excl. peaks | <i>Towards city centre</i> | 16 190 | -14% | -24% |
| | <i>Out from city centre</i> | 14 835 | -14% | -19% |
| | | | | |
| Charged period | <i>North (both directions)</i> | 13 850 | -14% | -16% |
| | <i>Southwest (both directions)</i> | 2 774 | -21% | -26% |
| | <i>South (both directions)</i> | 9 358 | -15% | -23% |
| | <i>West (both directions)</i> | 5 319 | -25% | -26% |
| | <i>Lidingö (both directions)</i> | 3 355 | -5% | -16% |
| | Charged period | TOTAL (both directions) | 34 656 | -16% |

When the charges were introduced, actual effects turned out to be larger than predicted, as shown in Table 1. The average reduction across the cordon during the charged time period became 20%, rather than the predicted 16%. Interestingly, the model predicted the effect to be almost equal in the two directions, while the actual effect was asymmetric in the two directions across the cordon.

It was mainly the effect during off-peak hours that was underpredicted, with an underprediction of 5-10 percentage points. There may be several reasons for the off-peak underprediction, but as will be shown below, the most important reason seems to be that the effect on leisure trips (which constitutes the bulk of off-peak traffic in the model) was underpredicted.

Looking at different parts of the cordon, the main source of underprediction was on the southern arterials. In relative terms, the largest underprediction was for the Lidingö traffic, but since this was a rather small traffic volume, it did not contribute much to the overall error.

Was the forecast good enough? In other words, had the system been designed differently if the prediction had been perfect? The most important concern during the design process was whether the traffic reduction would reach the target of 10-15% during peak hours. The model was sufficiently accurate to help planners reach the correct conclusions. The peak-hour charge turned out to be almost perfect, given the

² In this table, the peaks are defined as the 15 SEK and 20 SEK periods, i.e. the AM peak is 7:00-9:00, and the PM peak is 15:30-18:00. However, the highest charge (20 SEK) is charged only between 7:30-8:30 and 16:00-17:30, while the charge is 15 SEK during the rest of the peak periods. In the model, however, the charge is assumed to be 20 SEK for the whole peak. On the other hand, the model does not account for two factors: changes in the timing of the trips, and that the *flow* of traffic may actually increase during the extreme peak (due to reduced congestion) although *demand* ("incoming traffic") decreases. The latter factor makes it pointless to compare traffic measurements with model predictions for very congested periods, since the model-calculated demand may be higher than the physically possible traffic flow.

design target. Had the model given a perfect prediction, it is possible that the charge would have been slightly lower during off-peak hours.

3.2 Effects on travel patterns

An inevitable question is how the “disappearing” car drivers adapt to the charges. Strictly speaking, it is not necessary to know the answer to this to plan and design a charging scheme. As long as the traffic decrease is close to the target, without moving congestion problems to other new roads or creating severe transit crowding, it does not really matter what the “disappearing” drivers choose to do instead. Still, understanding the adaptation mechanisms is useful for checking that model results are reasonable, for improving the design of the scheme, and for communication with professionals and the public.

Table 2 shows how the model distributed the “disappearing” car drivers on other travel alternatives. According to the model, around 2/3 of the traffic reduction would consist of disappearing work trips, while the other 1/3 would be disappearing leisure trips. Half of the disappearing work trips would switch to other modes (mainly transit, but some walk/cycle), while a little less than half would choose other workplaces. Changes in workplaces were not expected to materialise in the short run, however, so the main response for work trips was expected to be mode-switching. The majority of the disappearing leisure trips were forecasted to change destination. A considerable share of leisure trips was also forecasted to disappear altogether, while only a small fraction would switch modes.

