Congestion Offsets: Transforming Cities by Letting Buses Compete

Matthew Bradley and Jeff Kenworthy

Better Cities are Possible

Transport planners used to dream big dreams. Gleaming ribbons of steel and concrete were going to take us, uninterrupted, anywhere we wanted to go. The automobile would allow people to have access to all the good things their cities had to offer, while also enjoying all the benefits of a freestanding house on a spacious suburban block. Sufficient roadway space would be provided so that everyone could share in this bright mobility-based future.[1]

Although this dream has had its share of critics in recent times, it was not without merit.[2] The dream of mobility, what might be called the great freeway project, ran into problems not because of any lack of good ideas or intentions, but because inherently it could not be made to work. In pursuit of this dream many cities drew up detailed transportation plans replete with highly ambitious road building projects. Indeed this was the era of the grand transportation plan, with transportation planning methods initially developed in Detroit and Chicago being rapidly exported not only to the rest of the USA, but also to much of the world.[3] Following such grand transportation plans many cities built extensive road systems, at great expense in terms of both money and damage to urban form (bisected neighbourhoods and widespread demolition of inner city housing and businesses), only to see traffic congestion turn the dream of mobility into something more like one long traffic jam.[4]

Not only has this outcome considerably dampened public enthusiasm for grand road building projects, it has also had its effect academically. In the 1980s a plethora of downbeat journal articles started to appear as numerous countries, but most notably the USA, descended into suburban gridlock; an 'honour' previously reserved for CBD-bound arterial roads. With the extreme decentralisation of work into 'salt and pepper' patterns across the landscape, served primarily by cars along non-radial routes, with hardly so much as a peak period bus service in operation, the Herculean task of trying to match the capacity of road infrastructure with demand simply became too much.[5] During earlier times, much academic focus had been on how we were going to achieve a car-based mobility future. By contrast, since the 1980s we have, to quote the well known transport academic Professor Phillip Goodwin, been trying to solve congestion "when we must not build roads, increase spending, lose votes, damage the economy, or harm the environment, and we will never find equilibrium."[6] Although the mobility dream may have been at best an overly enthusiastic venture and at worst a fool's errand, it was at least a project embarked upon with a positive outlook, not the grim task outlined by Professor Goodwin. Not that he is the most downbeat example. Bleak would be the only word to describe the astonishingly titled Transportation Quarterly article "Metropolitan Congestion: Towards A Tolerable Accommodation".[7] Feel the *joie de vivre.*

This paper rejects the proposition that the best we can do with our road networks is muddle through, or learn to tolerate congestion. Congestion is a solvable problem. Better cities are possible. Most urban networks, such as the telephone or electricity networks, function smoothly and efficiently in many cities around the world; making a telephone call is no more difficult than dialling a number and electricity is available at the flick of a switch. Roads are unusual in that they are the only urban network that can regularly be observed functioning very poorly, even in wealthy cities where all other urban networks are running well. In cities where the water, sewage, power, gas, internet and telephone networks are all running relatively smoothly, it is not uncommon to see road networks completely overwhelmed by congestion.

All urban networks other than roads use one of two techniques to manage 'traffic' on their systems; they overwhelm demand with capacity or they manage demand to keep it below capacity. The classic example of overwhelming demand with capacity is an electricity network designed with sufficient capacity to continue functioning on the hottest day of the year. Demand management is also used on many urban networks, if only inasmuch as networks have a price on their use, thus dampening demand. The great freeway project was an attempt to overwhelm demand with capacity, an attempt which, given widespread congestion, it would have to be said has failed. The alternative to overwhelming demand with capacity is demand control, but this approach is only presently used on a small number of roads, such as toll roads. With one traffic control approach, overwhelming demand with capacity, having failed, and the other, demand control, not in general use, widespread congestion has resulted.

So if the attempt to overwhelm demand with capacity has failed, why not widespread demand control? Why do we run one, and only one, of our urban networks without some protection against congestion? If congestion-proofing all other networks makes sense, why does it not make sense for roads?

The answer given here will be two fold. In the section "How Cars Monopolise Roads and Prevent Buses from Competing" it will be argued that the damage being done by not congestion-proofing roads is greatly underestimated, in that the result is not just a poorly performing road system, but a structurally inefficient urban transport solution. Secondly, in the following section "Simple and Fair Congestion Control", it will be argued that road pricing, the only congestion-proofing method under general discussion, has serious weaknesses. To summarise, this paper will argue that we do not congestion-proof roads because the damage being done is greatly underestimated and because the only congestion-proofing method under general discussion has serious weaknesses.

To address the first part of this problem, that we underestimate the damage being done by not congestion-proofing roads, it is enough to simply show the greater damage; something which will be done in the next section, "How Cars Monopolise Roads and Prevent Buses from Competing". To address the second part of this problem, that the only congestion-proofing method under general discussion has serious weaknesses, it is not enough to just show these weaknesses, an alternative must be offered. The following section, "Simple and Fair Congestion Control", will use an examination of both the strengths and weaknesses of road pricing as a basis from which to develop an alternative congestion control method, namely congestion offsets.

With a key objective for congestion offsets being to retain the strengths of road pricing, some readers may, in due course, end up seeing sufficient similarities between road pricing and congestion offsets that they view congestion offsets as just a variant of road pricing, particularly in cases where road pricing has been packaged along with the subsidising of public transport, such as in London. Although this paper acknowledges that road pricing has significantly informed the development of congestion offsets, it rejects the proposition that congestion offsets are just a variant of road pricing, and, in the section "Simple and Fair Congestion Control", will argue that congestion offsets are a distinct method of congestion control.

As will be discussed, a key difference between congestion offsets and road pricing is the mechanism each approach uses to achieve congestion control. With road pricing congestion control is achieved by constructing a market for road space, and then only selling the quantity of road space that is actually available. By contrast, with congestion offsets congestion control is achieved by altering the regulations under which roadways are managed; sufficiently punishing roadway overuse and sufficiently rewarding roadway underuse so as to reduce the quantity of road space used to the quantity of road space that is actually available. Congestion offsets retain roadways as a commons, with the objective being simply to better regulate that commons. Regulating a commons is something that everyone accepts, even in transport, as for example when one receives a fine for parking on a street beyond the designated time allowance, or for parking on a street in a location where parking is not permitted. By changing the regulations under which roadways are managed, so as to eliminate congestion, it will be argued that our cities could be not just improved, but transformed. In short, this paper will argue that we should not accept a 'tolerable accommodation' with congestion.

How Cars Monopolise Roads and Prevent Buses from Competing

All else being equal, it is not possible for a bus service to provide a faster journey time than a car. Along bus routes, bus services incur two unavoidable time overheads which cars do not; namely slowing down and speeding up from bus stops and waiting for passengers to board and alight. The impact of these factors can be demonstrated using a somewhat simplified model that compares a bus service to a car service, in the same setting. To give buses their best chance in such a comparison, to see buses at their best, cars will be compared to a Bus Rapid Transit (BRT) service, rather than to a traditional bus service. BRT systems maximise the speed performance of buses by using exclusive rights-of-way and by measures such as level boarding and pre-boarding ticketing. For cars, the model presented will, for now, assume roadways free of congestion, thereby also showing cars at their best. In the real world, cars and buses are often not seen at their best; buses normally operate without the benefit of an exclusive right-of-way, and cars are, in many cases, far from free of the effects of congestion. The impact of these real world factors on the service speed of cars and buses will be addressed later in this section. For now, a model of a simplified notional transport system between just two stops is shown in Figure 1.[8] This model is sufficient to show the lower speed performance of buses, as compared to cars.

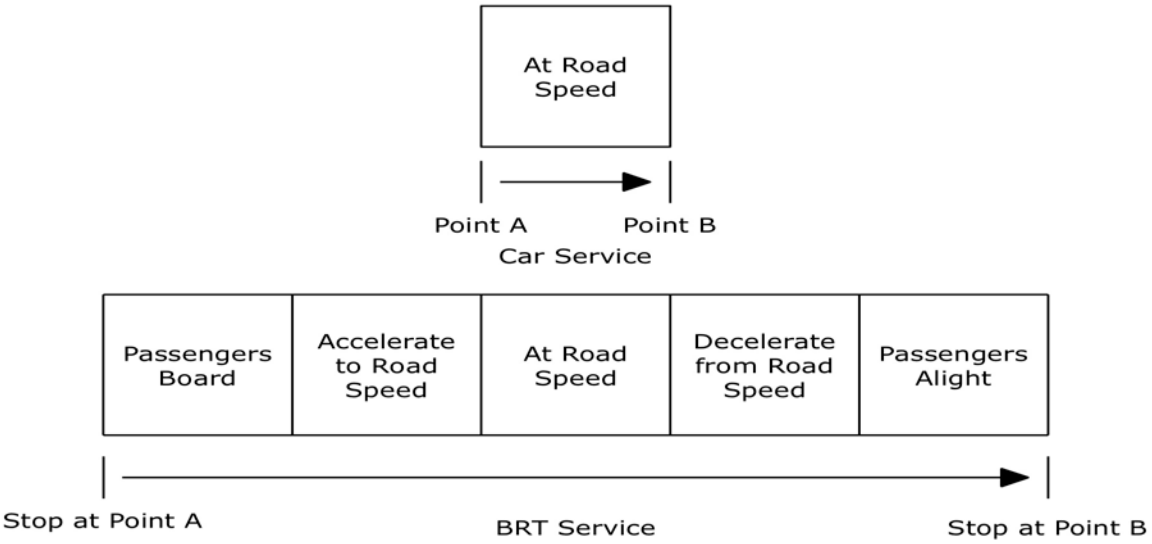


Figure 1: Comparison of Car and BRT Services Between Two Points

To be able to determine the comparative speed of the car and BRT services in Figure 1, it is necessary to know the speed of the roadway and BRT line, the distance between the two points, the bus boarding and alighting times and the bus acceleration and deceleration times. A major update of the *Millennium Cities Database for Sustainable Transport* shows that in forty-three cities covering the US, Canada, Australia, Europe and selected Asian cities, the average roadway speed in 2006 was 40.9 km/h.[9] For the model presented here, this figure will be rounded down, and 40 km/h will be used as the roadway speed for cars and for the speed of the BRT line. The distance between points A and B will be 800 metres, which is approximately the average inter-stop distance for BRT systems surveyed in the *Bus Rapid Transit Planning Guide*.[10] This 800 metre value is also very close to the often-stated value of half a mile between stations or stops for the transit-oriented development planning approach.[11]

For boarding and alighting times, each will be 12 seconds, approximately half the average dwell time of the BRT systems surveyed in the *Bus Rapid Transit Planning Guide*.[12] Bus acceleration and deceleration will both be 1 m/s2; typical average performance numbers.[13]

The roadway speed number is sufficient to calculate the speed of the car service, which is 40 km/h, the same as the roadway speed. Calculating the speed of the BRT service is somewhat more complex. With the distance of 800 metres between stops being known, the average speed of the BRT service can be determined by calculating how long a bus would take to travel between the stops. The total time is the sum of the boarding, alighting, acceleration and deceleration times, plus the time spent at the roadway speed. At the previously given acceleration and deceleration figure, a bus would reach the BRT line speed of 40 km/h in approximately 11 seconds over approximately 62 metres, and decelerate from the line speed over the same time and distance. With acceleration and deceleration both taking 62 metres, this leaves 676 metres of the 800 metre distance between stops occurring at the BRT line speed of 40 km/h, a distance that would be covered in approximately 61 seconds. Add to this time the combined acceleration and deceleration time of 22 seconds and the combined board and alight time of 24 seconds, then the total time taken to travel between the stops is 107 seconds. A distance of 800 metres covered in 107 seconds gives a speed of approximately 27 km/h, so about 67% as fast as a car.

