Road capacity change and its impact on traffic in congested networks: evidence and implications

Roger B Behrens & Lisa A Kane

This article reviews explanations of, and international empirical evidence for, 'induced' traffic as a result of increased road capacity and 'suppressed' traffic as a result of decreased road capacity. In essence, the former refers to new traffic appearing as a result of new road construction, while the latter refers to traffic disappearing as a result of road closure. Despite problems with the available data and their measurement, it is concluded that — with the caveats of either pre-existing congestion in the case of capacity increases or no spare capacity in the case of capacity decreases — the weight of evidence indicates that induced and suppressed traffic are indeed real phenomena. It is argued that the link between traffic and road capacity is therefore far more complex than previously understood. The implications this has for both urban passenger transport planning practice and policy formulation are discussed.

1. INTRODUCTION

Under-funding of transport infrastructure investment in South African cities in recent years and a concomitant increase in competition for limited transport project funding have led to considerable debate on the priority and appropriateness of proposed investments in transport infrastructure. In particular, debate has focused on whether increasing road capacity through adding new network links or widening existing roads constitutes the most effective and indeed the most appropriate means of addressing the problem of traffic congestion that besets cities in periods of peak travel demand. While different in scale and context, other parts of the world have had similar debates, and a considerable body of literature has now developed on the often counter-intuitive, empirically observed impacts that road capacity change has on the alleviation of traffic congestion problems. The literature documenting these empirical studies deals, in essence, with two phenomena associated with road capacity change — 'induced' and 'suppressed' traffic.

The purpose of this article is to review the international literature, and extract from it the implications for transport planning policies and practices that seek to address the vexing problems of identifying the appropriate nature of transport infrastructure investment and prioritising transport projects in the context of resource scarcity and competition.

The article begins by defining what is meant by the terms 'induced' and 'suppressed' traffic (Section 2). Section 3 reviews the explanations offered in the literature of the behavioural responses that give rise to induced and suppressed traffic phenomena.
Section 4 reviews the empirical evidence supporting the existence of these phenomena. Section 5 discusses the implications this evidence has for transport planning practice and policy formulation. The article concludes with a discussion on the relevance of these implications for the South African context (Section 6).

2. DEFINITION OF INDUCED AND SUPPRESSED TRAFFIC

The watershed publication in the debate over induced traffic was, without doubt, a report submitted to the United Kingdom's Secretary of State for Transport in 1994, by the Standing Advisory Committee for Trunk Road Assessment (SACTRA, 1994). Although focused primarily on trunk roads (i.e. roads in the national road system managed by the national government), the location of many of these routes through, or close to, conurbations meant that roads with a wide variety of traffic conditions were considered, from very congested to virtually free-flowing. The SACTRA committee was asked to advise 'on the evidence of the circumstances, nature and magnitude of traffic redistribution, mode choice and generation (resulting from new road schemes)' (SACTRA, 1994: 1).

At the outset, the committee recognised that the notion of 'roads generating traffic' was one that had gained widespread popular acceptance, but that had not been subjected to much rigorous investigation. There had been confusion and inconsistency over terminology in the literature, one reason being that the definitions of 'generated' traffic are not straightforward. Until the publication of the SACTRA report, problems in defining induced traffic had blurred real debate over the actuality of the phenomenon. There was a lack of professional consensus over what induced traffic was, and if it existed to any significant extent. The committee thus chose to address this issue of definition in detail (see Hills, 1996), and this section of the article summarises the final definition they chose to use. Others have suggested slight variants on this definition in the intervening years (e.g. Heanue, 1998; Litman, 1999), but for the purposes of this article SACTRA's definition is used.