Comparing forecasts with outcomes is difficult in this case, because adaptation patterns turned out to be extremely difficult to track. This is largely due to seasonal effects between the “before” and “after” travel surveys. A panel travel survey was carried out, with one wave in September 2004 (before the charges) and the other wave in March 2006 (with charges). Originally, the two waves were meant to be carried out during the same time of the year, but the start of the trial was postponed, so the second wave had to be postponed as well. This created problems with seasonal variations in modal split, destination patterns and trip frequencies. In the analyses, the travel surveys have been re-weighted to account for this, but results are still uncertain.

Table 2. Distribution of “disappearing” car drivers on alternative options.

| | Work trips | | Leisure trips | |
|------------------------------|------------|-----------------------------------|---------------|-----------------------------------|
| | Forecast | Outcome (<i>approximate</i>) | Forecast | Outcome (<i>approximate</i>) |
| Share of disappearing trips | 64% | ~50% | 36% | ~50% |
| Other destination, still car | 42% | ~0% | 58% | ~80-100% |
| Other mode, total | 49% | ~100% | 15% | ~0% |
| Disappeared (fewer trips) | 9% | | 27% | ~0-20% |

With these caveats in mind, forecast and outcome are compared in Table 2. Disappearing car trips turned out to consist of roughly equal shares of work trips and leisure trips; the effect on leisure trips was underpredicted by the model. Virtually all disappearing work trips switched to transit, which was consistent with expectations: the forecast that workplace changes would account for almost half of the disappearing work trips was not plausible in the short run. For leisure trips, the main response was forecasted to be changes in destination and trip frequency, while very few would change mode. This turned out to be a correct prediction.

The underpredicted effect on leisure trips explains most of the underpredicted off-peak effect noted in the preceding section. There may be several reasons for this. One hypothesis is that the model may underestimate how attractive the alternative destinations outside the inner city are. Another hypothesis is that the fact that the model does not model trip chaining underestimates the effects on leisure trips, since some leisure trips are undertaken on the way home from work, or during the day. Changing from car to another commuting mode may then possibly reduce the likelihood of making leisure trips, a phenomenon which the model would not take into account. A third hypothesis is that the model (incorrectly) assumes that if there are passengers in the car, the cost of the trip is divided between all the people in the car, reducing the actual cost, whereas in reality the driver pays the entire cost. Typically there are more passengers/vehicles in leisure trips than in commutes.

3.3 Increase in public transport ridership

The public transport system in Stockholm is well developed and carries a large share of trips, especially to/from the city centre during rush hours, where the mode share is around 75% of all trips. Since it was anticipated that the charges would cause a considerable share of car drivers switch to public transport instead, there were concerns that this would increase crowding problems. To alleviate some of the anticipated crowding problems, metro and commuter train frequencies were increased and new bus lines were introduced. Hence, forecasting the increase in transit ridership was also important to design these capacity improvements.

Table 3. Out-of-home trips by public transport per day; predicted and actual changes.

| | Model results | | Actual change (based on travel surveys) | Change (based on passenger counts) |
|-----------------------------------|----------------|----------------------|--|------------------------------------|
| | Before charges | Predicted change | | |
| Outside → inside cordon | 148 152 | 3.9% | 4.5 % | |
| North of cordon ↔ South of cordon | 23 130 | -0.5% | | |
| Inside → outside cordon | 34 110 | -1.6% | | |
| Within the cordon | 62 316 | 2.3% | | |
| Total across the cordon | 206 097 | 2.4% | 7 % | 6 % |
| | | <i>Adjusted: ~6%</i> | <i>(5 % excl. trend)</i> | <i>(4% excl. trend)</i> |

Table 3 shows model results in terms of out-of-home journeys per aggregate O-D relation. According to the model, transit passengers going from outside the cordon to the inside – which is the majority of the trips across the cordon – would increase by 3.9 %. But the other O-D relations show strange results. First, transit trips between the county south of the cordon and the county north of the cordon is forecasted to *decrease* by -0.5%. Second, transit trips from inside the cordon to the outside are also forecasted to decrease when the charges are introduced. These effects are due to the doubly constrained work trip demand model. Some of the work trips by car across the cordon are forecasted to switch to other destinations (i.e. other workplaces) because of the charges. The doubly constrained work trip model then forces some trips starting within the cordon to end within the cordon, to fill the “vacant” workplaces. This means that the total number of trips going from within the cordon to the outside decreases, leading to a decrease in the number of transit trips from within to the outside (the number of car trips in this relation obviously also decreases). This also explains part of the substantial increase in transit trips within the cordon.