The update of the *Millennium Cities Database for Sustainable Transport* shows, using the forty cities for which data are available, that the average ratio between bus speed and car speed is 54%.[14] The worst performing city is Chicago at 30% and the best performing city is Hamburg at 89%. To compare this real world data to the theoretical outcome produced by the model, roadway speed must be taken into account, as the model is based on a speed of 40 km/h whereas the forty real world cities have average roadway speeds ranging from 24.0 km/h for Berlin to 57.5 km/h for Chicago. Slower roadway speeds make it easier for buses to compete with cars, as the buses have to spend less time getting up to and down from the roadway speed, and also because board and alight times become a smaller proportion of the total trip time. Figure 2 shows, across the speed range 0 km/h to 60 km/h, the ratio between bus and car speeds produced by the model.

As can be seen in Figure 2, when all other factors are kept constant, slower roadway speeds result in better performance ratios for buses. If the cities from the Millennium Cities Database are compared to the model, at the average roadway speed for each city, an idea can be gained of how well the bus systems in these cities are performing. Figure 3 shows the forty cities from the Millennium Cities Database and the percentage by which they underperform

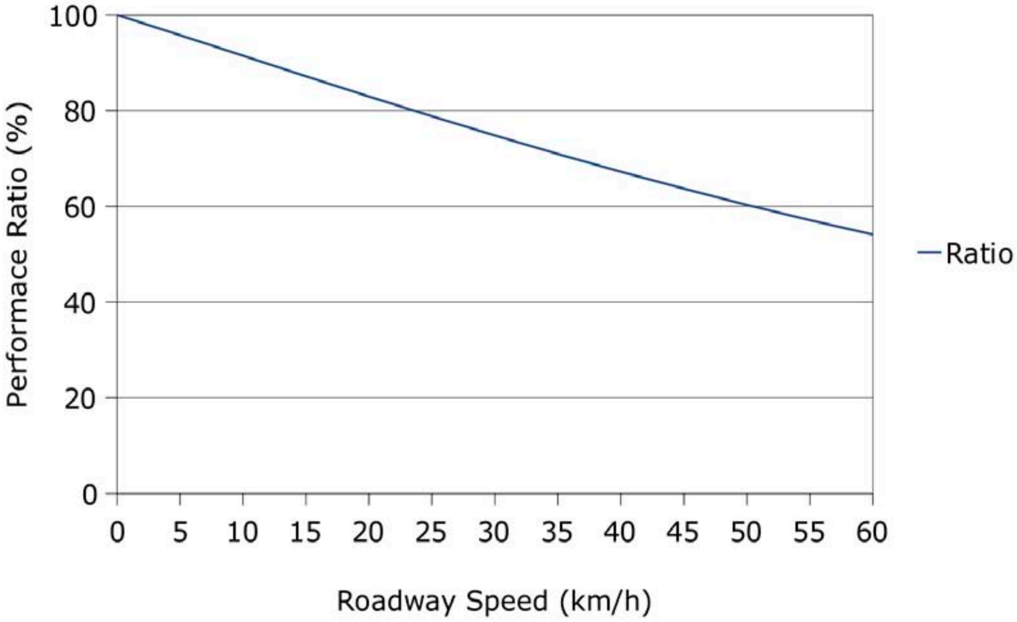
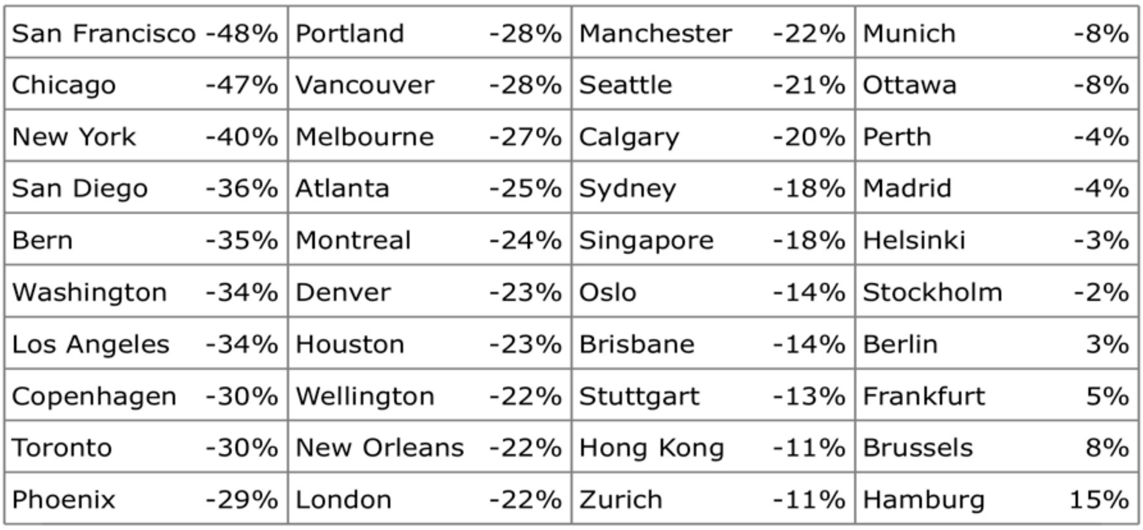


Figure 2: Ratio of Bus Speed to Car Speed Across a Range of Roadway Speeds



San Francisco	-48%	Portland	-28%	Manchester	-22%	Munich	-8%
Chicago	-47%	Vancouver	-28%	Seattle	-21%	Ottawa	-8%
New York	-40%	Melbourne	-27%	Calgary	-20%	Perth	-4%
San Diego	-36%	Atlanta	-25%	Sydney	-18%	Madrid	-4%
Bern	-35%	Montreal	-24%	Singapore	-18%	Helsinki	-3%
Washington	-34%	Denver	-23%	Oslo	-14%	Stockholm	-2%
Los Angeles	-34%	Houston	-23%	Brisbane	-14%	Berlin	3%
Copenhagen	-30%	Wellington	-22%	Stuttgart	-13%	Frankfurt	5%
Toronto	-30%	New Orleans	-22%	Hong Kong	-11%	Brussels	8%
Phoenix	-29%	London	-22%	Zurich	-11%	Hamburg	15%

Figure 3: Real World Performance Compared to the Model

or outperform the model. Negative values show cities that are underperforming the model and positive values ones that are outperforming it.

With the average road speed for each city used as the speed for the model, only four cities outperform the modelled BRT system, namely Berlin, Frankfurt, Brussels and Hamburg. Three of those four cities, Berlin, Frankfurt, and Brussels, only outperform the model slightly, with ratios that are three, five and eight percent higher than the model respectively. The only city to significantly outperform the model is Hamburg, which has a ratio 15% higher than the model. An assumption of the model is that cars and buses operate at the same speed, which is not the case in Hamburg. Due to the total removal of its light rail system, Hamburg now has most major bus routes operating in the middle of wide roads on dedicated bus lanes, and it also has pre-emptive green-wave traffic light technology in operation. Although some caution is warranted when comparing real world and theoretical data, the fact that only 10% of the real world cities outperform the theoretical model suggests that the presented BRT system is a higher than average performance bus system. This is as intended; the model presented here aims to show buses at their best, not to show an average bus system.

In terms of general trends, in Figure 3 it can be seen that the cities with the best performing bus systems are predominately in Europe, with eight of the ten best performing cities being European. It was in such European cities that the grand transportation plan era was at its most limited, and where serious efforts have been made to have transit as a competitive travel option. The cities with the worst performing bus systems are predominately in the USA, with seven of the ten worst performing cities being US cities. The USA is where the great freeway project achieved its zenith and where efforts to accommodate and provide for competitive transit systems, for example through the provision of bus lanes, reached their nadir.

For both the modelled case and the empirical data, the data represents only the transit line speed; the ‘in-vehicle time’ from the perspective of a passenger. Neither the average ratio of 54% bus speed versus car speed for the real world data, or the 67% ratio for the model include time spent waiting for a bus to arrive, or the speed difference between walking to a main road to catch a bus, as compared to driving there. Given the figures used in the model, the best a BRT system can deliver is a speed 67% as fast as a car. To achieve this outcome, a passenger would have to be travelling from a location right on top of a BRT stop to a location right on top of another BRT stop and arrive at their departure stop exactly as their bus is about to depart. This is far from the typical case. Normally, passengers have to walk some distance to catch their bus service, wait for that service, and then, after alighting their bus, walk to reach their final destination. Normal bus use is door-to-door, not stop-to-stop (or, in the case of traditional bus services, door-to-door, not kerb-to-kerb).