As a first step in its work, SACTRA decided that the word 'generate' was problematic to use, as 'trip generation' has a very specific meaning in a transport planner's vocabulary. Usually a household or individual is understood to generate trips, as the first stage in the four-stage modelling process. With this very particular definition in mind, the notion that 'roads generate traffic' (and implicitly, then, trips), seems nonsensical. To preclude any confusion over this particular issue, the SACTRA committee chose to replace the term 'generate' with 'induce', and to investigate whether the provision of roads induces (i.e. indirectly brings about) traffic. In summary, the definition adopted by SACTRA was that induced traffic is the additional daily private vehicle traffic that may occur on a network following some increase in road capacity. It does not, therefore, include additional traffic found on individual links if the total of network vehicle kilometres remains constant, nor does it include additional traffic found in the peak period if the total of daily vehicle kilometres remains constant.

The literature on the relationship between traffic levels and road capacity was broadened considerably by the more recent publication of a report entitled, Traffic impact of highway capacity reductions: assessment of the evidence (Cairns et al., 1998). This report was commissioned by London Transport and the Department of Environment Transport and the Regions. The premise of the study was that if we accept the notion of induced traffic, we must also consider the possibility of traffic being
suppressed when road capacity reductions are imposed. The report therefore reviewed
evidence of the traffic impacts of capacity reductions (e.g. as in the case of a bridge
or lane closure), rather than increases. It is important to note that in this report, as well
as in the earlier SACTRA report, the capacity increases or reductions in question
referred specifically to road capacity change for private vehicles. The impacts of
increased or reduced capacity for public transport were therefore not investigated in any
significant way.

This article reviews the literature on the explanations for, and empirical evidence of,
‘induced’ and ‘suppressed’ traffic conjointly. The earlier SACTRA definition is thus
expanded to include both phenomena as follows: Induced (or suppressed) traffic is the
additional (or reduced) daily private vehicle traffic that occurs on a network following
some capacity increase (or reduction) or, as discussed in the following section, some
other reduction (or increase) in the generalised cost of travel.

3. EXPLANATION OF INDUCED AND SUPPRESSED TRAFFIC PHENOM-
ENA

It is widely held in the literature on travel behaviour that motorists are most likely to
adapt their behaviour when faced with significant changes to the cost of, or constraints
on, their travel choices. The phenomena of induced and suppressed traffic have thus
been observed to occur in situations where a change in road capacity causes a
significant change in the generalised cost or attractiveness of motor car travel. In the
case of capacity increases, such cost changes would typically only occur in situations
where road space is added to networks already experiencing congestion, or in the case
of major new links leading to a dramatic change in generalised cost (e.g. a river
crossing). In the case of capacity reductions, such cost changes would occur typically
in situations where road space is taken away from networks that have little or no
existing spare capacity (Goodwin, 1996; Cairns et al., 1998; DeCorla-Souza & Cohen,
1999).

What happens when capacity is changed on a network with either congestion, or no
spare capacity, and thus how can the phenomena of induced and suppressed traffic be
explained? The literature offering such behavioural explanations identifies a variety of
possible travel adaptations, which tend to be noticeably different in the immediate,
short and long term (Kitamura, 1994; SACTRA, 1994; Goodwin, 1996; Cairns et al.,
1998; Dowling & Colman, 1998; DeCorla-Souza & Cohen, 1999; Litman, 1999;
Noland, 1999). In many but not all respects, behavioural responses to increased
capacity mirror, in inverse form, responses to reduced capacity. These temporally
differentiated responses are discussed below and are summarised in Tables 1a and 1b.

In the immediate term (i.e. the first few days), drivers often simply change their driving
styles in ways that adjust to the new traffic conditions. In the case of capacity
reductions in particular, and often depending on the amount of forewarning received by
media predictions of ‘traffic chaos’, they have been observed to drive slower and closer
together. However, the introduction of a £5.00/day congestion charge in London in
February 2003 – whereby the generalised cost of private cars was significantly altered –
suggests that immediate-term responses can extend beyond simply changed driving
styles. On the first day of the scheme (albeit a school holiday), instead of the traffic
‘nightmare’ predicted by the media, automated counting machines showed that the
Table 1a: Behavioural responses leading to induced traffic

