2010

DISCUSSION PAPERS

12

Joint Transport Research Centre

ROUND TABLE
4-5 February 2010

International Transport Forum

IMPLEMENTING CONGESTION CHARGING

SUMMARY AND CONCLUSIONS

OECD

ORGANISATION
FOR ECONOMIC
CO-OPERATION AND
DEVELOPMENT

International Transport Forum

JOINT TRANSPORT RESEARCH CENTRE

Discussion Paper No. 2010-12

Round Table, 4-5 February 2010

Implementing Congestion Charging

SUMMARY OF DISCUSSIONS

April 2010

International Transport Forum

The International Transport Forum is an inter-governmental body within the OECD family. The Forum is a global platform for transport policy makers and stakeholders. Its objective is to serve political leaders and a larger public in developing a better understanding of the role of transport in economic growth and the role of transport policy in addressing the social and environmental dimensions of sustainable development. The Forum organises a Conference for Ministers and leading figures from civil society each May in Leipzig, Germany.

The members of the Forum are: Albania, Armenia, Australia, Austria, Azerbaijan, Belarus, Belgium, Bosnia-Herzegovina, Bulgaria, Canada, Croatia, the Czech Republic, Denmark, Estonia, Finland, France, FYROM, Georgia, Germany, Greece, Hungary, Iceland, India, Ireland, Italy, Japan, Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Mexico, Moldova, Montenegro, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Kingdom and the United States. The Forum's Secretariat is located in Paris.

Joint Transport Research Centre

The OECD and the International Transport Forum established a Joint Transport Research Centre (JTRC) in 2004. The Centre conducts co-operative research programmes addressing all modes of transport to support policy making in Member countries and contribute to the Ministerial sessions of the International Transport Forum.

JTRC Discussion Papers

The JTRC Discussion Paper Series makes economic research commissioned or carried out at the Joint Transport Research Centre available to researchers and practitioners. The aim is to contribute to the understanding of the transport sector and to provide inputs to transport policy design. The Discussion Papers are not edited by the JTRC and they reflect the author's opinions alone. They can be downloaded from:
http://www.internationaltransportforum.org/jtrc/DiscussionPapers/jtrcpapers.html

The International Transport Forum's website is at:
http://www.internationaltransportforum.org/

For further information on the Discussion Papers and other JTRC activities, please email: itf.contact@oecd.org

TABLE OF CONTENTS

1. INTRODUCTION

The Round Table addressed the broad question of what research and experience tell us about how to arrive at a successful introduction of congestion charging schemes. Attention was limited mostly to urbanized areas where road traffic congestion is or may become an issue. "Success" means (a) that a policy is implemented, (b) that it works, (c) that it is accepted by actual and potential users, and (d) that it generates benefits for society overall. In order to shed light on these dimensions of success, lessons are drawn from more and less successful attempts to implement charges (Section 2). In addition, we ask if and how the evolving understanding of the economics of road traffic congestion charging might affect the assessment of congestion charging policy (Section 3). The conclusions in Section 4 summarize the main recommendations for policy-makers that contemplate the introduction (or the removal) of congestion charges.

Transport economists have long lamented the lack of policy interest for a tool that to them is as obviously welfare-improving as congestion charges. However, what is obvious in principle is less obvious in practice. Questions regarding the desirability and feasibility of congestion charges become apparent when policy constraints and costs are taken into account. How to convince voters and their representatives that it is a good idea to make travel more expensive when traffic is bad? How to set charges and deploy revenues so that the distribution of gains and losses constitutes a marketable political proposition? Section 2 discusses these and related issues by taking a careful look at practice, in particular in Singapore, Stockholm and London. In studying these cases we try to distinguish idiosyncrasies of the particular cases from general concerns.

Section 2.1 ends with some remarks on cordon pricing, value pricing and parking pricing. Cordon pricing and value pricing both are approximations to ideal congestion charges, but reflect different basic strategies. Cordons focus on maintaining an acceptable average level of service, while value pricing is about offering variety (low quality at a low or zero price, higher quality at a premium).[1] Parking pricing is part of the debate because parking charges could to some extent mimic congestion charges. Probably more important, however, is that removing subsidies to parking on public roads (as well as possibly through reducing relatively beneficial tax treatment of employer-offered parking) is likely to have major beneficial effects on the use of space and on congestion levels where space and road capacity are scarce, while involving very low cost.

[1] Though all forms of congestion charge offer a variety of response options (shifting time of travel, shifting mode, etc.), value pricing is different by allowing last-minute decisions in response to recent information on travel conditions.

The principal question in Section 2 is how to improve the chances of getting congestion charges introduced. Focussing on implementation implies a presumption that congestion charging is a sensible policy instrument to combat excessive congestion. There is indeed widespread agreement among transport economists that, where there is congestion, charges are required to ensure that potential road users take the marginal social cost of their trips into account when deciding if, how, where, and when to travel. When individuals decide on the basis of marginal social costs instead of marginal private costs, welfare increases because the marginal external cost of congestion is no longer ignored so that the social benefit of an additional trip will be (more or less) equal to its social cost.