To calculate the speed impact of these two additional factors, getting to and from BRT stops and waiting for services to arrive, the model presented must be expanded to be a model of a full urban transport system. Based on the 800 metre distance between stops, the model will be expanded to consist of an 800 metre grid of major roads along which BRT lines also run, with a BRT interchange stop at every intersection; a model that is similar to the Squaresville model of Dr. Paul Mees.[15] From the US 2009 *National Household Travel Survey,* the average trip length for this model urban transport system will be, as near as possible, 15.7 kilometres.[16] On an 800 metre grid, the average distance travelled to get to a stop, assuming uniform population density, would be 400 metres at both the departure and arrival end, so 800 metres in total. To, again, give buses their best chance in this comparison, a transit-oriented development style of urban planning will be assumed, with sufficient focusing of development around each transit stop to reduce total walking distance to 200 metres at both the departure and arrival end, so 400 metres in total, or half the distance at uniform density. This leaves 15.3 kilometres of the 15.7 kilometre distance to be travelled on the BRT system, which, at 800 metres between stops, is 19.125 stops. For this model, 19 stops will be used, yielding a total trip distance of 15.6 kilometres. As for passenger wait time; this is determined by bus service frequency. The service frequency for BRT systems surveyed in the *Bus Rapid Transit Planning Guide*, averaging the available peak and non-peak figures, is one bus on average every six minutes, which gives an average wait time of three minutes.[17]

The car speed on non-arterial roads will, for this model, be assumed to be 20 km/h, an aggressively traffic calmed level. At this speed, the 400 metre trip to and from an arterial road, 200 metres at both the start and end of the journey, would take a total time of 72 seconds. The remaining distance of 15,200 metres on arterial roads running at 40 km/h would take 1,368 seconds, yielding a door-to-door travel time of 1,440 seconds and, over the 15.6 kilometre total distance, an average trip speed of 39 km/h. As before, calculating the speed of the BRT service is somewhat more complicated. Also in a similar fashion as before, with the door-to-door trip distance of 15.6 kilometres being known, the average speed of the BRT service can be calculated by calculating the total time a person would take to make this trip using the service. The total time is the sum of the time passengers spend walking to and from their departure and destination BRT stops, the time spent waiting for bus services to arrive, plus the time spent traveling on the BRT system itself. With regards to wait times, a reasonable assumption for a grid of BRT lines is that a person would have to wait twice for a bus, once on a horizontal line and once on a vertical line. At the previously mentioned figure of an average wait of three minutes for a bus, the total wait time would be six minutes. As for the 200 metre walking trip at each end of the journey, at a brisk pace of five km/h this distance of 400 metres would be covered in 288 seconds. Finally, the time spent on the BRT system itself needs to be calculated. From previous calculations, the time between stops is 107 seconds, which yields, for the 19 stops being travelled, a BRT travel time of 2,033 seconds. Add all these values together, and the total 15.6 kilometre journey takes 2,681 seconds, at an average door-to-door speed of approximately 20.9 km/h.

So, in a congestion free environment, a BRT system, on a grid of BRT lines spaced at a typical transit-oriented development level of 800 metres with development focused around stops, is able to provide a door-to-door travel speed approximately 54% as fast as a car. Slight further improvements might be possible to this theoretical outcome via the running of express buses, though their speed impact tends to be curtailed by the concomitant increase in wait times at stops, as passengers watch other people's express buses whoosh by. Analysis showing, in an uncongested setting, BRT delivering door-to-door travel speeds only a little better than half as fast as a car, is not the slightest bit controversial. There is no analysis anywhere showing BRT systems being even close to competitive with cars in terms of speed, in an uncongested setting, because there cannot be such analysis. With regards to door-to-door trip times, there are four areas in which buses incur time costs that cars do not; namely time accelerating from and decelerating to stops, time boarding and alighting, time waiting for buses to arrive, and the extra time needed to walk to and from bus stops.

There are no areas in which cars incur time costs that buses do not. As a consequence, regardless of how the figures are 'tweaked' in a model such as the one given, no model whose method and data has any grounding in the real world is going to show anything except even very good BRT systems only providing a door-to-door speed around half that provided by a car, in an uncongested setting. Indeed, the overall speed outcome from the model, which has BRT providing a service speed 54% as fast as a car, is based on a transit line speed that is 67% as fast as a car. Only seven out of the previously discussed forty cities from the *Millennium Cities Database* update had a bus performance outcome greater than 67%.[18] All of these cities are European, and all have an average roadway speed lower than the 40 km/h speed used for the model, with the average roadway speed for this group of seven cities being 29.4 km/h. As previously discussed, lower roadway speeds result in better bus performance ratios.

It is the above type of analysis that underpinned the great freeway project; the post-war attempt to provide automobile mobility to all via vast road building projects. After all, if it can be demonstrated in a few pages that buses cannot possibly compete with the speed of cars, why not deliver the 'best' option to all. It is, though, an assumption of the above analysis that there are no capacity constraints and therefore there is no congestion. If analysis such as that above is conducted with capacity constraints, a very different picture emerges.

Consider, for example, a typical urban road. The capacity of the road lanes that make up such roads depends on many factors. Here, a capacity figure of 700 passenger cars per hour per lane will be used; the mid-point of the capacity range given in *Urban Transit Systems and Technology* for private autos on a street (as opposed to on a freeway) which is "heavily loaded but somewhat below capacity."[19] At an average occupancy of 1.67 people per car, the average level recorded by the US 2009 *National Household Travel Survey*,[20] one lane with a capacity of 700 passenger cars per hour (pcph) can service a load up to 1,169 people per hour using cars. For the sake of simplicity, the measure "people per hour" will be referred to, from here on, as passengers per hour (pph), the metric used for public transport. When reporting values for cars, all passengers per hour levels will, of course, include the driver of a car. The occupancy level being used here is somewhat generous since it is an overall average occupancy for cars, which includes more heavily loaded cars on weekends and other off-peak times when people go out together for social purposes and roads are often less busy. Roads are most heavily loaded during peak or near peak times, when average occupancies tend to be only a little over one person per car, due primarily to the widely acknowledged issue of single occupancy driving when traveling to and from work.

By contrast to the 1,169 pph capacity using cars, a BRT service could support far higher loads in a single lane. For example, a simple BRT service might consist of one all-stop and one express bus service, both running at three minute intervals. If such a service were run using typical twelve metre buses, then the crush-load capacity of each bus would be of the order of 85 passengers.[21] Unlike car services, the normal mode of operation for buses is for strangers to travel together and so all of the capacity of these buses could be utilised. The maximum capacity level of the example BRT service would only be touched for one moment and at one point along its route, with other times and areas running at lower levels due to passenger load variation, but this is also what occurs with at-capacity car services. Two services each running at three minute intervals, gives a total service level of 40 buses an hour. With each bus having a maximum capacity of 85 passengers, the overall BRT service would have a maximum capacity of 3,400 pph.

The 40 buses an hour needed to run this 3,400 pph bus service would only use a very small proportion of the capacity of a BRT lane. By comparison, any attempt to service this 3,400 pph load using cars on the example 700 pcph roadway lane, which is heavily loaded when servicing 1,169 pph, would very quickly reduce the lane from full speed to a zero or near-zero speed. Using the previous door-to-door journey speed calculation method, Figure 4 shows the speed of a car service operating from full roadway speed, the previously used 40 km/h figure, to the roadway being jammed, at 0 km/h. The roadway speed scale in Figure 4 is shown, left to right, from highest speed to lowest speed, reversing the typical presentation, so as to show the impact on service speed as congestion increases. For comparison purposes, the speed of the example BRT service is also shown, unaffected by the car-related congestion degrading the speed of the roadway that the cars are running on.

Presented separately, the superior speed performance of cars at higher car roadway speeds, which would occur at lower load levels, is clear. Also clear, though, is the superior speed performance of buses at higher loads, once congestion has taken its toll and the speed of the car roadway has dropped. In this example, once the roadway speed has dropped to below approximately 21 km/h, the example BRT service is providing a faster door-to-door journey time than a car. But cars and buses do not typically run independently as is shown in Figure 4; in the general case, buses share a roadway with cars. By running buses in an exclusive lane, BRT in many ways has more in common with heavy rail, as those who conceived this mode intended, than it does with buses on normal shared-use roads. Indeed, a strong case could be made that the running of buses in an exclusive lane, rather than in traffic, is the defining feature of BRT, and the reason we have a separate name for this transit mode. Figure 5 shows what happens when the car and bus services are run on the same roadway; in other words what happens when the BRT buses are taken off their exclusive right-of-way and mixed in with car traffic.

With cars and buses sharing the roadway, there are no points at which buses outperform cars. As can be seen in Figure 5, although it is the car users who are causing the congestion, the bus users are bearing the cost of congestion along with the car users. Operating in a congested environment imposes two costs on bus services, the time cost to passengers of longer journeys and the extra cost to operators of running a slower service, with more buses and drivers being needed to service routes at lower speeds and, of course, higher generalised operating costs such as fuel and wear and tear. By bearing a part of the burden of the congestion caused by cars, buses provide a worse service and cost more to run. Instead of cars be-