<table>
<thead>
<tr>
<th>Changes in:</th>
<th>Induce person trips/day?</th>
<th>Induce vehicle km/day?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Timing</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Mode (to private car)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Vehicle occupancy (decreasing)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Trip frequency (increasing)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Trip destination (becoming more remote)</td>
<td>(Yes)¹</td>
<td>Yes</td>
</tr>
<tr>
<td>Trip origin (becoming more remote)</td>
<td>(No)²</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Notes: ¹In the longer term, land-use changes as a result of capacity increases may result in new destinations on offer, and hence new trips.
²The source of trip origins is the household, and while the location of the household may change, there will not be any new trips simply as a result of this relocation. There may, however, be induced traffic due to the need to undertake the previously planned trips to new destinations, via new routes.

Table 1b: Behavioural responses leading to suppressed traffic

<table>
<thead>
<tr>
<th>Changes in:</th>
<th>Suppress person trips/day?</th>
<th>Suppress vehicle km/day?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Timing</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Mode (from private car)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Vehicle occupancy (increasing)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Trip frequency (decreasing)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Trip destination (becoming less remote)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Trip origin (becoming less remote)</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

amount of morning traffic moving into central London was 25 per cent lower than on a normal working day (Clark, 2003).

In the shorter term (i.e. the first couple of months), behavioural responses tend to take the form of re-routed trips or rescheduled departure times. In the case of capacity increases, trips are attracted from other previously quicker, but now slower, routes within the network, or trips are rescheduled to a preferred departure time (sometimes referred to as the 'return-to-peak' effect) in response to the initial relief in congestion. In the case of capacity reductions, trips are re-routed to neighbouring streets or are rescheduled (i.e. departing a little earlier or a little later) to avoid the worsening congestion. It should be noted that changes in departure time or route, while clearly increasing or reducing the amount of traffic on the particular links subject to capacity change at a particular time, do not necessarily lead to induced or suppressed traffic as defined earlier. They may simply cause traffic to either divert from or reappear on other equidistant links within the network or at other times on the same links. The total amount of traffic on the network as a whole over the whole day would thus remain relatively constant before and after the capacity change. Only if the re-routed trips involve significantly shorter or longer trip distances does induced or suppressed traffic occur. Rescheduled departure times, on the other hand, would not strictly induce or
suppress traffic, even though Noland (1999) does observe that the ‘return-to-peak’
effect may induce new trips by freeing up capacity at other times of the day and
theoretically, at least, the inverse would be true for shifts to the off-peak times.

In the longer term (i.e. up to five to ten years after the capacity change), behavioural
responses tend to take the form of changed mode, trip frequency or trip end. In the case
of capacity increases, the reduced generalised cost of travelling by car may lead to
people taking trips by car that were previously undertaken by other modes, taking more
frequent trips, or because of improved travel speeds, taking trips to preferred destinations further away. In the case of capacity reductions, the failure of shorter-term
behavioural adjustments to avoid unacceptable congestion delays may lead to people
using non-motorised or public transport modes instead of their cars, suppressing
non-essential trips or at least linking previously separate trips into chains, or selecting
nearer destinations. In some instances, the change in travelling conditions may ‘tip the
balance’ in decisions that were being made for other reasons, such as buying or selling
a car, moving house or moving job. All these behavioural responses potentially
contribute to induced or suppressed traffic effects. In addition to these behavioural
responses, increased road capacity (and accessibility) can stimulate unforeseen changes
in land-use patterns, which can generate further unexpected traffic (Headicar, 1996).

In summary, it can thus be expected that – with the caveats of occurring on networks
with either congestion or no spare capacity – an increase or reduction in road capacity
may result in a change of route; timing; mode (or at least vehicle occupancy); trip
frequency; trip destination; or, over the longer term, trip origin. As illustrated in Table
1, these behavioural responses have an impact on either the number of person trips or
the distance of existing trips, both of which potentially induce or suppress traffic.