Translating this key principle into practical recommendations for policy is, however, less than straightforward. Section 3 deals with some complicating factors. Congestion charging systems are not costless. To ensure a welfare-improving outcome, the efficiency gain from charging needs to be greater than the cost of the system. The size of the efficiency gain depends on getting the prices right. Determining the level and structure of charges requires an understanding of the physical and behavioural aspects of congestion. The mechanisms involved are far more complex than is suggested by the argument that travel times increase with traffic volumes, as used in the basic rationale for charges. Charges are always an approximation to the theoretical ideal, so that decisions must be made on what approximation is best. Experience demonstrates that analytical approaches using disaggregated network models are more likely to produce an efficient result than prices based only on common sense. Such models are better placed to capture complex network activity across modes and reveal impacts on traffic flow that cannot easily be anticipated. A period of model testing and iteration before charges are set is indicated.

Charges are introduced in a world that is rife with other market imperfections and where other policy objectives than efficiency matter. Again, the question is if and how this affects recommendations for setting congestion charges. This is a subject of some controversy, but the emerging view is that charges should remain closely tied to marginal external congestion costs rather than adjusted to compensate for one or a range of other imperfections. There was some support for making the charge deductible from income taxes for commuters. However, this needs to be determined within the broader taxation framework and existing deductions for commuting expenses. Moreover, other forms of revenue use may have a better claim than tax deductions.

2. LESSONS FROM PRACTICE

2.1. Some general principles

2.1.1. Acceptance

A successful congestion charging scheme is one that works technically and reduces congestion, is acceptable, and generates net socio-economic benefits. Acceptability is the overriding concern for policy makers as without it no lasting implementation is possible. It follows that acceptability affects system design and that the way impacts, in particular benefits, are perceived is critical. It also tends to lead to congestion pricing systems being modified and expanded to contribute to a variety of other policy goals, including broader tax reforms and environmental protection. In extreme cases there could be perceived benefits but social losses. This could be the case, for example, where “congestion” charges are introduced to improve the environment but where there is actually relatively little congestion. The improvement in the urban environment that results may well be real but smaller than the cost of running a congestion pricing system. In the absence of major congestion mitigation the net benefits are likely to be negative. At the other extreme, a perception of unacceptably high system costs may prevent implementation where charges would in fact produce net benefits. This is the more common scenario because of the difficulty of presenting the benefits of congestion management to the public in tangible terms.

Modifying the design of congestion charging systems to promote acceptability by targeting goals other than congestion management usually results in a trade-off between efficiency benefits and acceptability. This can lead to a sterile divergence of views, with one side arguing that focusing on acceptability results in forgoing too much of the potential benefits of charging and the other side saying that too much emphasis on economic measurement of benefits jeopardize acceptability to the point of becoming counterproductive. A more productive intermediate position is that acceptability comes first even though it may be costly in the sense that some benefits are given up.

Rather than taking perceptions as immutable, communication and marketing can help shape them, paving the way for the introduction of congestion charging and for increasing its benefits over time, and potentially reducing the need to sacrifice efficiency in system design. As the existence of effective congestion charging systems in Singapore, Stockholm and London shows, the tension between acceptability and welfare potential is not irredeemable.

Certainly, it makes no sense to drop a feasible and productive solution because in principle a better one is available.

Acceptability evolves over time. Surveys of public opinion reveal a typical pattern whereby acceptance first rises as the general idea is discussed, then deteriorates as details become known and implementation approaches, but rises to its highest level once the system is operational. Higher acceptance of operational systems may relate to sharply increased awareness of now tangible benefits or to cognitive adaptation. At any rate, the dynamics of public acceptance make politics a critical factor. As is well known, the Mayor of London, Ken Livingstone, risked his political future by proposing the introduction the London Congestion Charge in the run up to a Mayoral election. In Stockholm, conditions imposed by the junior party in the governing coalition drove reluctant political partners to introduce congestion charging for a trial period. It appears that no political champion is needed for continuing a scheme (likely partly because support tends to rise after implementation). The citizens of Stockholm voted to reinstate congestion pricing on expiry of the charging trial. In London, although the Western Extension Zone to the Congestion Charge will be discontinued following a change in government and other extensions are no longer under consideration, there are no plans to discontinue charging in the Central Zone.

2.1.2. Prime objective

The primary goal of a congestion charging system is to reduce congestion to a more efficient level. This level is determined by the cost of the scheme, by the behavioural response to it, and by the relation between external congestion costs and traffic volumes. This is straightforward but has some important implications. First, congestion can only be reduced where it is (perceived to be) excessive to begin with. Introducing a congestion charging system in anticipation of excessive congestion in the future is a very hard proposition to sell, because perceived benefits will be low in the absence of an acute problem (certainly after discounting). This notion that drastic policy changes are easier at times of crisis is familiar from environmental and safety policy, as well as from macroeconomic policy.

Second, setting a congestion charge low and gradually increasing it is a risky strategy for gaining acceptance. If the charge starts too low to have a visible impact on congestion the strategy will backfire. Similarly, when congestion charging is part of a larger reform of charging for use of the roads, the congestion element may be critical to the new system being perceived as a success; leaving differentiation of a new charge to deal with congestion to a second stage of reform may undermine rather than ease acceptance.