While this may conceivably happen in the long term, it is highly unlikely to happen in the short term. Since it was unfeasibly complicated to fix the work trip destinations in the model, the model results had to be reinterpreted and adjusted. Adjustment of the model results in Table 3 resulted in an expected increase in the number of transit passengers of around 6 %; this was calculated applying the new mode shares for the “inside → outside” and “North ↔ South” to the old trip volumes. These additions are relatively small since the base volumes are comparatively small, so the magnitude of the total increase does not change much even if these additions are somewhat misjudged.

The measurements of the actual effect are only available at an aggregate level. According to the passenger counts carried out by the public transport authority, the volumes crossing the cordon increased by 7%. According to calculations based on the travel surveys carried out before and during the charges, the increase was around 6 %. The number of transit trips across the cordon increased by 2 % from the fall of 2004 to the fall of 2005, i.e. before the charges had been introduced. Looking at longer time series of the number of trips across the cordon, it is likely that this 2 % increase was due to external factors (increased population etc.). So the actual increase of transit ridership due to the charges can be estimated to around 4-5%.

Concluding, did the model give sufficiently accurate answers? On the one hand, allowing work trip destinations to change while at the same time imposing a doubly-constrained demand model clearly produced counterintuitive results. For the purposes of this study, keeping the work trip matrix fixed would have been a better option. On the other hand, the most important conclusion was that the increase of transit ridership would be limited, certainly staying well below 10%, more likely being in the range of 6%. This was a good enough precision to dimension the capacity expansions, and for assuring planners that the increase in transit ridership would not be lead to disastrous crowding.

3.4 Traffic on circumferential links

A common concern with congestion charging schemes is that traffic will be rerouted, creating problems in new places, such as circumferential roads around the charged area and within residential areas. This concern was also aired in Stockholm. In particular, there were concerns about the Essinge bypass (the only major link connecting the southern and northern parts of Stockholm, see Figure 1) and the Bergshamra link (connecting the E4 and E18 arterials just north of the inner city). Both links are heavily congested, and the topology of the network indicated that rerouting effects would increase traffic volumes there. On the other hand, the general traffic reduction was expected to offset some of the rerouting effect.

Table 4. Traffic effects on circumferential links: forecasts and outcomes.

| | AM peak | | PM peak | |
|---------------------------------|------------------|---------------|------------------|---------------|
| | Predicted effect | Actual effect | Predicted effect | Actual effect |
| Essinge bypass (Gröndal bridge) | -4% | 0% | -2% | 4% |
| Bergshamra link | 2% | 0% | 1% | 1% |
| Kymlingelänken | 2% | -1% | 2% | 6% |
| Magelungsvägen | -3% | -9% | 2% | 1% |
| Älvsjövägen | 0% | 6% | -1% | 2% |

Table 4 compares forecasts and outcomes for five circumferential links. In general, the forecasted effects were of the right magnitude, especially considering random variations. In the Essinge case, a trend increase in traffic contributed with 2-3% increase between 2005 and 2006. Hence, the model results indicating a small negative net effect (of increasing trips from rerouting and decreasing trips from a general trip reduction) are in line with the observed effect.

One lesson from ex-ante and ex-post analyses was that model results were potentially dependent on the assignment parameters – the values of time and the cost per distance for the five classes in the multiclass assignment (see section 2.2), and the relative sizes of the classes. This problem had not received much attention in the past, since it is less important for analysing consequences of transport investments, for example. But for congestion charges, a vehicle's value of time and driving cost per kilometre will determine whether it takes a detour around the charged area or not. These assumptions may thus have a potentially important effect on the forecasted link flows, especially for circumferential roads. On the other hand, the topology of Stockholm implied that rerouting mainly affected the bypasses. The fact that the cordon was located on bridges entering the inner city limited the number of rerouting possibilities.