The behavioural responses discussed above have been argued to be consistent with the
basic microeconomic theory of supply and demand (SACTRA, 1994; Goodwin, 1997;
Litman, 1999; Noland, 1999). In terms of the microeconomics of supply and demand, any increase in supply (i.e. road capacity) results in a reduction in price (i.e. the
generalised cost of travel). The theory holds that when any good (i.e. travel) is reduced
in price, demand for that good increases. Hence, increased road capacity leads to
generalised costs going down, and so demand for motor car travel increases and traffic
is induced. Figure 1 illustrates the induced traffic effect in terms of this theory. Line $s_1$
represents road capacity supply before new road construction. Line $s_2$ represents
road capacity supply after an increase in capacity that results in a lower generalised
cost of travel – from $g_c1$ to $g_c2$. Following the dashed line that plots increased demand
in relation to decreased cost, at $g_c1$ on the y-axis the quantity of travel (measured in
vehicle kilometres travelled) is $q_1$ on the x-axis. At $g_c2$ the quantity of travel is $q_2$. The
quantity of travel therefore increases from $q_1$ to $q_2$ as the change in supply lowers the
cost of travel from $g_c1$ to $g_c2$. The increase in the quantity of travel from $q_1$ to $q_2$
represents the induced travel effect.

Conversely, reduced road capacity leads to generalised costs going up, and so demand
for motor car travel decreases and traffic is suppressed. Given that such microeconomic
theory underpins so much modelling and road scheme appraisal practice, as Goodwin
(1997) observes, ‘one wonders why the phenomena are greeted with surprise’.

4. EMPIRICAL EVIDENCE

Documented empirical evidence on induced traffic has been available in the literature
Figure 1: Induced traffic explained in terms of the microeconomic theory of supply and demand

Note: VKT = vehicle kilometres travelled.

since the early 1960s, but it was not until the 1990s that a body of authors began investigating the topic in detail (Kitamura, 1994). The seminal SACTRA report of 1994 has already been mentioned, but perhaps equally influential in the United States was the San Francisco Bay Area Lawsuit (Garrett & Wachs, 1996; Weiner, 1997). In June 1989, two environmental organisations claimed that the State of California, the Metropolitan Transportation Commission (MTC) of San Francisco, and other regional agencies had violated the provisions of federal clean air legislation by not doing enough to meet clean air standards. The case focused on the general issue of the effects of increased road capacity on reducing public transport use, increasing traffic speeds and enabling the spread of urban sprawl, all of which were argued to contribute to greater air pollution emissions. The MTC had undertaken a conventional analysis to determine the emissions impacts of its transportation plan. The environmental groups argued that conventional travel forecasting models overstated the emission benefits of road building in that, while they fully reflected the impact of speed improvements on reducing emissions, they showed little or none of the air-quality impacts of travel induced by speed improvements. The phenomenon of induced traffic thus lay at the heart of the court case. At the time of the trial, the American Transportation Research Board was non-committal about the link between roads, induced traffic and environmental pollution, but later evidence emerged from California demonstrating that increased road supply did appear to increase vehicle kilometres travelled (Hansen & Huang, 1997). The verdict on whether or not this increase in vehicle kilometres is detrimental to environmental pollution is still undecided. Most recently, both British and American authors have concerned themselves with how to model the induced traffic that is now acknowledged to occur (Coombe et al., 1998; DeCorla-Souza & Cohen, 1999; Noland, 1999).

In this section, empirical evidence which the authors believe has at least swayed the academic debate towards an acceptance of induced and suppressed traffic phenomena,
is presented. As there is a considerable volume of empirical evidence available, only the pieces of evidence that are particularly enlightening, or that appear particularly relevant to South Africa, are presented.

4.1 Evidence of induced traffic from before-and-after traffic counts

Probably the most obvious means of examining this phenomenon is to inspect the results of actual road improvements or closures and to compare these with what was expected. It is enlightening to examine the case of the M25, London’s orbital freeway that was finally completed in 1988. The press at that time branded the road as a ‘transport fiasco’ and ‘obsolete before it was opened’, for reasons that will become evident through scrutiny of Table 2, giving the annual average daily two-way (AADT) traffic on selected links of the M25 in 1992, and the equivalent design year forecast.