The third implication relates to policy targeting. The introduction of congestion charges is often defended for reasons other than congestion, including environmental benefits and meeting a revenue need. The point is not that such ancillary benefits don’t exist or are unimportant, but rather that they should not be turned into the first priority. Overemphasizing environmental benefits may have contributed to the rejection of congestion charging proposals in New York in 2008, although equity concerns probably were at least as important. The primary goal for the proposed congestion charge in Gothenburg is raising revenue for investment in a road tunnel, as congestion in the city is relatively light. Oslo, Bergen and Trondheim in Norway successfully introduced toll rings to raise revenues for infrastructure investments but without confusing the objective. There is a risk that if a revenue-targeted system is implemented as a congestion measure, acceptance will decline rather than improve after implementation. A related issue concerns shifting from fuel taxes to road pricing (in the sense of distance charges rather than congestion charges) for the collection of public revenue as discussed in the USA, mainly in relation to the revenue shortfall in the Highway Trust Fund. This may be a reasonable policy but whatever its merits, it is not mainly a congestion charging policy and *if* marketed as such would probably result in problems of acceptance. Indeed, advocates of distance-charges focus on revenue needs more than on congestion management, as the latter is seen to be particularly problematic in terms of acceptance.

As noted, there is a potential for ancillary benefits to congestion charges and these should be taken into account when designing, marketing, and assessing the scheme. Some environmental impacts of car use are strongly correlated with congestion. The damage from noise and pollutants like particles, as well as general discomfort caused by dense traffic, serve as examples. It makes sense to take these effects into account when setting charges. For a local scheme, it does not make sense, however, to set congestion charges in relation to greenhouse gas emissions: the effects of emissions of greenhouse gases are decidedly non-local, and much better instruments than congestion charges exist for tackling them (notably, fuel taxes). This, of course, does not mean that congestion charges do not have an effect on CO2 emissions: emissions will fall if traffic levels fall and will decline with smoother traffic flows at more optimal speeds.

2.1.3. Revenues

Pricing reforms in transport, including but not limited to congestion charges, are sometimes accompanied by commitments to revenue neutrality.[2] For example, at the time of introduction of congestion charges in Singapore, revenue neutrality was promised and (more than) achieved by reducing vehicle taxes; revenue neutrality became of lesser concern as the system became firmly implanted. The Dutch proposal to replace vehicle ownership-based charges with usage-based charges was designed to be revenue neutral.

Revenue-neutrality is not a universal objective or constraint. Discussions on introducing usage-based charges in the USA are often inspired by prevailing revenue shortfalls in the Highway Trust Fund (fed by capped fuel taxes), so revenue-neutrality is not a concern.[3] The

[2] We refer here to practical proposals, not to analytical work. Assuming neutrality is a useful analytical device, but this as such does not indicate whether it is justifiable or not.

[3] The US discussions on usage-based pricing are closely tied to infrastructure expenditures. This is less so in Europe. In London and Stockholm, the introduction of charges did not change plans for road expansion, despite the expected effect that the need for infrastructure becomes less pressing.

London Congestion charging scheme produced additional revenues, although its prime objective was to reduce congestion. Discussions on potential charges in Moscow look to increased revenues from transport. In all cases, however, discussions on what to do with revenues are at the core of system design and of public debate. The odd man out here is Stockholm, where revenue concerns were not crucial, and decisions on what to do with the new funds came almost as an afterthought.[4]

Commitments to revenue neutrality are common when changes in the way taxes are collected are proposed. When the only aim is to *replace* one tax with another that is less costly in terms of efficiency or collection costs, or more difficult to avoid or otherwise preferable, opting for neutrality is relatively straightforward.[5] When the change involves internalizing an external cost the picture changes, as now a tax is *added*. Pricing the externality inevitably involves raising new revenues. The issue then becomes what to do with those revenues. They can be recycled into reducing other taxes, and if these are transport taxes then the overall change might be neutral in terms of transport tax revenue. There are, however, two important respects in which the change cannot be neutral for all users. First some road users will be worse off financially from the change (those paying the highest rates of the congestion charge most often) and some better off.[6] And charging for congestion requires expenditure on the system for monitoring traffic and collecting and processing payments. Experience in Singapore, London and Stockholm suggests the costs of the system represent 15 to 30% of the gross charge income and that the figure cannot be compressed to below around 10%, given their definition of costs. The precise ratio is partly a function of the level of charges that bring congestion to the optimal level (the higher the charge the larger the revenue), but the key message is that congestion charges are a relatively expensive way of raising revenue.

Any calculation of neutrality should be based on net revenues and not gross income from congestion charges. This should not be too difficult to communicate. Achieving the large benefits of congestion relief requires a system that has to be paid for.[7] And if the objective is to maximize net revenues rather than manage congestion, taxes on fuels and vehicles are much cheaper to collect. The communications mission becomes more complicated when congestion charging is part of a wider reform of transport charges and taxes. The answer may be to put the emphasis on transparent use of revenues and be precise

[4] There are earmarks on how revenues of the Stockholm scheme are spent, which might suggest revenue concerns are crucial. However, the earmarks are mainly an accounting issue with little impact on prevailing revenue streams.

[5] Even here neutrality is not entirely straightforward, as it can be argued that more revenues should be raised in order to expand the supply of public services when raising revenues becomes cheaper.