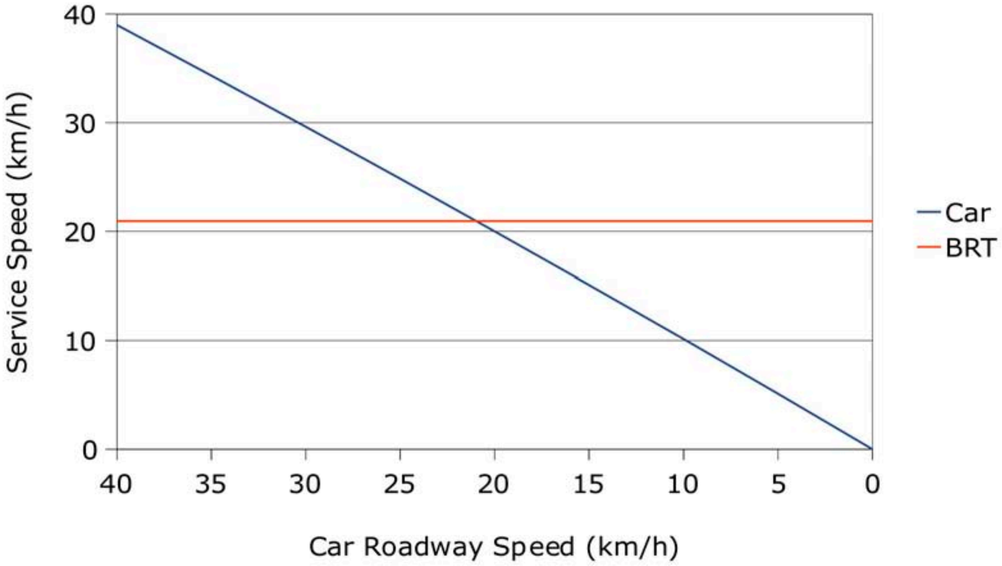


Figure 4: Car and BRT Service Speed on Separate Roadways

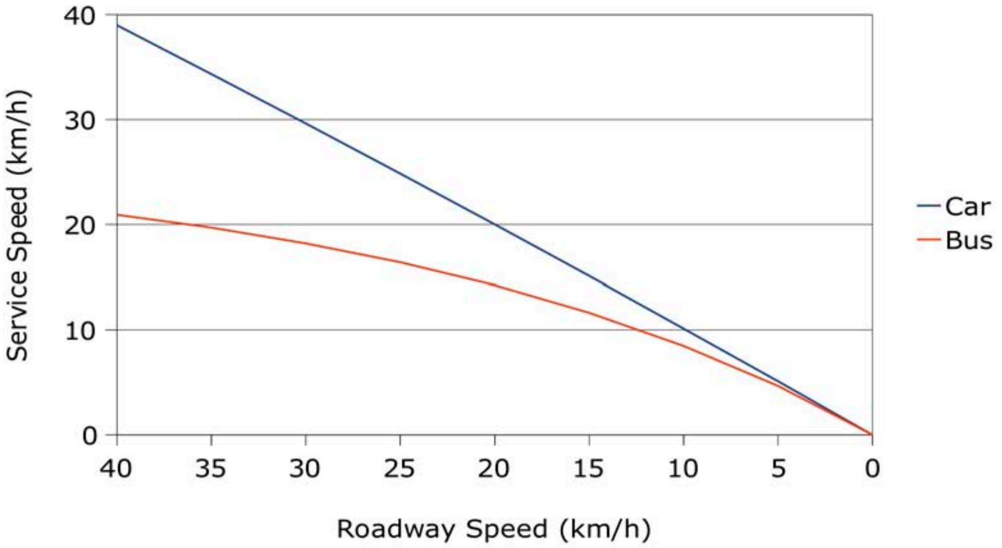


Figure 5: Car and Bus Service Speed on a Shared Roadway

ing faster at lower loads and buses being faster at higher loads, cars become faster at all loads by offloading part of their congestion cost onto buses. In effect, the car mode has been allowed to 'poison' the bus mode.

The role of the bus mode, the mode crippled by cars, is not one that the car mode can effectively substitute for. At a level of 1,169 pph, where cars are heavily loading a single lane, there is just enough load to run a high frequency bus service. For example, a load level of 1,169 pph would fill approximately 14 standard twelve metre buses to capacity every hour, yielding a service frequency of one bus approximately every four minutes.[22] To show the amount of road space that would be used by a 14 buses per hour service, it is necessary to know how much road space a bus takes up compared to a car. From the *Transport Research Board's Highway Capacity Manual*, buses will be calculated here as taking up twice the road space of cars, the mid-point between the figures for level and rolling terrain.[23] Using this conversion rate, a 14 buses per hour service would require 28 of the 700 pcph 'slots' available on the example road lane. In other words, the described 14 buses per hour service would use up 4% of the available road space, one-twenty-fifth of the road space required by cars to service the same load. Another way of looking at this is to say that to provide the same service using cars would require twenty-five times more roadway space than using buses.

The capacity range that buses on normal roads can comfortably service, and the percentage of roadway used, is a critical issue. If you reduce the above described bus service from 14 buses an hour to seven buses an hour, one bus approximately every nine minutes, you reduce the load serviced to 595 pph, and the percentage of roadway used to 2%. At this nine minute frequency, you are close to the border of where people stop just turning up and waiting for a bus, and start looking at timetables;[24] a characteristic that will be used here to mark the low end of the load range which buses on normal roads can be said to comfortably service. At the high end of the load range, if you increase the frequency of the all-stop service to one bus every three minutes and run three parallel express services at the same service frequency, you now have a 6,800 pph bus service using approximately 23% of the available roadway slots. Use 18 metre, 110 passenger buses,[25] which will use slightly more roadway space, then capacity increases to 8,800 pph. This 8,800 pph level is in the same general capacity range as a three lane freeway running just cars, reflective of the massively better roadway utilisation of buses in comparison to cars. Although, in principle, the load that buses on normal roads are able to service could be pushed higher than this level, it is at these higher load levels that BRT and rail solutions begin to come into their own.

In terms of infrastructure, buses on normal roads are able to comfortably service mid-level loads using only a small slice of the roadway capacity of one lane. In the above examples, 595 to 8,800 pph loads were serviced using approximately 2 to 23% of the capacity of a heavily loaded roadway lane. If the approximate mid-points of these values are taken, then buses on normal roads can be characterised as a transport mode that can service loads of the order of 5,000 pph, and which require approximately one-eighth of the capacity of a heavily loaded road lane to do so. Attempt to service these mid-level loads with cars, and the roadway resources required are of the order of twenty-five times higher, due to the car's inefficient use of roadway space; hence the chronic congestion problems of the car mode. Attempt to service these loads using BRT and, with the defining characteristic of BRT being the use of an exclusive right-of-way, you incur 100% of the road lane infrastructure cost, eight times that used by buses on normal roads. Attempt to service these loads with heavy rail and you again incur 100% of the right-of-way infrastructure cost, plus you are trying to service mid-level loads with a mode that excels at high-loads. As with cars, in servicing mid-level passenger loads, BRT and heavy rail also have massively higher right-of-way infrastructure needs than buses on normal roads, due, perversely, to their passenger carrying efficiency. BRT and heavy rail take 100% of a lane and then, at mid-level passenger loads, use that lane so efficiently that most of it is left unused. Cars on normal road lanes, and buses and trains in exclusive lanes, both, for divergent reasons, use vastly more right-of-way infrastructure when servicing mid-level passenger loads than buses on normal roads. From the previous calculations, cars use twenty-five times the right-of-way infrastructure of buses on normal roads, and both BRT and heavy rail use eight times as much.

Heavy rail, the running of trains in an exclusive right-of-way, is not, of course, the only way to operate trains; trains can also be run on normal roads. Such services are known variously as trams or light rail, and although some such systems run in sections at least on exclusive lanes, others run on normal roads, mixed in with traffic. Light rail services differ from bus services in a number of important respects. Perhaps the most important of these differences is that light rail vehicles have a maximum passenger capacity far higher than that of buses. *Urban Transit Systems and Technology* lists light rail vehicles as having a maximum passenger capacity more than six times that of buses, with the maximum capacity of regular buses being listed as 120 passengers, whereas multi-carriage light rail vehicles are listed as having a maximum capacity of 750 passengers.[26] Not only does this greater capacity mean that light rail can service the same passenger load as buses using less drivers, it also reduces all costs that increase with fleet size, such as vehicle and driver scheduling costs. The ubiquitous use of either pre-boarding ticketing or conductors to issue tickets is a further advantage of light rail, though, in this case, one that could be adopted more widely by traditional bus services. For all of the differences between buses and light rail, there is one stark similarity; when light rail is run in traffic it has exactly the same car-caused congestion problems that buses on normal roads have. In a situation where congestion was controlled, the performance of light rail systems running in traffic would be enhanced, and much of the discussion in this paper, which applies to buses, may also apply to such light rail systems.

To summarise, currently all of the public transport options for servicing mid-level loads either have their speed performance crippled by the congestion caused by cars, or have their financial position seriously degraded by having to pick up 100% of their right-of-way cost, eight times the average amount required by buses on normal roads to service mid-level loads. Due to the congestion caused by cars, cities lack a really effective mid-level transport mode, one that is both reasonably fast and reasonably cheap.

Cars are normally occupied by one or two people, buses by many tens of people and heavy rail trains by many hundreds of people. The proper role for cars, buses and heavy rail trains as, respectively, low, mid, and high load services could not be more apparent. It is a simplification to say that cars work best between 0 and 1,000 pph, buses work best between 1,000 and 10,000 pph and heavy rail trains work best between 10,000 and 100,000 pph; but it is not much of a simplification. Although, with multiple lanes, cars can provide many times the previously calculated 1,169 pph capacity of a single lane, such multi-lane roads can be seen, in many situations, as merely a desperate attempt to keep the low capacity car mode functioning, an issue that will be discussed in greater detail in the next section.

The 1,000 to 10,000 pph mid-level load range is not a minor one; vast swathes of the urban transport task occur in this range. Cars, though, via the mechanism of congestion, have stopped buses on normal roads, the obvious transport mode for servicing these mid-level loads, from competing on its merits. Consequently, a capability chasm has opened up between low and high loads, with no mode being available to effectively service mid-level loads. The result has been disastrous. The story of many modern cities, particularly low to mid-density cities, is the car mode running well over capacity and the heavy rail mode running well under capacity, as transport planners desperately try to make each mode service mid-level passenger loads that neither is suited to. A train line operating at a fraction of its potential capacity, running alongside a heavily congested arterial road, is the signature image of the modern transport planning dilemma. This dilemma will not be resolved until cars are no longer permitted to offload their congestion costs onto buses, or, in other words, until buses are allowed to compete.

Simple and Fair Congestion Control

What would lead a person to believe — even for a moment — that a sensible approach to urban transport was possible without some form of roadway congestion control? For cities without railways, roads are the entire motorised transportation solution, and even for those cities that have significant rail systems, roads still play a critical role. For example, London, home of the well-known London Underground, as well as a substantial surface rail system, still had 67% of its motorised travel happening in private cars as of 2006.[27] Few if any cities have such widespread alternatives to road transport that they can afford to let their road networks operate in a dysfunctional manner. Nevertheless, such operation is the norm. It is, after all, dysfunctional to run a network, any network, where the service provided regularly collapses due to load exceeding capacity, without instituting some form of congestion control. Furthermore, in both the popular and academic literature, there seems to be a distinct absence of either lay people or planning experts rushing to defend congestion as the sensible outcome; when congestion is talked about it is normally in tones of condemnation not praise. So if, as a matter of logic, it would seem running road networks without congestion control makes no sense, and if congestion is widely disparaged, why do the overwhelming majority of cities continue to run their road networks without congestion control?

The answer that will be given here is that there is only really one idea, road pricing, in the 'marketplace of ideas' about roadway congestion control, and that road pricing has two serious weaknesses, which will be discussed.[28] There are many ideas in the related disciplines of urban planning and transportation planning — transit-oriented development, balanced transport, sustainable planning — but, other than road pricing, none of them are directly aimed at congestion. None can show, as road pricing can, in a short, clear and compelling manner how chronic congestion could be eliminated. Ironically, the free market idea as to what to do about congestion lacks competition.