Before the charges were introduced, there was little evidence as to how to set the assignment parameters. In the end, parameters were chosen based partly on value of time studies based on stated preferences and on correspondence with link flow measurements. ((Börjesson & Eliasson, 2007) provide an ex-post study of drivers' values of time and cost per kilometre, based on data from the Stockholm charges.)

3.5 Environmental effects

The charges were partly marketed as an environmental measure, intended to improve air quality in the inner city. According to model results, vehicle kilometres travelled (VKT) in the inner city would decrease by 7%, and emissions somewhat less, since heavy traffic would decrease less than private car traffic, and heavy traffic emit more per kilometre.

Measuring VKT in a relatively large area such as the Stockholm inner city (around 20 km²) is very difficult, since only a subset of the links can be measured. Two methods were tried: one based on a random sample of links that were measured before and after the introduction of the charges, and one based on calibration of matrices from the transport model using link counts, extrapolating non-measured link flows from the calibrated matrices. The two methods gave rather different results, and both turned out to depend on what link measurements were included in the calculations, raising concerns about sampling biases.

Bearing these uncertainties in mind, the best estimate of VKT reduction is 10-15% - that is, more than the forecasted 7%. The difference seems to be due to flows on small, non-measured links within the inner city, so it is difficult to judge whether the model was "sufficiently correct" or not – or even whether the 7% or the 10-15% figure are closest to objective reality. From a design and policy point of view, this discrepancy would not have mattered: no decisions would have been different had the true figure been known in advance.

3.6 Timing of trips

During the design process, it was anticipated that changing departure time would be the most common adaptation mechanism for drivers. However, the transport model did

not model departure time choice – trips were allocated to one of four time periods in fixed shares (see section 2.2). Therefore, the time-differentiation of the charges was designed based on the traffic time-profile from traffic measurements. There were concerns that there might appear congestion problems immediately before and after the charges, or before and after the highest charge periods. However, calculations based on scheduling elasticities (taken from the literature) indicated that the overall traffic decrease would mean that traffic demand would be lower during all time periods.

Figure 2 shows the distribution of traffic across the day. In the 2006 distribution, there are visible peaks immediately before and after the charged period; these have disappeared in later measurements (2007 and later).