[6] Whether consumer surplus rises or falls is a broader question, discussed in Section 3

[7] A counterargument here is that the revenue flows triggered by charges are very large compared to the net benefits from congestion relief. This argument has some merit but tends to be overstated, as discussed in Section 3.

about what is meant by revenue neutrality if it is committed to.[8] Public support can be improved by a neutrality pledge that allays suspicions of hidden tax increases. Political support is increased by revenue neutrality as it alleviates concerns regarding public finance shortfalls.

2.1.4. Whatever the technology, running a congestion charging system is not cheap

Once it is clear what a congestion charging system is supposed to achieve, a technology that meets the requirements can be found. It deserves emphasis that setting goals and then choosing means to attain them is the logical order of things. Taking the opposite approach, aiming to make congestion charging an application for a technology seeking to develop a market, should be avoided. Choosing technology as a function of system design also avoids false choices between for example dedicated road-side communications (DRSC) and GPS/GSM-based systems.[9] These systems fit different contexts, e.g. depending on whether there are many or just a few charging points.

Experience with congestion charging schemes suggests they require costly investments and are expensive to operate. Operating costs generally outweigh start-up costs (often by a factor of 10 according to Bernhard Oehry) and should not be ignored in the design stage, as otherwise there is a risk of investing too little in capital. Acceptability increases costs for various reasons. First, as emphasized in the discussion of the Stockholm experience, policy-makers will not incur the risk of going ahead with a system unless they can be convinced that it will work as announced from day one. Minimizing the risk of malfunctioning leads to duplication of components or even systems, which inflates costs. A functioning system also is one that is enforceable. Legal constraints, for example on what constitutes proof of non-payment when payment was due, tend to make enforcement expensive. Adaption of congestion charging systems can save costs over time and much can be learned from the early adopters. Political risk is reduced when local policy-makers can argue that congestion charges are standard and good practice elsewhere. This helps avoid excessive risk aversion and over-specification in system design. Stockholm began with a system that installed DRSC transponders on all vehicles. But as the legal requirements for proof of identity for enforcement pushed development of the camera based automatic number plate recognition system to levels of performance far beyond capabilities in the early stages of design, transponders became redundant eliminating one cost element.

8 Revenue neutrality is difficult to define in a dynamic sense (as it is uncertain how revenues would evolve were there no reform). It was noted that opposition to congestion charges may arise for fear of high future charges, even if current charges do not lead to higher average tax burdens.

9 The view that technology is neutral is supported by evidence that (a) operating costs dominate fixed costs over the life cycle of the investment and (b) the costs of an additional user in a system are largely independent of technology, being roughly equivalent to the cost of an internet or mobile phone subscription (i.e. 10 to 20€ per month in 2010).

Acceptability requires that the congestion charging system accommodates occasional users.[10] These users are not familiar with the system and can't be expected to subscribe to cost-cutting services in the same way as frequent users. Making occasional users pay less than frequent users poses an incentive problem, in that everyone will try to look like an occasional user, and raises fairness concerns. Making occasional users pay more may pose political or legal issues[11] although it is a frequent feature of public transport and many other service products. Accommodating occasional users is a key driver for costs and for the overall design of the system in the sense of limiting system complexity and versatility. Occasional users may generate little revenue, particularly in relation to the large costs they impose on the system.

Costs are also affected by interoperability requirements. The costs are not so much driven by technological as by procedural requirements, i.e. who is responsible for what part of handling transactions.[12] Since incentives for interoperability are weak from the point of view of an individual system, progress has been slow in the EU, inducing European authorities to mandate it.

2.1.5. Differing approaches: value-pricing and area-pricing

The Singaporean and European examples of congestion charging use cordons to charge for entry or travel in a congested zone. This is different from value pricing systems used in the US, where travellers on a particular (segment of a) facility are given a choice between using faster toll lanes and slower free lanes.[13] The introduction of value pricing was a response to shortages of funds to provide new capacity but interest for applying it to fund maintenance or make better use of existing infrastructure is rising ("value pricing 2.0", Poole, 2009).

One potential explanation for the different approaches lies in the typical spatial structure of urbanized (congestion-prone) areas on either side of the Atlantic. European cities correspond more closely to the monocentric, radial archetype, while US cities are more polycentric and grid-like.[14] The European pattern lends itself more to the introduction of cordons, as there are "natural" boundaries, which are less self-evident in the typical US context. While there is some truth to this argument, it is not complete. It is, for example,

10 The argument applies to passengers, not to trucks where competition concerns make equal payment for all users imperative.

11 This is mainly an issue that concerns truck km charges, where discrimination on the basis of country of registration of the vehicles is not tolerated by agreements such as the European Union's treaties.

12 While overall technology costs don't dominate, there may be losers from standardisation.

13 We use the term "value pricing" to indicate "partial facility pricing", because of the close connotation with product differentiation. The U.S. government use of the term is different, as value pricing refers to both partial facility pricing and cordon pricing – value pricing is in fact the same as congestion pricing.

14 There are of course differences within both continents, with newer cities in the US being closer to the grid patterns, and spatial structures in Europe evolving away from the monocentric pattern to some extent. Nevertheless, as far as averages make sense, differences between both continents remain.

hard to see why value-pricing could not be introduced *if desired* on the A1 Highway from Charles De Gaulle airport to central Paris, when it is possible to create a dedicated taxi lane. Similarly, the bus lane from Heathrow airport to London could be turned into a HOT-lane. Spatial structure does not prevent facility pricing from being introduced in Europe. Similarly, even in grid-like cities such as Los Angeles or Atlanta, centres do emerge making the design of cordons possible in principle. The London Congestion Charge covers less than 3% of the area of Greater London and many cities have congested hubs.