That the economics profession has something significant to say about congestion is not surprising; the management of scarcity is commonly said to be the central focus of economic thought.[29] Roads are a scarce resource and so, as a matter of inevitable economic logic, if you do not implement a rationing scheme then queuing (i.e. congestion) becomes your default rationing scheme. Not having a formal rationing scheme in place does not, after all, change the scarce nature of the roadway resource. Economists highlight the fact that the use of queuing as a rationing scheme leads to a wildly inefficient allocation of roadway resources. For example, an eye surgeon, whose time might be billed out at hundreds of dollars an hour sits in traffic the same as, for example, a junior office clerk, whose time might only be billed out at fifteen dollars an hour. The economic solution to this problem is road pricing,[30] a market based approach, which is really what you would expect any mainstream economist to say, regardless of the road context. Present to an economist a general problem where the primary context was resource scarcity, and the primary symptom was queuing, and then ask for a 'diagnosis' and they would likely say that the correct course of treatment is to establish a market and may also comment that you have a common, but readily treatable ailment, called socialism.

The basic mechanism of road pricing is that people pay for the road space they use. If the objective is simply congestion control, then these payments may be limited to times and places where there is a congestion problem, with free access to the roads allowed otherwise. This mechanism has one key strength, namely that it works. Set a price for being on a road, with higher prices during times and places of higher demand for road space, and traffic levels will be reduced. Make the price sufficiently high and congestion will be eliminated. No one needs to be convinced of this, it is both intellectually obvious and can be seen on a number of tollways which have been built above free access roadways, where the tollway is running below capacity and the free access roadway is hopelessly congested. Along with the key strength of road pricing — that it works — it will be argued here that road pricing has two key weaknesses, one political and one organisational. The objective of the rest of this section will be to outline these two weaknesses and thereby develop an alternative approach, namely congestion offsets, which retains the key strength of road pricing — that it works — but which does not have its weaknesses. Both the weaknesses of road pricing relate to the mechanism of its operation; that money is paid to buy access to road space.

Paying for something is an implicit acknowledgement that you do not own what you are paying for. People do not pay rent on a house that they own; as to do so would be to accept that they do not own it. Local council rates are generally accepted, but these payments are used to provide local services, not for the right to reside at a house, as such. In a similar manner, to pay to use road space is to implicitly acknowledge that you do not own that road space. As with local council rates, road taxes are widely accepted as these payments are used to maintain the roads, not for the right to use the roads as such. Currently roads are owned by the people and held in trust by the government; they are a commons. Road pricing involves transferring ownership of the roads from the populous to either the government or private interests, who then charge people for access to what was previously theirs. People's dislike of being dispossessed of what is theirs, and therefore of road pricing, would be considered by some to be proper, and really should be considered by most to be expected. Land is not a second-tier issue, and a strong case can be made that public roads and streets represent the last great area of common land. Whether ownership of this land is transferred to the government or to private interests, as required by road pricing, both options involve people's disenfranchisement from arguably this last great commons. This requirement for a loss of ownership has made road pricing all but impossible to sell politically, and is a key political weakness of road pricing as a congestion control method.

An alternative to the loss of ownership required by road pricing is to keep the roads as common property, held in trust by the government, but to change the regulations under which that commons is managed. The current roadway regulations, most everywhere, distill down to saying: "you may take more than your fair share of available road space." Given that not taking more than your fair share is such a basic moral principle that it tends to be one of the first taught to young children, it is little short of bizarre that we run our roads this way. What might the roadway commons look like if it were instead managed under the principle: "you may, on average, take only your fair share of available road space?"

At the moment people claim road space by taking a vehicle onto a roadway. If that vehicle is a car, one road slot is claimed; for larger vehicles, such as buses and trucks a greater number of slots are claimed, and for smaller vehicles, such as motorcycles, a lesser number of slots are claimed. Congestion occurs when the number of road slots claimed is in excess of the number of road slots that are available. The number of people inside a vehicle has, in the general case, no effect on the amount of road space that can be claimed. It is difficult to see how an approach based on vehicle size could even begin to meet any criteria of fairness; it is just by default what happens. Instead, consider the simplest notion of a fair share; that a fair share is an even share. In the classic example, if there is a cake to be divided up between people then it is 'fair' that everyone gets an even slice. This is a simple approach, and it does have its weaknesses. For example, what if one of the people who the cake is being shared between is currently starving to death; should that person only get an even share? Similarly, in a transport context, few people object to disabled drivers receiving privileged parking access. Nevertheless, an even slice, an even share, is a widely used moral principle, and will be used here to examine what taking one's fair share of road space might involve. How does road use look if "a fair share" is taken to mean that people are entitled to an even slice of a roadway they are on?

To examine this question, it is useful to return to the roadway capacity example of the previous section, where one lane had a capacity, when heavily loaded, of 700 pcph. If such a road lane is loaded equally with people travelling in cars and those travelling in buses, using the same occupancy levels and other values as the previous section, how much road space is used by each group? To load the example single lane to capacity equally with car and bus users requires approximately 2,250 roadway users per hour, approximately 1,125 users for each mode. With the users so split, Figure 6 shows the road space utilisation levels for these 2,250 roadway users on a road lane that has a capacity of 700 pcph.

As shown by the average bar in Figure 6, when a roadway lane with 700 available slots is loaded with 2,250 users, the fair share of road space is approximately 0.311 slots per user. Figure 6 also shows that car users are using far more of the roadway than the average, at a level of approximately 0.599 slots per user, and that bus users are using far less, at a level of approximately 0.024 slots per user. Even if the bus occupancy level is reduced to be the same as that for cars, based on the assumption that a car has five seats, car users still come in well above the average, and bus user still well below. At an equal 33% occupancy, the car users' slot use remains at approximately 0.599 slots, and the bus users' slot use increases from approximately 0.024 to 0.071 slots. So, even with equal occupancy, car users are still taking approximately eight times the road space of bus users. The overuse of road space by car users can be seen more clearly in Figure 7 which, using the same occupancy levels as those for Figure

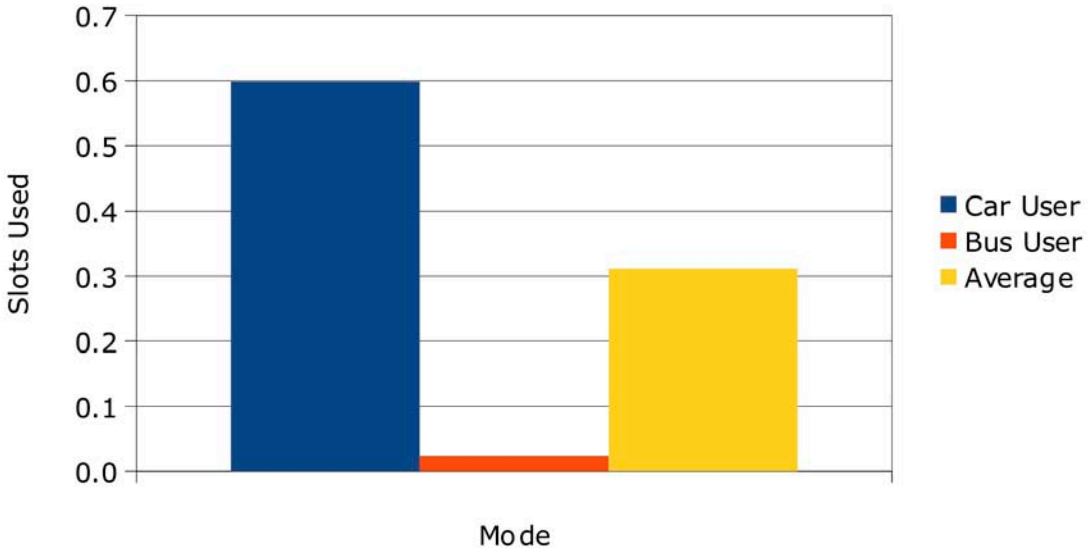


Figure 6: Use of Slots by Mode

6, shows the difference from the average, with positive values showing slot overuse and negative values showing slot under-use.

In Figure 7, it can be seen that, for this example, car users are taking approximately 0.288 slots more than their fair share and bus users are taking approximately 0.288 slots less than their fair share. At this point, with the roadway heavily loaded, each bus user is generously 'donating' approximately 0.288 slots to a car user. Even in the absence of congestion, as is the case here, the above situation is highly unfair. The car users' congestion-free trip is only being made possible by the bus passengers' decision to take a bus, thus freeing up road space which allows for the car users roadway overuse. Tip over the line into congestion, and not only are car users taking more than their fair share of roadway space, they begin dumping part of the cost of the congestion they are causing onto bus users. All statements about buses being "unable to compete" with cars should be viewed in this light. Due to the way roads are regulated, or not regulated, bus users are 'punished' with congestion while also donating road space to those that are dishing out that punishment. Take the bus; no good deed will go unpunished. In this situation, of course buses "can't compete." Saying that buses cannot compete is like saying a boxer who has been given tranquillisers cannot compete against another who has been given amphetamines; technically true but highly misleading.

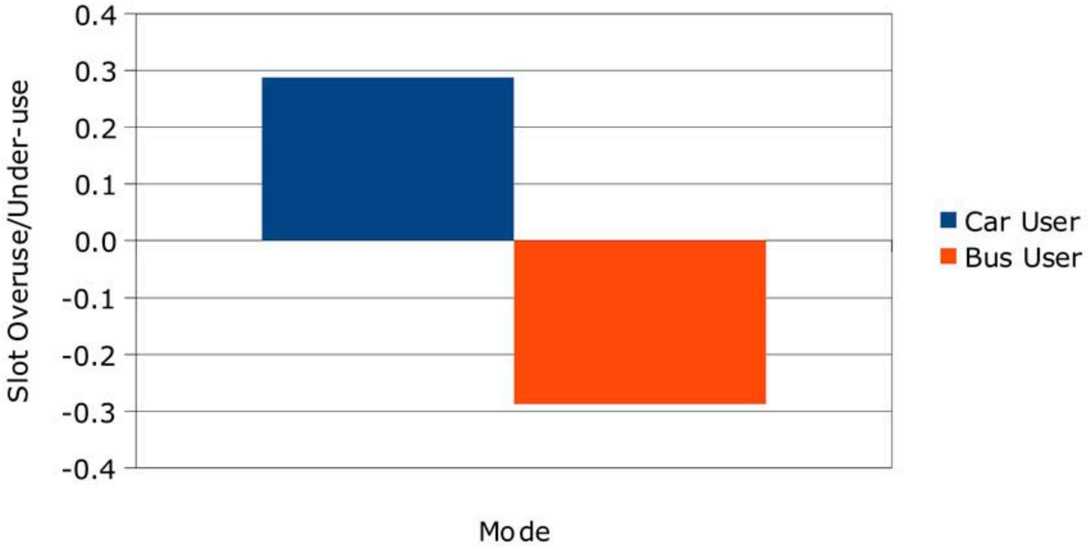


Figure 7: Overuse and Under-use of Slots by Mode

Road pricing gives a lead as to how the above situation might be addressed, such that roadways are managed under the principle: "you may, on average, take only your fair share of available road space." With road pricing, roadway users are customers, not owners, and so everyone using roadways pays and there is a net flow of value from the customers to the external roadway owners, be they private interests or the government. If, though, roadways are a commons, such transfers should be made between the owners, from those who are over-using a roadway to those who are under-using it, with no net flow of value outside the group of owners, namely the public. This could be done, during times when road space is scarce, by charging private vehicle users, who are on average overusing the roads, and by paying public bus users, who are on average under-using the roads.