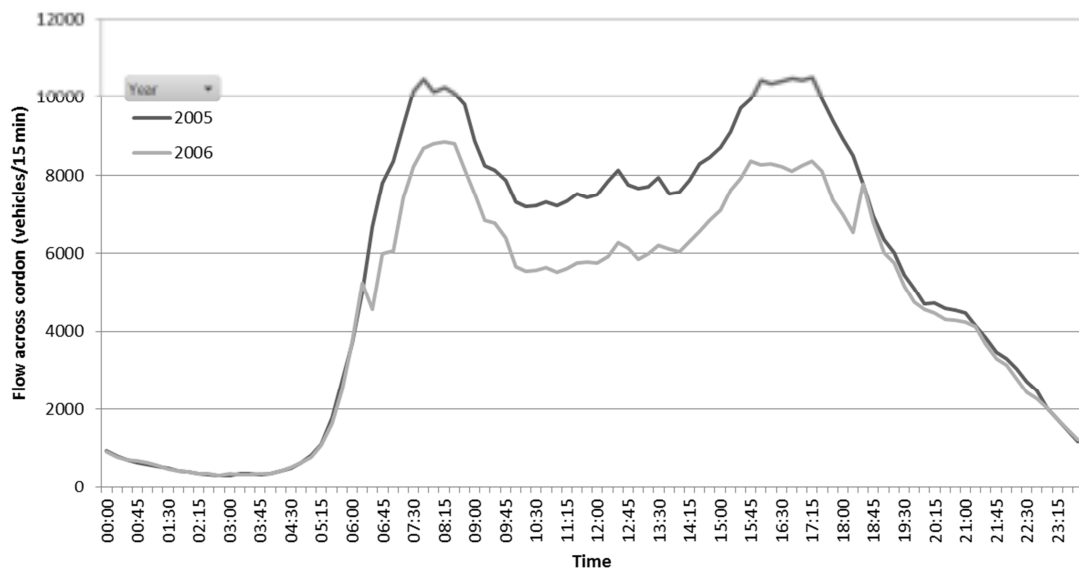


Figure 2. Time distribution of traffic across cordon (both directions).

The aggregate time profile shows a surprising feature: there is no immediately visible effect on the time profile of traffic, despite the time differentiation of the charges. However, (Karlström & Franklin, 2009) show that departure times indeed were affected. Using travel survey data, they show that morning commuting trips shifted away from the peak. For other trip purposes and times of day, the seasonal and random variations of the travel patterns were too large for any conclusions to be drawn.

There are two reasons that the aggregate time profile changed so little. First, the reduced congestion meant that traffic flow stayed below capacity for most parts of the peak, which was not the case before the introduction of the charges. This implied that the peak-spreading due to incoming flow above capacity declined and almost vanished, making the peak of the flow, measured at a specific point, more pointed. Second, departure times seem to have shifted in both directions: while some travellers shifted away from the peak to avoid the higher charges, other travellers who had avoided the congested peak hours may have shifted into the peak when travel times decreased. This effect is shown and illustrated by (Börjesson & Kristoffersson, 2012), using a dynamic model based on data from before and after the charges.

In summary, the lack of departure time modeling caused no major surprises, although the time differentiation of the charges could perhaps have been designed more efficiently if a proper departure time model had been available. The indicative calculations predicting that traffic would decrease during all time periods were correct:

the overall traffic reduction did offset any rescheduling of trips from peak to off-peak periods. The belief that the charges would affect departure times was confirmed, but the effect was considerably smaller than anticipated. However, as shown by (Börjesson & Kristoffersson, 2012), taking (re)scheduling of trips into account may make a large difference when the economic efficiency of a congestion pricing scheme is estimated.

3.7 Travel times

The *Sampers* model used in the design process uses the static network assignment model EMME/2 to calculate link flows and travel times from the car O-D matrices. It is well known that static network models cannot handle severe congestion well for a number of reasons, such as inability to handle upstream propagation of queues and difficulties with handling congestion at intersections and junctions. Hence, it was known that the model tended to underestimate changes in travel times in the network.

Table 5 shows travel times for a few examples of routes, each comprising a number of links in the static model. The routes have free-flow travel times between 3 and 12 minutes. The static model generally underpredicts travel times, although in some cases it is close to reality. When the charges are introduced, however, the difference between model and reality increases. This is most easily seen in Table 6, where changes in travel times are compared between model and reality. The difference is substantial. The model fares somewhat better in terms of relative reduction of “delay time”, i.e. travel time minus free-flow travel time. For the first four routes, model results are close to actual effects. For the last three routes, model results have the right sign and indicate substantial changes, but actual changes are much larger than predicted. For the fifth route, the model predicts a negligible increase of the travel time, while the actual effect is a substantial reduction of travel time. This is because the queue on this route is due to a bottleneck on a different link, causing a queue that propagates back and spreads onto this route.