If differences in spatial structure provide a very partial answer at best, what explains the differing approaches to congestion charges?[15] A widely held view is that the philosophy underlying charges differs fundamentally between the US and the EU. Value-pricing reflects the view that the charge is in return for higher quality service, whereas the cordon approach is more in line with an internalization perspective. Of course these points of view are not irreconcilable as such: value-pricing is an approximation to internalization and cordons provide higher-quality travel. Nevertheless, there is a real difference: with a cordon there is no avoiding payment for car users, whereas at a value-pricing facility drivers can decide at the last second which lane to use. Public transport could be seen as a lower quality alternative to car use within a cordon, and using interregional roads instead of tolled highways could be seen as offering a similar choice, but in these cases choices to use either alternative need to be made at an early stage in the trip and involve more than just using one or the other lane on a particular facility. One more real difference between cordons and value pricing is that value pricing up to now has been associated with providing new capacity or with providing access to spare capacity (on HOV lanes) at a premium, whereas cordons are about charging for the use of existing, congested roads. In sum, value pricing emphasizes the supply of premium service while retaining the option for an easy switch to a lower quality alternative; cordon pricing emphasizes maintaining reasonable service levels overall. Economic assessment (e.g. Small and Yan, 2001) shows that both options produce better outcomes than when there is no pricing. Value pricing can perform better than cordon pricing by offering greater discrimination in the choices users can make. This can be used to maximize benefits as users differ in how much they are willing to pay to gain time. Evidently, in terms of making efficient use of the road network, value pricing (partial facility pricing) is outperformed by well-designed full facility pricing schemes.

2.1.6. Pricing parking

When not in use, cars are parked. Parking requires space that could be used for other purposes so there is a resource cost associated with it. Many drivers do not pay the resource cost of parking even when there is a charge for on street parking. They may park for free at work or they may pay a fee that is below resource costs. Most stores with parking lots do not

[15] It could be argued that offering travellers a choice between tolled motorways and free but slower roads for interurban travel (as is done e.g. in France) is a form of value pricing. The principle is not extended to urban areas because building alternatives is too costly. By consequence, congestion will need to be managed by area or cordon pricing. In this sense, one difference between Europe and the US may be that extra capacity in Europe is more (and too) expensive.

charge directly for parking. These parking policies lead to inefficiencies in the transport market, creating excessive demand for underpriced parking spots a resulting in more traffic than if parking was charged for directly. These inefficiencies can be very large: some evidence (Calthrop *et al*., 2000) suggests that the gap between private and social costs per kilometre is larger for the parking inefficiency than for the congestion inefficiency.

In debates on congestion charging this raises the following questions: what would be the effect on congestion if parking were priced differently, and could parking charges be used instead of congestion charges or in conjunction with congestion charges to tackle the congestion externality?

Direct charges for parking would have a number of effects. Congestion would fall, optimal congestion tolls would be lower[16] and the net benefits from congestion tolls would be lower, as congestion is lower to begin with. This is of importance given the high costs of running a congestion charging system compared to the lower cost of modifying parking charges.

That parking charges are relatively cheap suggests they might be used instead of congestion charges to manage congestion in particular locations. Parking charges then would rise above parking costs and congestion would fall because the demand for travel from those wishing to park in centre would fall. How effective such an approach is depends on how much through-traffic there would be, i.e. how many drivers would use roads in congested zones but not park there, and on how much “cruising-for-cheaper-spots” would be induced.[17] With limited through-traffic and little extra cruising, parking charges can mimic congestion charges well and attain similar efficiency gains. However, excessive parking charges would also generate efficiency loses and equity concerns – very high charges for those that park and no charges for those that don’t – reduce the appeal of this approach.

The key message from the debate on parking pricing is not so much that parking charges could mimic congestion charges (that is possible but very much context-dependent), but that restructuring the way parking is paid for would lead to better use of space and to less congestion where capacity is scarce.

2.2. Insights from Singapore and Stockholm

The cases of Stockholm and Singapore were debated extensively at the Round Table, on the basis of background papers on those charging systems.[18] This section presents highlights

16 Congestion would decline because there is less traffic and because fewer drivers would be cruising for parking (a type of driving behaviour that is particularly disruptive to traffic flows).

17 This in itself of course depends on how many cheaper spots would be available, an issue susceptible to public management.

18 http://www.internationaltransportforum.org/jtrc/roundtables.html#RTCongestion_Charging.

from the discussion. It does not contain complete descriptions of the Singapore and Stockholm schemes and does not summarize the background papers.

2.2.1. Singapore

The Singapore congestion charging scheme (see Chin, 2010 for a detailed description) is a demand management system, not a revenue generating device. It is one component of a broad transport policy that also relies on a vehicle license quota, on infrastructure planning and public transport provision to offer high-quality transport options to a growing number of users at a reasonable cost. The vehicle quota system is intended to keep the growth of the vehicle stock roughly in line with the planned expansion of available road space (with allowed growth of the stock recently reduced from 3% to 1.5% per annum). Surveys indicate that car users are mainly interested in high-quality road transport (more road investment, effective congestion management) than in cheaper public transport alternatives.