Paying public bus users directly would risk people chasing the payment as income, but such payments could be made indirectly, by reducing the price of bus tickets via a discount. There is no real limit to how large such discounts could be, as ticket prices could be raised to accommodate higher discount levels, with the extra funds being used to improve bus system infrastructure, or to fund other costs associated with providing a better bus service.[31] These discounts are not a subsidy, they are a payment for the use of a scarce resource from the people who are overusing that resource to the people who had a right to that resource but gave it up. The community could decide how much traffic would be allowed on a given roadway, econometric professionals determine what transfer price would need to be charged at given times to hit that target, and each person overusing the roadway pays each person under-using it. The higher the transfer price from car users to bus users, the less attractive the car option becomes and the more attractive the bus option becomes. As more people use buses, overall vehicle load drops. If the outcome is roadway under-use, the transfer price was set too high. If the outcome is roadway overuse, the transfer price was set too low.

Such an approach would be financially neutral for a person who was not over-using the roads. Take, for example, someone who travelled to work and back Monday to Friday, where there was only one-fifth of the required road space available to make that journey in a car everyday. To achieve a given traffic level, the cost of taking a car on this trip to work might be set at $8 a day, charged in a manner to be discussed shortly. Bus users would, for this example, receive a $2 discount on their fare each day for their roadway under-use. If a person took the bus to work four days a week and their car one day a week they would pay $8 on the day they took their car and, on the other four days, have their bus ticket discounted by $2, for a total discount of $8. The requirement to pay $8 and the total discount of $8 cancel each other out. For those currently using a fair share of roadway space, the introduction of congestion offsets would have a negligible direct effect; the money lost and the money gained would balance out. For those using more than their fair share of roadway space, they would pay those using less than their fair share. In both cases, the public would remain owners of the roads. Such a system would cost money to administer, but as with local council taxes or road taxes, such payments would be used to provide a roadway congestion control service, not for the right to use a fair share of the roads, as such. Congestion offsets do not require the loss of public ownership of the roadway commons, a key political weakness of road pricing.

The second key weakness of road pricing is organisational, and again relates to the fact that, under road pricing, money is paid to buy access to road space. Payments come in many forms, such as purchases, taxes and fines, and each payment type has a different character. With road pricing the payment type is a fee for a service, with each driver acting in the role of a customer who is purchasing road space. When making a purchase, people expect to pay for what they use, not for more than they use. Few people would be happy if they went to a shop to buy a litre of milk and the shopkeeper tried to charge them for two litres. Because road pricing works via the payment of a fee for a service, precise metering is needed. In the few cities that have implemented road pricing, sophisticated toll gate systems are the norm, and, in significant part due to this complexity, such road pricing systems tend to be limited to the centre of the city.

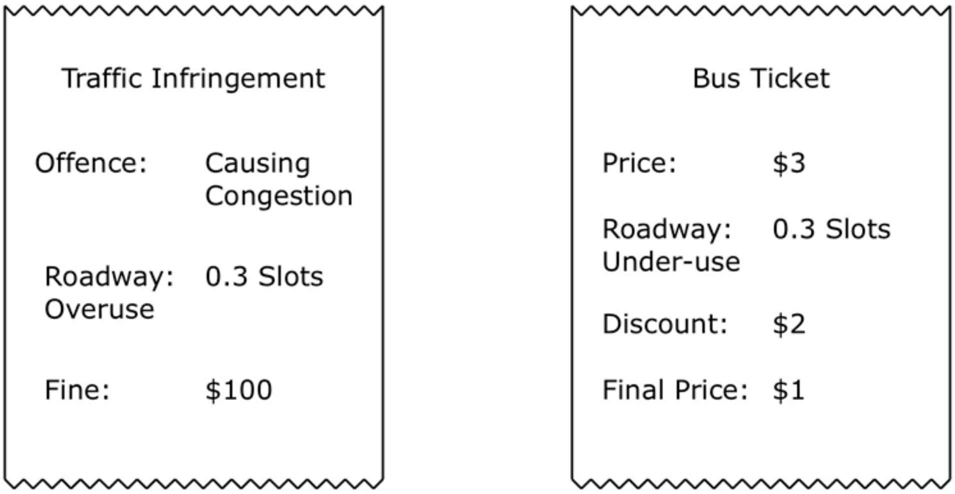
If the objective is instead to regulate the roadway commons, a fee for a service is not the only payment method that might be used. Another method, already widely used in other areas of roadway commons regulation, is available; namely the penalising of misbehaviour. For roadways we already have a wide spectrum of penalties for different types of misbehaviour, from, for example, small fines for parking in a clearway to jail time for causing death by the reckless use of a motor vehicle. There is no reason that the offence of “Causing Congestion” could not join the long list of all of our other traffic offences, with appropriate penalties attached. Tickets for causing congestion could be issued using the same general approach as parking tickets; any private vehicle on a given stretch of road, on a given side of the road, at a given time, incurs a penalty. In contrast to parking tickets, though, such penalties would be incurred by moving vehicles, rather than by parked vehicles. The moral basis for issuing such penalties is straightforward. Taking a car onto a congested roadway worsens the speed performance of that roadway for all. It is contributing to the despoiling of a public good. It is generally accepted that it is immoral to despoil a public good by taking more than your fair share of it. Immoral does not become moral just because everyone is doing it.

Congestion offsets treat people as citizens, not customers. Citizens are expected to behave properly and to, amongst other things, not wantonly despoil public goods. For those who do not behave properly, socially agreed sanctions are imposed. A penalty approach is no different in principle to controlling speeding; we do not have a market for speeding, we just penalise it sufficiently to reduce it to a socially acceptable level. It is currently the case that driving a car onto a congested road is a widely accepted practice; a person who would be horrified to see someone drop litter in a public place, does not think twice about driving onto a road that they know to be congested. Attitudes have changed in the past though; drink driving used to be widely socially tolerated, smoking in public buildings used to be legal in many places where it is now banned. Attitudes can and do change.

Penalising people for "Causing Congestion" has a major organisational advantage over charging a fee for a service, namely that people do not usually complain when society fails to punish them for their misbehaviour. Unlike prices, fines are fairly randomly imposed, with 'speed traps' and 'booze buses' catching people sometimes, and not at other times. With fines, there is no need for the precise metering of road pricing. As with parking fines, a city government could just erect signs, citywide, saying that in these places, at these times, congestion is a problem and so such-and-such a penalty will be imposed. These penalties could be enforced sufficiently to achieve the desired outcome, and no more. A person might drive all month in a congestion-prone area and only be fined once during that month, but if that fine is a significant percentage of the average monthly wage, behaviour will soon change. Congestion offsets do not require precise metering of roadway use, a key organisational weakness of road pricing.

To bring the above two points together; congestion offsets punish people who cause congestion and, in equal measure, reward people who alleviate congestion. Unlike road pricing, money is collected via fines for misbehaviour rather than a fee to purchase a service. Also unlike road pricing, the money collected via fines is explicitly paid to those helping to reduce congestion, in the form of cheaper bus services. Figure 8 shows mock-ups of how these fines and payments might look. The traffic infringement shown in Figure 8 would, on average, be issued once for every 50 of the bus tickets issued, thus both penalising and rewarding by the same amount, namely $100.

The approach shown in Figure 8 is one of offsets, thus the congestion offsets name. Due to widespread public discussion of climate change, the best known form of offsets are carbon offsets. Organisations who emit carbon dioxide as a result of their operations, who are said to have a certain 'carbon footprint', may choose to eliminate that footprint by buying carbon offsets, thus becoming 'carbon neutral'. The money paid for carbon offsets is used to plant trees, or to fund other activities that remove carbon dioxide from the at-



Traffic Infringement

Offence: Causing Congestion

Roadway: Overuse 0.3 Slots

Fine: $100

Bus Ticket

Price: $3

Roadway: Under-use 0.3 Slots

Discount: $2

Final Price: $1

Figure 8: Road Overuse Fines and Under-use Payments

mosphere, with the outcome being no net carbon dioxide entering the atmosphere. Congestion offsets are the same concept, except rather than planting trees to remove carbon, people use buses to remove congestion. Importantly, unlike the case of carbon offsets, where no single government controls the atmosphere, cities normally operate as a single political area and so congestion offsets could be mandated for all by a single government.

The defining characteristic of an offset is that it enables something negative to be wholly undone. With carbon, if enough offsets are purchased, an organisation's carbon emissions can be wholly undone, and they can become carbon neutral. Many types of bad behaviour cannot be undone. For example, when a person parks in a clear-way, causing two lanes of a road to drop to being only one lane, great inconvenience is caused. The fine received for this type of bad behaviour does not undo the damage; it is purely a punishment, a deterrent. By contrast, congestion offsets are a way of ensuring, in effect, that people only take their fair share of road space, by responding to any attempt to take more than a fair share in a way that brings road use back under an agreed traffic level. Under a system of congestion offsets, if people wish to take what would be more than their fair share of road space they are required to sufficiently fund bus services such that the unfair portion of road space they are taking becomes available. Congestion offsets would require all road users to be congestion neutral.

With the damage caused by congestion being wholly unwound, congestion offsets could be optionally charged as, in effect, a fee, instead of a fine. Take the example previously shown in Figure 8 where a congestion fine of $100 is levied, but where the fine is incurred, on average, only once out of every fifty times that a given road section is driven on. In this example a $2 congestion fine could alternatively be charged each time, with GPS and mobile communication technology now making this easily achievable. Indeed, many commercial vehicles are already fitted with live location tracking equipment. Congestion offsets could be charged as, in effect, a fee, for those who value determinacy more than privacy, or as a fine, for those who value privacy more than determinacy. It would be likely that most commercial vehicle operators, including freight operators, would opt for the determinacy of a fee. Under congestion offsets, charging the fine as a fee would be an auxiliary measure, available on a vehicle-by-vehicle basis, with fines as the default mechanism.