Table 5. Link travel times (minutes), model and measurements.

| | Model | | | Reality | | |
|---|-----------|------------|--------------|-----------|------------|--------------|
| | Free flow | No charges | With charges | Free flow | No charges | With charges |
| Lövstavägen-Fridhemsplan | 12.5 | 19.1 | 17.7 | 11.0 | 25.8 | 21.6 |
| Johanneshovsbron N- Klarastrandsleden N | 11.8 | 23.5 | 20.4 | 11.8 | 27.1 | 23.4 |
| S:t Eriksplan-Valhallavägen | 5.1 | 6.8 | 6.5 | 3.2 | 8.1 | 7.0 |
| Valhallavägen: Lidingövägen-Roslagstull | 3.2 | 4.4 | 4.3 | 2.1 | 6.2 | 5.5 |
| Stocksund-Ulvsundaleden | 6.1 | 10.3 | 10.4 | 5.3 | 10.9 | 9.5 |
| Klarastrandsviadukten - Örbyleden | 11.4 | 16.8 | 15.7 | 11.0 | 24.6 | 11.7 |
| Sveavägen: Sergels torg - Sveaplan | 6.5 | 7.6 | 7.4 | 3.4 | 7.0 | 5.1 |
| Valhallavägen: Roslagstull-Lidingövägen | 3.2 | 5.8 | 5.3 | 2.4 | 8.0 | 5.0 |

Table 6. Changes in travel times, model and measurements.

| | Travel time reduction, minutes | | % reduction of "delay time" | |
|---|--------------------------------|---------|-----------------------------|---------|
| | Model | Reality | Model | Reality |
| Lövstavägen-Fridhemsplan | -1.4 | -4.2 | -21% | -29% |
| Johanneshovsbron N- Klarastrandsleden N | -3.0 | -3.6 | -26% | -24% |
| S:t Eriksplan-Valhallavägen | -0.3 | -1.1 | -18% | -23% |
| Valhallavägen: Lidingövägen-Roslagstull | -0.1 | -0.7 | -11% | -16% |

| | | | | |
|---|------|-------|------|------|
| Stocksund-Ulvsundaleden | 0.1 | -1.5 | 3% | -26% |
| Klarastrandsviadukten - Örbyleden | -1.0 | -12.8 | -19% | -94% |
| Sveavägen: Sergels torg - Sveaplan | -0.2 | -1.9 | -17% | -53% |
| Valhallavägen: Roslagstull-Lidingövägen | -0.5 | -3.0 | -19% | -53% |

The assignment model hence gets effects on delay times about right in some cases, but in other cases they are wide of the mark. An analysis by (Engelson & Van Amelsfort, 2012) shows that it does not help much to adjust the volume-delay functions: the problems arise because the model does not account for spillback queues, blocking of intersections and similar network effects.

The shortcomings of the static network model for travel time calculations meant that travel time-based measures such as consumer surplus were not very useful during the design process. Instead, the charging system was designed using various variants of model-based congestion indices. The most important target was the calculated congestion index (calculated travel time divided by free-flow travel times) at the major bottlenecks. The target was that bottlenecks close to the cordon, where the most significant bottlenecks are located, should have congestion indices less than 1.7, while bottlenecks further from the cordon should have congestion indices below 2. This combination of model results, traffic engineering rules-of-thumb and manual checking of results turned out to work satisfactorily in the Stockholm case.

Had the system or the topology of the city been more complicated, however, it is probable that these crude methods and the static model had given less satisfactory outcomes. It would have been very difficult to use the model to fine-tune the charges with respect to specific times and places. The design process concentrated primarily on avoiding to cause severe second-best problems (namely increasing traffic in residential areas and on congested bypasses), and satisfactory traffic flow reductions in the bottlenecks.

Moreover, the present model does now allow for social appraisal of the system, since consumer surplus measures cannot be trusted. This means that if there had been several acceptable schemes accomplishing desired traffic reductions and minor second-best problems these could not have been ranked according to economic efficiency. Hence, moving from static assignment models to dynamic assignment models should be a high priority whenever policies to tackle severe congestion are to be designed or evaluated.