The approach to system design is pragmatic. The system has become gradually more refined over time, moving from manual to electronic fee collection and enforcement, and covering more of the city as roads become busier, with initially 33 and now 66 gantries for automated control. Charges differ between gantries and vary with the time of day. In 2009, active management of the morning peak through changes in charge levels was extended to cover the evening peak hours too. Rates are revised every three months in order to keep speeds between 45 and 65 km/h on the freeway links in the charged area.[19] Rate changes respond to perceptible changes in congestion levels. Extensions to the system and significant changes in rates are accompanied by major communication efforts. Changes in revenue and concerns over revenue neutrality are of secondary concern, although the scheme is intended to be revenue-neutral in principle. Revenues from congestion charges are dwarfed by those from vehicle licenses, so that changes in congestion charge revenues are not a major policy concern. Revenue neutrality was ensured at the time of the introduction of congestion charges through reduced vehicle taxes (government actually lost revenue as congestion charge revenues were overestimated). Revenues are not earmarked and are in fact lower than road and public transport spending.

The incremental development of the congestion charging system has had the benefit of making the more sophisticated later evolution easier for users to adapt to. Starting with a simple system and making it gradually more complex might be expected to help improve acceptance, but at the same time too simple a system may lead to too large a share of benefits forgone (cf. the discussion on Stockholm). The gradual expansion of the Singapore system reflects pragmatic responses to evolving circumstances more than an explicit strategy.

The Singapore case is atypical in a number of respects, some by virtue of the geographical situation of the city and some by design. First, the number of foreign vehicles

[19] It is likely that these speeds are below benefit-maximizing speeds (i.e. that tolls are below marginal external costs or below second-best cordon levels).

(occasional users) is small and easily identifiable with foreign licence plates, allowing design choices geared towards more frequent local users that may be hard to duplicate elsewhere. Second, the ability to change prices every three months cannot be replicated everywhere -- the example of Stockholm was given, where a rate change might easily take two years because of legal requirements. By contrast, tolls in some value-pricing systems in the US (e.g. the I-15 in San Diego) are adapted every six minutes. This is accepted both in Singapore and California as the goal is to maintain free-flow speeds. Such a rule-based pricing system may be easier to accept than one where prices are determined by discrete decisions. More frequent changes allow better congestion management and they may also serve acceptability as large rate changes can be avoided. Finally, the charging system in Singapore is just one component of a broader system that manages supply (road infrastructure) and demand (vehicle licenses and charges) with a view to what performance levels need to be reached. Ownership policies are more restrictive than in many other countries.[20] Otherwise said, if ownership policies were less restrictive, the congestion charges would have to be higher to attain similar service levels (travel speeds).

The Singapore approach can be read as one where congestion charges are used to fine-tune overall transport prices so as to obtain acceptable service levels throughout the system. Other proposed or existing congestion charging schemes are not limited to fine-tuning to the same extent. Given the weaker degree of integration of the various components of transport policy elsewhere, the mismatch between demand during peak hours and capacity may be larger than in Singapore in many cases.

2.2.2. Stockholm

The Stockholm congestion charging system (see Eliasson, 2010 and Hamilton, 2010 for an in-depth discussion) is effective in reducing traffic volumes and increasing travel speeds.[21] Traffic volumes have declined in all time periods, suggesting there is more trip elimination than rescheduling. Commuters tend to reschedule, but trip purposes are strongly diverse, with about 40% non-commuting trips and many occasional users. Cost-benefit analysis suggests that the system produces net benefits; 75% of gross benefits come from time savings, the remaining quarter mainly from better air quality. The Stockholm evidence also suggests only limited direct gains came from the expansion in public transport expenditure that was part of the overall transport policy reform.

[20] The aggregate level of government revenue raised from road users/car owners in Singapore also is unusually high, at around 20% of total revenue.

[21] There is evidence for Stockholm that volume effects are highly persistent over time while travel time effects tend to wear off to some extent. This has happened because of an increase in road maintenance works and because some road space has been allocated to other use (neither development related to the pricing system), so it is not in itself evidence of declining welfare gains.

Public acceptance at this time is sufficiently broad to expect indefinite continuation of the system's operation.[22] Acceptance is related to congestion reduction, but also to the improvement of the urban quality of life (e.g. less traffic, less pollution) and with the perception that support for the scheme reflects – and signals – green preferences. In this sense, the Stockholm case illustrates that perceived benefits are not necessarily entirely the same as benefits included in traditional appraisal.

With respect to the relation between acceptance, political risk, and system costs[23], the Stockholm case is a clear example of the cost-inflating effect of strong risk-aversion (cf. Section 2.1). In addition, it illustrates how political risk is transformed into risk for the administration responsible for designing and implementing the scheme, and how this administration shifts risk to the companies selected for executing the plans. The administration passes on risk by adding features to the charging system and assigning legal liabilities to contractors who respond by "over-specifying" system components and building redundancy into the system, all of which inflates costs. It deserves emphasis that concerns about acceptance creates these risks. It follows that if acceptance of a project can be won early, risk is lower throughout design and implementation stages, and this allows costs to be cut. If acceptance is won late, as in the Stockholm case, then higher costs are incurred in early stages as a form of insurance. Once it becomes clear that risks do not materialize, insurance expenses can decline, and project costs can gradually be trimmed down.[24] However, higher costs were incurred in the past and irreversible design choices partly drive future costs, so that gradual cost cutting does not allow recovering all expenses associated with high initial risk.