It should be noted that a deterministic outcome could be achieved without a high-tech, live-location metering solution. We have a well-established technology to deal with outlier risks, namely insurance. It comes with an administrative overhead, and the unpleasant but necessary need for fraud prevention measures, but it works. In a similar vein, sophisticated technology such as that used in red-light cameras would no doubt be very useful in issuing congestion fines, but the task could also be done by a police officer with an infringement pad. Likewise, super-computer modelling to determine optimal congestion offset levels would no doubt be beneficial, but crude try-and-adjust methods would achieve an acceptable outcome. Today's high-technology — live location tracking, automatic fine cameras, super-computer modelling — would be very useful in reducing cost and increasing efficiency, but such technologies are not necessary to deploy congestion offsets. We could have deployed congestion offsets at any time since the invention of the automobile; paper fines, insurance and crude try-and-adjust modelling methods are all technologies that pre-date the automobile.

Living for so long without congestion control has led to the most perverse results, something that can be best demonstrated by considering the alternative. Imagine if it had been normal practice for towns to introduce a system of congestion offsets a generation or so before they had any chronic congestion problems. Car drivers in such towns would, from time-to-time, receive road space use notices, stating that on a given date and time they were recorded using one road slot when, say, two road slots were available; giving a roadway overuse level of zero slots. This notice would record that the fine for overusing zero slots is zero dollars, and thank the driver for staying within the available road space. Similarly, on the small bus services run by such towns, bus tickets would record that passengers had underused zero of the contested slots, due to a surplus of one slot per vehicle, and that they were receiving a discount of zero dollars on their tickets.

During the twenty or thirty years before these towns had sufficient traffic to cause chronic congestion, no fines or discounts would be issued, but both car and bus users would see, as their towns grew, the excess road slots slowly being used up. For car drivers this would draw them closer and closer to the day where they were going to start being fined for their roadway overuse. By contrast, the diminishing number of free roadway slots would draw bus passengers closer and closer to the day where they were going to start being paid for their roadway under-use. One day, which people would have seen coming for decades, car drivers would start receiving occasional small fines for their roadway overuse, and bus passengers would start receiving tickets which stated that they had underused a tiny percentage of the contested road slots, and that they were therefore receiving a discount of one cent on their tickets. Over time, both the size of the fines being issued to drivers and the size of the bus ticket discount would increase. Buses would slide into greater use as people found their share of the roadway commons dropping in line with urbanisation. There would be a slow increase in public transport use, simply as a matter of rational decision-making. Chronic congestion would never be a feature of such towns' transport planning landscape, as they grew from towns to cities.

Would a city with such a history ever talk about building large arterial roads simply to carry car traffic, or heavy rail lines in places where the passenger load simply did not justify the construction of such lines? It seems most unlikely. Far more likely, would be that both heavy road and heavy rail infrastructure would be built sparingly, and only when circumstances fully justified such construction. Not so in our cities today; once two lanes, one in each direction, are no longer sufficient to soak up traffic, four lane roads, or even heavier grade roads are built. Why, though, should private vehicles get to use up all this space when there are other users of street space? Why not have more bike lanes or wider footpaths? The same point could be made for space use within a single road lane. Throughout this paper, all calculations have been made on the assumption that all of the capacity of a road lane was to be used by vehicles. If a road lane could take 700 cars, it did take 700 cars, and without congestion control this is indeed what happens. A street with more traffic, though, is a worse street for everyone using it except the vehicle users.[32] Under a system of congestion offsets it would be possible to consciously run road lanes significantly below their maximum capacity. In the absence of congestion control, many of our cities have become awash with both traffic and bitumen.

One option for congestion offsets would be to implement them in a minimal way; to eliminate chronic congestion, but to still provide as much mobility as possible using cars. This would be an improvement over the current situation. Alternatively, congestion offsets could be used in a transformative way; to slowly squeeze a great deal of both the traffic and the bitumen out of our cities. Such an approach would see chronic congestion become a thing of the past, indeed many roads may be run well below capacity, so as to balance the needs of vehicle users and other street users. Many road widening projects would most likely be deemed no longer necessary, and all existing road grades above two lanes, one in either direction, could be reviewed to see if the extra lanes were really necessary, or whether the land could be put to better uses. Mass mobility could be provided by a grid of high frequency, BRT-style bus services running on normal congestion-free roads. Under such a system most people would travel on buses most of the time, with cars playing an important, but far smaller role than they do now. In short, congestion offsets could be used to help rehabilitate our cities from their current degraded state, where they are heavy with both traffic and bitumen. Congestion offsets could be an important new tool to help transform our cities.

With congestion offsets now described, the question raised in the introduction as to whether congestion offsets constitute a distinct congestion control method can be revisited: are congestion offsets just a variant of road pricing? From a conceptual standpoint, road pricing and congestion offsets are clearly different congestion control methods. The central concept behind congestion offsets is to keep roadways as a commons, but to regulate that commons better, whereas the central concept behind road pricing is to enclose the roadway commons, with the road space then being ‘parcelled up’ for sale. Even with congestion offsets being a conceptually distinct congestion control method it is nevertheless possible for two conceptually distinct methods to yield very similar policy prescriptions. If there are no significant differences between the policy prescriptions of congestion offsets and those of road pricing, then it would be difficult to maintain a position that congestion offsets are a distinct form of congestion control.

In cases where road pricing does not come packaged along with policy measures to funnel raised funds towards public transport, congestion offsets are clearly a distinct form of congestion control; congestion offsets take funds from road users and also give funds to road users, whereas road pricing only takes funds from road users. This is a very significant difference in terms of policy prescriptions between congestion offsets and road pricing. But what about circumstances where road pricing does come packaged along with policy measures to funnel raised funds towards public transport, such as in London. Are congestion offsets a clearly different form of congestion control in comparison to “road pricing plus public transport funding?” Putting aside the fact that we are no longer talking about road pricing, as such, but road pricing with an auxiliary policy bolted on the side, there are still key differences in how road pricing and congestion offsets take funds from road users and give funds to road users. Although the money raised from road pricing in London is ploughed back into improving their public transport system, one is not reminded every time one gets on a bus, via an explicit financial reward, that one was underusing his or her fair share of road space. Also, on the penalty side of the penalty/reward equation, road users in London are not fined for their misbehaviour of overusing the roads; they are instead offered the chance to purchase access to road space. The whole behavioural mechanism, the reward/penalty system and therefore psychology of what this paper has proposed is very different from road pricing, even in cases where road pricing has been coupled with the funding of public transport.

Finally, there is the impact that using fines, instead of fees, has on ease of deployment. This fines-versus-fees difference between congestion offsets and road pricing stems directly from the fact that, unlike road pricing, congestion offsets have a built-in notion of what constitutes a fair share of road space, and can therefore advance the proposition that people who wantonly despoil a public good by taking more than their fair share of it should be punished, via a fine. Fines can be randomly imposed, fees cannot. The fact that fees cannot be randomly imposed saddles road pricing with the requirement for precise metering, which has drastically restricted its deployment, usually to small central city areas, even when congestion is a citywide problem. This different policy prescription, the use of fines instead of fees, yields a concrete and highly beneficial outcome; it allows congestion control to be easily deployed citywide. Congestion offsets are not a variant of road pricing; they are a distinct method of congestion control.

For readers who remain unconvinced, who continue to view congestion offsets as simply an aggressive repackaging of road pricing into a more ‘palatable’ form, these readers are invited to consider the possibility that an aggressive repackaging of road pricing into a more ‘palatable’ form is, in fact, desperately needed. In terms of the number and scope of deployments, road pricing has been a spectacularly unsuccessful policy. The idea of road pricing has been around for decades, during which time congestion has grown to be an enormous worldwide problem, and yet only a handful of cities have adopted road pricing schemes. In a continuation of this trend, a road pricing scheme for New York has been recently shelved by the New York State Assembly.[33] The normal response, when a ‘product’ fails to ‘sell’ well over an extended period of time, is to redesign the product into something that people actually want to buy. Instead, those advocating road pricing appear to be doggedly trying to sell something that, the evidence suggests, not many people want to buy. The difficulties inherent in this course of action are highlighted by one of the questions to be discussed at an urban congestion conference scheduled to be held some months after the finalisation of this paper. One panel at this conference, somewhat surprisingly, plan to, amongst other things, engage themselves with “delving deeper into the road pricing scheme and answering the ultimate question of whether it is a death wish or real solution?”[t4] That congestion control through traditional road pricing approaches could possibly be equated with a “death wish”, gives considerable pause to thought and highlights the difficulties faced in trying to introduce road pricing in its present form on a large scale in cities. The authors suggest that this course of action should be reconsidered in light of the demonstrated history of its results, and the lack of any solid evidence suggesting that there might be a significant improvement of those results in the future.

Whether the approach presented in this paper is called congestion offsets, or road pricing with fairness built in, or social road pricing, or road pricing 2.0, is not a matter of great consequence. As the previous section of this paper detailed, the urban transport situation is far worse than generally acknowledged. Buses on normal roads, an entire transportation mode, and one that cannot be effectively substituted for, has been crippled by congestion, with disastrous results. This terrible situation is a matter of great consequence, and one that will only be resolved by the widespread deployment of an effective form of congestion control. Consequently, there is a desperate need for a congestion control method that works, that is easy to deploy and that has the concept of fairness built into its fundamental design principles, so that it has some hope of being acceptable to the general public. Road pricing meets one of those three requirements, namely that it works; congestion offsets meet all three.

The urban transport problems of cities, no matter how dire, are, of course, only one part of the bigger urban sustainability picture. However effective congestion offsets might be in addressing the particular sustainability problems of urban transport, they would, on the whole, be an ineffective tool for addressing the broader sustainability problems of cities. Tools are normally only any good at solving the problem they were designed to solve, and congestion offsets are a tool that has been designed to solve transportation problems, not wider social, economic or environmental problems. Introduce a system of congestion offsets in a socially unjust, economically stagnant and environmentally ruined city that also happened to have terrible traffic problems, and the likely outcome would be a socially unjust, economically stagnant and environmentally ruined city that no longer had terrible traffic problems. Ways of addressing these wider social, economic and environmental issues have been extensively described by others, and the broader sustainability of cities, their capacity to become regenerative in a wider sense, rests on a broad array of issues and strategies for action.[35]

Conclusion

A major part of the urban transport problem today is a failure from the very beginning to acknowledge that congestion is fundamentally inequitable and unfair, impractical to construct away, and therefore must be properly charged for and controlled to eliminate the transport system dysfunction which is systemic in cities today. By permitting free access to roadways as a matter of policy, governments have allowed cars to cripple the natural competitor for mid-level loads, namely bus services on normal roads. Cars dominate urban transport in most of the world not due to any inherent performance superiority, but because they have what amounts to a government-granted monopoly. In the absence of an effective mid-level transport mode, the next best option for transport planners has been to try and make low and high load modes look like good mid-level modes. To this end, the load capacity of roads has been greatly increased via road building projects and the load capacity of railways has been greatly reduced by operating them well under capacity. Running railways well under capacity at least has the advantage of only costing cities money. Greatly increasing the load capacity of roads via road building projects has blighted our cities, leaving many of them bloated with both traffic and bitumen.