4 CONCLUSIONS

Comparing model predictions with the actual effects of the Stockholm congestion charges, it can be concluded that most effects on travel demand were predicted fairly accurately. Effects on travel times, on the other hand, were underpredicted due to inherent deficits of the static assignment model. The main conclusions from the comparison are:

- Traffic reduction across the cordon was close to the forecast, especially during rush hours. This is especially remarkable considering that the forecasted traffic reduction was considered unrealistically large by the design team that were responsible for the forecasts.
- The effect on leisure trips was somewhat underpredicted, leading to some underprediction of off-peak traffic effects. The predicted substitution pattern

- turned out to be correct, however, with leisure trips changing destinations or being cancelled rather than changing modes.
- Effects on circumferential traffic were predicted to be almost negligible, which turned out to be correct.
 - The increase in public transport ridership was close to predictions, once the unrealistic effects of the doubly constrained work trip model had been taken into account.
 - That the model did not account for changes in departure times meant that it could not be used to design or evaluate the time-differentiation of the charges. Somewhat surprisingly, however, aggregate traffic effects still came out about right.
 - The parameters controlling the network assignment – values of time and cost per kilometre of the different user classes – may have substantial impact on the forecasts, depending on the network topology. Unfortunately, there is little evidence as to how these parameters should be set, especially in a network with road user charges.
 - 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.

Two main conclusions can be drawn regarding practical applications. First, 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 in 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.

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 widespread 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.

The second conclusion has to do with the transferability of the Stockholm experiences. 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 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.

5 ACKNOWLEDGMENTS

We are grateful to Mats Tjernqvist and Tobias Thorsson (WSP Analysis & Strategy) for running the transport model scenarios, and to VINNOVA for financial support.

6 REFERENCES

- Anguera, R. (2005). The Channel Tunnel – an ex-post economic evaluation. *Proceedings of the European Transport Conference*.
- Bain, R. (2009). Error and optimism bias in toll road traffic forecasts. *Transportation*, 36, 469–482.
- Bain, R. (2011). *An investigation of the causes of over-optimistic patronage forecasts for selected recent toll road projects*. GHD Report.
- Brundell-Freij, K., Jonsson, L., & Källström, J. (2009). Accepting charging – a matter of trusting the effects? *Proceedings of the European Transport Conference*.
- Börjesson, M., & Eliasson, J. (2007). The value of time of car drivers choosing route: evidence from the Stockholm congestion charging trial. *Proceedings of the European Transport Conference*. Leiden, Netherlands.
- Börjesson, M., Eliasson, J., Hugosson, M. B., & Brundell-Freij, K. (2012). The Stockholm congestion charges—5 years on. Effects, acceptability and lessons learnt. *Transport Policy*, 20(0), 1–12.
- Börjesson, M., & Kristoffersson, I. (2012). *Estimating welfare effects of congestion charges in real world settings* (CTS Working Paper No. 2012:13). Centre for Transport Studies, KTH Royal Institute of Technology.
- de Palma, A., Lindsey, R., & Niskanen, E. (2006). Policy insights from the urban road pricing case studies. *Transport Policy*, 13(2), 149–161.
- Eliasson, J. (2008a). Lessons from the Stockholm congestion charging trial. *Transport Policy*, 15(6), 395–404.
- Eliasson, J. (2008b). Lessons from the Stockholm congestion charging trial. *Transport Policy*, 15(6), 395–404.
- Eliasson, J. (2009a). A cost-benefit analysis of the Stockholm congestion charging system. *Transportation Research A*, 43(4), 468–480.