22 Support can decline over time, as evidenced by a Norwegian survey on toll roads. This is because public memory of why the system exists erodes, so that charges – and more strongly charge increases – may meet with increased resistance.

23 Assessing whether a system is expensive or not should not be done using revenue / cost ratios as revenues are endogenous through toll levels.

24 The Stockholm system initially used ANPR as well as transponders. Once it became clear that ANPR was sufficiently reliable – which it became partly because of technology development related to the Stockholm project – transponders were no longer used. This avoided the costs of managing transponders, which were higher than expected.

3. NEW LESSONS FROM TRANSPORT ECONOMICS?

3.1. Refining the basic argument for charges

The basic economic argument for congestion charges is well established. In a nutshell, it says that since travel times increase with traffic volumes, an additional car on the road slows down all other cars, increasing time costs for all the occupants of all the cars. The decision to travel by the occupants of an additional car is based on their own travel costs (their private or internal costs). They ignore any increase in travel costs for all other car users (the external costs).[25] This is inefficient when private costs are below the full social cost of the decision to travel. When decisions are made on the basis of "underestimates" of costs, too much of a good (in this case: travel) will be consumed. A congestion charge is intended to confront users with costs imposed on other users, so as to align private costs with social costs. The charge will suppress part of demand, reduce congestion and increase surplus.

This simple rationale for congestion charges is based on a range of explicit or implicit simplifying assumptions. Research on what happens if these simplifications are dropped is progressing rapidly. The question here is what recent research insights tell us about congestion charging policy. In answering that question, Fosgerau and Van Dender (2010) focus on design issues more than on acceptability concerns. The key message is that allowing for more complexity in the analysis strengthens the economic case for congestion charging.

A first remark on the basic argument is that it relies on a flow model of congestion, where speed declines because distances between cars decline with increasing traffic density. An alternative and at least as relevant model of congestion focuses on bottlenecks, where queues appear when demand exceeds the capacity of some part of the road network. Bottleneck congestion models highlight the possibility of trip rescheduling. In basic bottleneck models, tolls are used to affect users' decisions on departure times, so that queues disappear. Drivers' costs of waiting in line are replaced by toll costs, which generate revenues that are not lost to society in contrast to the costs of waiting. This model illustrates that the rescheduling of trips can generate very large social benefits, a dimension of social

25 It is sometimes argued that congestion is not an externality because transport users ultimately bear the cost of travel. While the latter is true, one can wonder whether the distinction between transport users and the rest of the economy is artificial. Of more direct importance is that who bears the cost is irrelevant: when individual users make travel decisions and they ignore costs imposed on others, there is an external congestion cost.

gains from pricing that is obscured by the standard model of flow congestion. A full appraisal of congestion charges should take explicit account of rescheduling effects.

A second research strand concerns heterogeneity among travellers in terms of their values of time. Here empirical evidence has produced a stylized fact: there are typically many travellers with low values of time and fewer travellers with high values of time. The range of values is huge and the upper end of the distribution has a long tail. The presence of strong heterogeneity has some immediate implications. First, introducing a toll increases the value of time of the average road user by suppressing trips associated with low time values. It follows that the time losses imposed by one driver on other users increase, so the equilibrium toll is higher than would be derived on the basis of the pre-toll average value of time. Because the distribution of values of time among users is not symmetrical, the effect on the average value can be quite large. Second, congestion charging schemes that maximize the number alternative responses, as is the case with value pricing schemes in the US, can be seen as forms of product differentiation that are strongly welfare improving when people differ. The point here is not that value pricing would not produce gains in the absence of heterogeneity (it does) but rather that gains are stronger when there is heterogeneity.

A third body of research focuses on defining and measuring the value of reliability. Whereas the basic argument for congestion charges focuses entirely on how travel time increases as congestion rises, in practice travellers care about expected travel time and about travel time risk. The expected cost of a trip, on which travel decisions are based, is higher when the trip is expected to take longer, and it is higher when the probability of deviations from the expected travel time is larger (i.e. reliability declines). Travel time risk is positively correlated with expected travel time, but it is not the same. Focussing on travel time and ignoring reliability implies underestimating time costs, so again tolls are higher when reliability is taken into account even if the relationship is not simply linear (see ITF/OECD 2010).

Fourth, the understanding of what congestion means in an economic sense is evolving. For example, the standard model of charging systems presupposes that more traffic flow leads to higher travel time. However, in practice hypercongestion can and does occur (the stage at which traffic slows to such a degree that flow - and not just speed - decreases as the rate of vehicle arrivals at a bottleneck increases). Standard traffic models ignore the possibility of hypercongestion, rendering their relevance to the preparation of congestion policy problematic and implying in particular that these standard models underestimate how high tolls should be set to alleviate congestion by any given amount.[26] Again, this points in the direction of higher congestion charges for optimal outcomes.