A review of the experience of the M25 concluded that in the years immediately following the opening of the freeway, reassignment, redistribution and mode shift had been important. The SACTRA report speculated that induced development traffic (i.e. induced traffic as a result of new or changed destinations) may become important in future. SACTRA concluded that the M25 appears to confirm the notion that roads induce traffic although the exact size of this effect could not be established at that time.

Although the M25 is a widely quoted example, and one which instigated public interest in the issue of induced traffic, there are of course many smaller road improvements where comparisons similar to the one above can be made. Goodwin (1996) summarises evidence from the SACTRA report and elsewhere, and compares the growth of traffic on improved corridors with the growth of traffic in a control corridor. If a control corridor was not available, then average growth rates were used for comparison. The results from this work have been adapted and are reproduced in Table 3.

The conclusions from this work were that the growth in traffic on the improved corridors, which appeared to be associated in some way with the road capacity change, was on average 25 per cent. The traffic growth element associated with the capacity change was found to increase over time, with unweighted averages of 9.5 per cent for less than a year, and 33 per cent for intervals of more than five years. Finally, while a reduction in traffic was observed on alternative routes to the improved route, this was on average only half as great as the increase on the improved route itself. In other words, the relief on alternative routes was not as great as had been forecast.

Despite the weight of evidence in support of the induced traffic phenomena, the lack of controlled experimentation means that it can never be categorically proven in the scientific sense. Although the definition of induced traffic requires a consideration of network-wide effects, in practice this is rarely possible and a consideration of corridor effects is often the best available evidence. Furthermore, it will never be possible to control the day-to-day variations that are symptomatic of urban traffic, nor to provide entirely satisfactory control sites and so eradicate the effect of changes in traffic due to exogenous variables such as economic growth or local migration (Bonsall, 1996; Hansen, 1998). This is the reality facing all transport planners – our transport ‘laboratories’ are not our own. Nevertheless, in conclusion to its own report to the UK Department of Transport in 1994, the SACTRA committee stated that (SACTRA, 1994: ii):

Considering all these sources of evidence, we conclude that induced traffic
Table 2: A comparison of design year AADT\(^1\) and 1992 actual AADT on London’s M25 orbital freeway

<table>
<thead>
<tr>
<th>M25 section</th>
<th>Design year</th>
<th>Design year two-way AADT</th>
<th>1992 actual two-way AADT</th>
<th>Difference between 1992 actual AADT and design year AADT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junction 13–14(^2,3)</td>
<td>1997</td>
<td>97 100</td>
<td>162 000</td>
<td>64 900</td>
</tr>
<tr>
<td>Junction 21a–22</td>
<td>2001</td>
<td>41 500</td>
<td>106 000</td>
<td>64 500</td>
</tr>
<tr>
<td>Junction 11–12</td>
<td>1995</td>
<td>82 800</td>
<td>146 000</td>
<td>63 200</td>
</tr>
<tr>
<td>Junction 22–23</td>
<td>2001</td>
<td>56 500</td>
<td>114 000</td>
<td>57 500</td>
</tr>
<tr>
<td>Junction 20–21</td>
<td>2001</td>
<td>59 000</td>
<td>113 000</td>
<td>54 000</td>
</tr>
<tr>
<td>Junction 10–11</td>
<td>1998</td>
<td>75 900</td>
<td>129 000</td>
<td>53 100</td>
</tr>
<tr>
<td>Junction 8–9</td>
<td>2000</td>
<td>55 200</td>
<td>107 000</td>
<td>51 800</td>
</tr>
<tr>
<td>Junction 19–20</td>
<td>2001</td>
<td>59 500</td>
<td>110 000</td>
<td>50 500</td>
</tr>
<tr>
<td>Junction 14–15</td>
<td>2000</td>
<td>103 000</td>
<td>152 000</td>
<td>49 000</td>
</tr>
<tr>
<td>Junction 15–16</td>
<td>2000</td>
<td>100 000</td>
<td>143 000</td>
<td>43 000</td>
</tr>
</tbody>
</table>

Notes:  
\(^1\)AADT = annual average daily two-way (traffic).  
\(^2\)The ten sites showing the greatest difference between design year two-way AADT and actual 1992 AADT were selected for inclusion in this table.  
\(^3\)Total number of junctions for which data were available was 19.  
Sources: SACTRA (1994: 46); Coombe (2000).

can and does occur, probably quite extensively, though its size and significance is likely to vary widely in different circumstances.