- Eliasson, J. (2009b). Expected and unexpected in the Stockholm Trial. In A. Gullberg & K. Isaksson (Eds.), *Congestion taxes in city traffic. Lessons learnt from the Stockholm Trial*. Nordic Academic Press.
- Eliasson, J., Hultkrantz, L., Nerhagen, L., & Rosqvist, L. S. (2009). The Stockholm congestion - charging trial 2006: Overview of effects. *Transportation Research Part A: Policy and Practice*, 43(3), 240–250.
- Eliasson, J., & Jonsson, L. (2011). The unexpected “yes”: Explanatory factors behind the positive attitudes to congestion charges in Stockholm. *Transport Policy*, 18(4), 636–647.
- Eliasson, J., & Mattsson, L.-G. (2006). Equity effects of congestion pricing: Quantitative methodology and a case study for Stockholm. *Transportation Research Part A: Policy and Practice*, 40(7), 602–620.
- Engelson, L., & Van Amelsfort, D. (2012). *The role of volume-delay functions in forecast and evaluation of congestion charging schemes, application to Stockholm* (CTS Working Paper No. 2012:X). Centre for Transport Studies, KTH Royal Institute of Technology.
- Flyvbjerg, B., Holm, M. K., & Buhl, S. L. (2005). How (in) accurate are demand forecasts in public works projects? *Journal of the American Planning Association*, 71(2), 131–146.
- Franklin, J., Eliasson, J., & Karlström, A. (2010). Traveller Responses to the Stockholm Congestion Pricing Trial: Who Changed, Where Did They Go, and What Did It Cost Them? In Saleh & Sammer (Eds.), *Demand Management and Road User Pricing: Success, Failure and Feasibility*. Ashgate Publications.
- Fridström, L., Minken, H., Moilanen, P., Shepherd, S., & Vold, A. (2000). *Economic and Equity Effects of Marginal Cost Pricing in Transport* (Research Reports No. 71). Government Institute for Economic Research Finland (VATT). Retrieved from <http://ideas.repec.org/p/fer/resrep/71.html>
- Gullberg, A., & Isaksson, K. (2009). The Stockholm Trial. In A. Gullberg & K. Isaksson (Eds.), *Congestion taxes in city traffic. Lessons learnt from the Stockholm Trial*. Nordic Academic Press.
- Kain, J. F. (1990). Deception in Dallas: Strategic Misrepresentation in Rail Transit Promotion and Evaluation. *Journal of the American Planning Association*, 56(2), 184–196.
- Karlström, A., & Franklin, J. P. (2009). Behavioral adjustments and equity effects of congestion pricing: Analysis of morning commutes during the Stockholm Trial. *Transportation Research Part A: Policy and Practice*, 43(3), 283–296.
- Kickhöfer, B., Zilske, M., & Nagel, K. (2010). *Income dependent economic evaluation and public acceptance of road user charging*. Working Paper, TU Berlin. Retrieved from http://scholar.googleusercontent.com/scholar?q=cache:Sa3or1g1XswJ:scholar.google.com/+matsim+public+acceptance&hl=en&as_sdt=0,5&as_vis=1
- Li, Z., & Hensher, D. A. (2010). Toll Roads in Australia: An Overview of Characteristics and Accuracy of Demand Forecasts. *Transport Reviews*, 30(5), 541–569.
- Parthasarathi, P., & Levinson, D. (2010). Post-construction evaluation of traffic forecast accuracy. *Transport Policy*, 17(6), 428–443.
- Pedersen, K., Denver, L., Daly, A., & Rohr, C. (2001). Great Belt Link - Traffic before and after. *Proceedings of the European Transport Conference*.
- Petersen, T. (2010). The validity of the Öresund forecasts. *Petersen, T.: Essays in spatial economics*, PhD thesis. Stockholm: KTH Royal Institute of Technology.
- Rich, J., & Nielsen, O. A. (2007). A socio-economic assessment of proposed road user charging schemes in Copenhagen. *Transport Policy*, 14(4), 330–345.
- Santos, G., Newbery, D., & Rojey, L. (2001). Static Versus Demand-Sensitive Models and Estimation of Second-Best Cordon Tolls: An Exercise for Eight English Towns. *Transportation Research Record: Journal of the Transportation Research Board*, 1747(-1), 44–50.
- Van Wee, B. (2007). Large infrastructure projects: a review of the quality of demand forecasts and cost estimations. *Environment and Planning B: Planning and Design*, 34(4).