It should be underlined that the shortcomings of standard traffic models do not mean they are useless. Improved understanding of congestion allows better use of conceptually

[26] Ignoring reverse causality in estimation implies such underestimation.

simple models, e.g. by establishing the direction of error. More sophisticated models can replace the simpler ones when they become available but there is no need to postpone the use of models in policy design until then. It was pointed out repeatedly at the Round Table that model-based judgment on where to locate tolling points and what tolls to charge performs better than common sense judgment.

3.2. Congestion charges in a broader economic context

The basic justification for congestion charges is that confronting travellers with costs they impose on others that they would otherwise ignore improves welfare. The charges remove an inefficiency. If congestion were the only inefficiency in the economy, the argument would be complete. However, as the economic theory of “second-best” suggests, congestion charges potentially trigger complicated interactions with other inefficiencies and these interactions might affect policy recommendations.

An extreme example of such interactions concerns labour markets: if all travellers are commuters that have no choice other than to use their car if they want to get to work and flexitime is not an option, should a congestion charge be introduced if taxes on labour are already high? The answer is no, unless the labour tax were too low to cover marginal external congestion costs. A slightly more realistic model would allow different travel modes and different trip purposes. The core message remains the same: avoid increases of higher effective taxes on labour if that is possible, e.g. by making charges paid for commuting trips deductible from income taxes, as is effectively done in Stockholm.[27] Care should be taken, however, to treat different commuting modes on an equal footing, e.g. season tickets for commuting by public transport are also eligible for tax rebates. Tax systems where commuting tends to receive favourable tax treatment in general may be more suitable to allowing tax deductions of congestion charges than systems where no such treatment exists.

Discussions about deductibility are part of a larger debate on how revenues from congestion charges should be used. As pointed out before, charges generate substantial amounts of revenues, and misusing them could easily dwarf the gains from reduced travel times and ancillary benefits. If revenues were simply burned, no welfare improvement would be possible from congestion charges – time losses would simply be converted to monetary losses. Revenue use should be at the core of the design process as it affects to what extent society is better off overall with charges. How revenues are used has an impact on how benefits from charges are distributed, which in turn has direct links with acceptability.[28]

[27] Deductibility did not exist during the trial but was introduced afterward.

[28] It is sometimes argued that relations between , e.g., labor taxes and transport taxes should not concern policy design too much, given the different political responsibilities However, analysis points out that accepting such constraints may be costly.

There is often tension between what economics suggests concerning revenue use and what is seen as practically feasible and desirable. For example, from a practical point of view it is often proposed to return revenues to car users by reducing other taxes. But car users gain from faster travel as a result of congestion pricing, so they would be overcompensated if all revenue also accrued to them.

4. CONCLUDING REMARKS

Next to the subsidization of parking, the failure to charge for external costs of congestion is one of the main inefficiencies in metropolitan transport systems. Recent economic evidence strengthens the case for using charges to bring congestion closer to efficient levels. The key to successful implementation of congestion charges is to get the policy accepted. Acceptance is dynamic. It can be managed to an extent and depends on a number of factors, including reduced congestion. Ensuring acceptance may require giving up some benefits of a closer-to-ideal system but less-than-ideal systems (simple cordons, value pricing schemes) can still be satisfactory. Rule-based systems for changing prices (e.g. maintaining pre-determined levels of speeds) appear to be more popular than those requiring political discretion. There can be some trade-off between perceived and assessed benefits of charges. The extent to which such trade-offs are made should not, however, be allowed to undermine the core objective of charges – which is to cut congestion. Ancillary benefits, including reduced environmental impacts, can in some cases have an impact on how much to charge and should always be included in assessments, but they are not the principal goal of congestion charging mechanisms.

Congestion charges potentially raise substantial amounts of revenue, but the systems are costly to run as well. This renders statements about revenue neutrality with new congestion charges risky, as equalising gross revenues implies lower net revenues when the unavoidable costs of congestion charging systems are taken into account. In general, emphasizing revenue neutrality may reduce policy flexibility. It may, however, be a requirement for getting public and political support. Transparency and accountability in revenue use is at least as important for acceptance. This, rather than revenue neutrality, was one of the keys to success in London.

REFERENCES

Calthrop E., S. Proost and K. Van Dender, 2000, Parking policies and road pricing, Urban Studies*,* 37**,** 63–76

Chin K.-K., 2010, The Singapore experience: The evolution of technologies, costs and benefits, and lessons learnt, JTRC Discussion Paper 2010-1

Eliasson J., 2010, So you're considering introducing congestion charging? Here's what you need to know: a FAQ based on Stockholm's experiences, JTRC Discussion Paper 2010-4

Fosgerau M. and K. Van Dender, 2010, Road pricing with complications, JTRC Discussion Paper 2010-2

Hamilton C., 2010, Revisiting the cost of the Stockholm congestion charging system, JTRC Discussion Paper 2010-5

ITF/OECD, 2009, Improving reliability on surface transport networks – Summary document, OECD/ITF, Paris

Oehry B., 2010, Critical success factors for implementing road charging systems, JTRC Discussion Paper 2010-3

Poole R., 2009, When should we provide separate auto and truck roadways?, JTRC Discussion Paper 2009-24

Small K. and J. Yan, 2001, The Value of Value Pricing on Roads: Second-Best Pricing and Product Differentiation, Journal of Urban Economics, 49, 310-336