It may be argued that road capacity increases have indirect benefits; that is, they lead to commercial activity that may not have otherwise taken place. The SACTRA report is clear on this point, and suggests that while indirect benefits may occur that are not
Table 3: Summary of traffic impact of capacity increases at individual locations

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Interval¹</th>
<th>Result (after corrections, where necessary)²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barnstaple Bypass</td>
<td>3 years</td>
<td>+ 20% overall</td>
</tr>
<tr>
<td>M62</td>
<td>5 years</td>
<td>+ 19%</td>
</tr>
<tr>
<td>York Northern Bypass</td>
<td>Not clear</td>
<td>Redistribution, modal diversion and new trips 2% of interviewed drivers</td>
</tr>
<tr>
<td>Severn Bridge</td>
<td>1 year</td>
<td>Authors suggest induced traffic is 44%</td>
</tr>
<tr>
<td>Westway (London)</td>
<td>4 months</td>
<td>Corridor + 14% (control³ + 2%)</td>
</tr>
<tr>
<td>M11 (London)</td>
<td>9 years</td>
<td>Corridor + 38% (control + 29%)</td>
</tr>
<tr>
<td>A316 (London)</td>
<td>12 years</td>
<td>Corridor + 84% (control + 66%)</td>
</tr>
<tr>
<td>Blackwall (London)</td>
<td>1 year</td>
<td>Screenline + 15%</td>
</tr>
<tr>
<td>M25/Lea (London)</td>
<td>4 months</td>
<td>Corridor + 9%</td>
</tr>
<tr>
<td>Rochester Way</td>
<td>2 years</td>
<td>Corridors: West + 26%, East + 24%, Transverse + 30%</td>
</tr>
<tr>
<td>Leigh Bypass</td>
<td>1 year</td>
<td>Screenline + 20%</td>
</tr>
<tr>
<td>Manchester Ring</td>
<td>1 year</td>
<td>Corridors: East–west + 23%, North–south + 15%</td>
</tr>
</tbody>
</table>

Notes: ¹ 'Interval' indicates the time between capacity increase and analytical study.
² 'Result' indicates changes in traffic flow after capacity increases, as a percentage of previous flow on the link, corridor or screenline.
³ 'Control' indicates similar unimproved site, used for comparison purposes.

4.2 Evidence of suppressed traffic from before-and-after traffic counts

A useful summary of the impact of capacity reductions on network flows is provided by Cairns et al. (1998). The researchers collected evidence from over 40 locations on the impact of capacity reallocation as a result of bus lane implementation or pedestrianisation; maintenance or structural repairs; or natural disasters such as earthquakes. The results were wide-ranging, but the unweighted average overall reduction in traffic on the network was 25 per cent of the traffic previously using the road or area subject to capacity reduction. A randomly selected extract from their analysis is provided in Table 4. Figure 2 summarises the data on induced and suppressed traffic presented in Tables 3 and 4, to illustrate the fairly consistent impact the road capacity increases and reductions have on daily network traffic flows.