As a solution to the urban transport problem, road pricing has been near universally rejected, and for good reason. Road pricing requires precise metering, making it difficult to deploy, and disenfranchises the public from the roadway commons, making it all but impossible to sell politically. Road pricing is, in short, a complex and unfair form of congestion control.

Congestion offsets have been presented here as an alternative to road pricing. In brief, congestion offsets involve penalising private vehicle users for overusing the roads and rewarding public bus users for under-using them. The design principle behind this approach is that roadway congestion should be addressed not by constructing a market, but by regulating a commons. Under congestion offsets, a person who, on average, uses only their fair share of roadway space would not incur any direct net costs. Even in the case of overuse, there is no transfer of value outside of the users of the roads, with value simply being transferred from over-users to under-users. Unlike road pricing, congestion offsets do not disenfranchise the public of the roadway commons. Also following from the design principle that congestion should be addressed by regulating a commons, congestion offsets ‘lean’ on roadway over-users not by charging them for roadway space, as such, but by punishing them for their misbehaviour via fines. Consequently, unlike road pricing, congestion offsets do not require precise metering. Congestion offsets are, in short, a simple and fair form of congestion control.

Better cities are possible. A tolerable accommodation with congestion is not the only option. Congestion offsets, a simple and fair alternative to road pricing, would allow congestion to be squeezed out of our cities, and for bus services to emerge as an efficient and effective mode for servicing mid-level transport loads. The massive distortions we have allowed to build up in our cities, in the absence of such a mode, could be slowly unwound. Lean roads on multi-purpose streets could become the norm, as they should have always been. Simply by deciding to let buses compete, we could help to transform our cities in significant ways by addressing two key issues: the modal split between private motorised and public motorised modes and the endemic congestion that plagues cities. Like the great freeway project this is a big dream, but our cities are our collective homes and we should dream big dreams about them. Congestion offsets could be a key tool in helping to shift cities towards a new, regenerative model.

Author details:
Matthew Bradley
matthew.academic@gmail.com

Jeff Kenworthy
Curtin University Sustainability Policy Institute (CUSP)
Curtin University of Technology
Perth, Western Australia

j.kenworthy@curtin.edu.au

1 Jacobs, J. (1961) The Death and Life of Great American Cities. Random House. New York.

Schneider, K. (1979) On The Nature of Cities: Toward Enduring and Creative Human Environments. Jossey-Bass Publishers, San Francisco.

2 Kunstler, J.H. (1994) The Geography of Nowhere: The Rise And Decline of America’s Man-Made Landscape. Touchstone. New York.

Stretton, H. (1975) Ideas of For Australian Cities. Georgian House. Melbourne.

3 Kenworthy, J. (2012) Don’t Shoot Me, I’m Only The Transport Planner (Apologies to Sir Elton John). World Transport Policy and Practice, Issue 18.4.

4 Ibid.

5 Poole, R.W. (Jr) Resolving Gridlock in Southern California. Transportation Quarterly 42 (4), 499-527.

Orski, C.K. (1987) “Managing” Suburban Traffic Congestion: A Strategy for Suburban Mobility. Transportation Quarterly 41 (4), 457-476.

Cervero, R. (1984) Managing the Traffic Impacts of Suburban Office Growth. Transportation Quarterly 38 (4), 533-550.

Gleick, J. (1988) National Gridlock: Scientists Tackle the Traffic Jam. The New York Times Magazine, May 8.

Pratsch, L.W. (1986) Reducing Commuter Traffic Congestion. Transportation Quarterly 40 (4), 591-600.

6 Goodwin, P. (1997) Solving Congestion (when we must not build roads, increase spending, lose votes, damage the economy, or harm the environment, and we will never find equilibrium). Inaugural lecture for the professorship of transport policy. University College London, 23rd October 1997.

7 Larson, T. D. (1988) Metropolitan Congestion: Towards A Tolerable Accommodation. Transportation Quarterly. 42 (4), 489-498.

8 The approach shown in Figure 1 draws significantly on material from the Travel Time Equations and Diagrams section of Professor Vukan Vuchic’s Urban Transit Systems and Technology book.

9 Kenworthy, J. (2012) Update of The Millennium Cities Database for Sustainable Transport. Unpublished.

10 Wright, L., Hook, W., (eds) (2007) Bus Rapid Transit Planning Guide. Institute for Transportation & Development Policy. New York. pp. 767-778.

11 Dittmar, H., Ohland, G.(eds) (2003) The New Transit Town: Best Practices In Transit-Oriented Development. Island Press. Washington, p. 20.

12 Wright, L., Hook, W. (eds) (2007) Bus Rapid Transit Planning Guide. Institute for Transportation & Development Policy. New York. pp. 767-778.

13 U.S. Department of Transit, Federal Transit Administration (1992) Characteristics of Urban Transportation Systems. Chapter 3: Bus Transit. Table 3-15.

14 Kenworthy, J. (2012) Update of The Millennium Cities Database for Sustainable Transport. Unpublished.

15 Mees, P. (2000) A Very Public Solution: Transport in the Dispersed City. Melbourne University Press. Victoria, Australia. pp. 138-142.

16 U.S. Department of Transport, Federal Highway Administration (2009) Summary of Travel Trends: 2009 National Household Travel Survey. p. 10. Table 3. Average person trip length, 2009.

17 Wright, L., Hook, W. (eds) (2007) Bus Rapid Transit Planning Guide. Institute for Transportation & Development Policy. New York. pp. 767-778.

18 Kenworthy, J. (2012) Update of The Millennium Cities Database for Sustainable Transport. Unpublished.

19 Vuchic, V. (2007) Urban Transit Systems and Technology. John Wiley & Sons, Inc. Hoboken. New Jersey. p. 77.

20 U.S. Department of Transport, Federal Highway Administration (2009) Summary of Travel Trends: 2009 National Household Travel Survey. p. 33. Table 3-16. All Purposes, 2009.

21 U.S. Department of Transit, Federal Transit Administration (1992). Characteristics of Urban Transportation Systems. Chapter 3: Bus Transit. Table 3-16.

22 Even using an average car occupancy level of 1.2 people per car, a typical figure for peak hour use, which would result in a load level of 840 pph, one could still all but fill ten standard buses to capacity every hour, delivering a six minute service frequency, better than most cities provide today.

23 Transport Research Board (2000) Highway Capacity Manual 2000. National Research Council. Washington D.C. Exhibit 21-8: Passenger-Car Equivalents on Extended General Highway Segments.

24 Ceder, A. (2003) Public Transport Timetabling and Vehicle Scheduling. In: Lam, W.H.K., Bell, G.H., Advanced Modeling for Transit Operations and Service Planning. Elsevier Science Ltd. Pergamon. p. 37.

25 U.S. Department of Transit, Federal Transit Administration (1992) Characteristics of Urban Transportation Systems. Chapter 3: Bus Transit. Table 3-16.

26 Vuchic, V. (2007) Urban Transit Systems and Technology. John Wiley & Sons, Inc. Hoboken. New Jersey. p. 76.

27 Kenworthy, J. (2012) Update of The Millennium Cities Database for Sustainable Transport. Unpublished.

28 It could be argued that the current Certificates of Entitlement systems in operation in Singapore and Shanghai, run as monthly auctions for the right to buy a car, with the number of certificates issued being set roughly to the attrition rate of old cars, are an indirect form of congestion control. The Singaporean system has been successful in keeping car ownership at about 100 cars per 1000 people, despite a GDP per capita that would imply much higher ownership. This clearly limits the growth of congestion in the medium to long term. The city of Beijing has also introduced basically the same system, but runs it as a lotto scheme each month, so it is pure luck as to who gets a ticket to buy a car. These approaches are, though, such blunt policy instruments that they will not be considered further here. Congestion could also be controlled by banning cars from cities outright.

29 Wessels, W. J. (2006) Economics: Barron's Educational Series (4th Edition). Barron's Educational Books. New York. p. 2.

30 Santos, G. (ed) (2004). Road Pricing: Theory and Evidence. Research in Transport Economics: Vol. 9. Elsevier Press. The Netherlands.

31 Although road pricing often comes packaged along with policy measures to funnel some or all of the funds raised towards public transport, such measures are auxiliary to the central mechanism of road pricing; the commodification and sale of road space. By contrast, the discounting of bus tickets is an integral part of congestion offsets, with the issuing of discounts for roadway under-use being the mirror policy to charging for roadway over-use. Congestion offsets have a built-in concept of what constitutes a fair-share of road space, and consequently are able to have built-in ways of dealing with (cont.) roadway use both above and below that fair amount. Congestion offsets are 'carrot' and 'stick' in equal measure. As the name would suggest, road pricing is a 'stick' only mechanism, which is why it is so commonly found packaged with 'carrot' measures, including the funnelling of raised funds towards public transport. The frequency with which road pricing is found packaged along with the funding of public transport, suggests the need for a congestion control method, such as congestion offsets, whose fundamental design principles naturally encompass both 'carrot' and 'stick' measures.

32 Appleyard, D. (1981) Livable Streets. University of California Press. Berkeley. California. pp. 15-28.

33 New York Assembly Shelves Manhattan Congestion Charge, The Guardian Website, Available: http://www.guardian.co.uk/politics/2008/apr/08/congestion-charging.pollution (Accessed 20 August 2012)

34 Tonkin's Urban Congestion Conference, The Tonkin Corporation Website, Available: http://www.tonkincorporation.com/images/Eureka/ENG40E.pdf (Accessed 30 August 2012)

35 Newman, P. and Kenworthy, J. (1999) Sustainability and Cities: Overcoming Automobile Dependence. Island Press, Washington, DC.

Schiller, P. Bruun, E. and Kenworthy, J. (2010). An Introduction to Sustainable Transportation: Policy, Planning and Implementation. Earthscan, London.

Newman, P. and Jennings, I. (2008) Cities as Sustainable Ecosystems: Principles and Practices, Island Press, Washington DC.