Since the publication of the reports by SACTRA and Cairns et al., the general consensus in the United Kingdom has been that induced and suppressed traffic phenomena occur. The current debate is over their magnitude and importance. In the United States, the debate so far has focused more on induced traffic, but the consensus also appears present there (Hansen, 1998).
Table 4: Summary of traffic impact of capacity reductions at individual locations

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Interval</th>
<th>Result (after corrections, where necessary)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tower Bridge closure</td>
<td>1 month</td>
<td>− 80%</td>
</tr>
<tr>
<td>Bologna city centre</td>
<td>8 years</td>
<td>− 51%</td>
</tr>
<tr>
<td>Edmonton-Kinnaird Bridge closure</td>
<td>3 weeks</td>
<td>− 42%</td>
</tr>
<tr>
<td>Partingdale Lane local area</td>
<td>3 months</td>
<td>− 30%</td>
</tr>
<tr>
<td>A13 closure</td>
<td>1 day</td>
<td>− 23%</td>
</tr>
<tr>
<td>Luneburg</td>
<td>3 years</td>
<td>− 15%</td>
</tr>
<tr>
<td>Cambridge city centre</td>
<td>5 months</td>
<td>− 11%</td>
</tr>
<tr>
<td>Freiburg ring road</td>
<td>10 months</td>
<td>− 7%</td>
</tr>
<tr>
<td>Edinburgh—New Town cordon</td>
<td>3 months</td>
<td>− 3%</td>
</tr>
<tr>
<td>Ring of Steel London ‘Square Mile’</td>
<td>1 year</td>
<td>− 0.2%</td>
</tr>
<tr>
<td>Frankfurt am Main bridge closure</td>
<td>Not clear</td>
<td>+ 2%</td>
</tr>
<tr>
<td>Aarau</td>
<td>1 day</td>
<td>+ 14%</td>
</tr>
</tbody>
</table>

Notes: The table indicates changes in traffic flow after road capacity reductions, as a percentage of the previous traffic flow in the area.


5. IMPLICATIONS

What, then, are the implications of this general consensus for transport planning practice, and for transport policy formulation? These are discussed in turn.

5.1 Practical implications

When considering the practical implications of these findings – and more particularly of induced traffic – it is necessary to reflect on how transport modelling and economic appraisal are generally undertaken; to examine whether current practice adequately reflects induced traffic; and to ask whether, in practical terms, this is important.

Induced traffic, as explained earlier in Table 1, is not ignored completely in the planning process. Indeed, some components are well entrenched in modelling practice (such as changes in route); others can be included theoretically, but are often not used in practice (such as changes in destination). This inclusion of induced traffic phenomena in the modelling process is described in Table 5.

Data providing an estimate of the scale of the traffic not generally included in the four-stage modelling process are scarce. Goodwin (1996) suggests, from a comparison of modelled and actual flows on nine urban road schemes, that unpredicted traffic in the first year of a scheme is almost 6 per cent. Heanue (1998), who examined traffic growth in Milwaukee during almost 30 years, suggests that the growth which could be attributed to road capacity improvements is between 6 and 22 per cent. Coombe (1996) tackles the problem rather differently, attempting to model all induced traffic effects and then to compare the results with conventional model outputs. In the congested urban case studies used, he finds that in the cases where traffic was induced, the overall short-term increase in trips was between 1 and about 3 per cent. One could argue whether these findings are significant, given the large errors inherent in conventional transport models (Atkins, 1986), but this train of debate is to an extent futile, as the
Figure 2: Changes in daily network traffic flows after road capacity changes
Note: Known time intervals of before-and-after studies are indicated in parentheses.

Table 5: Inclusion of induced traffic in the modelling process

<table>
<thead>
<tr>
<th>Changes in:</th>
<th>Induced traffic included? (vehicle km/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route</td>
<td>Generally yes</td>
</tr>
<tr>
<td>Mode</td>
<td>Generally no1</td>
</tr>
<tr>
<td>Vehicle occupancy (decreasing)</td>
<td>Generally no</td>
</tr>
<tr>
<td>Trip frequency (increasing)</td>
<td>Generally no2</td>
</tr>
<tr>
<td>Trip destination (becoming more remote)</td>
<td>Generally no3</td>
</tr>
<tr>
<td>Trip origin (becoming more remote)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: 1Assumptions about mode-choice changes may be made in forecasts, but a feedback loop between changes in generalised cost on the vehicle network, and the mode-choice model, are not common.
2Although in theory the trip destination model can be used for forecasting, with new inputs of future costs, in practice this is not often the case (Coombe, 1996).
3Generally, development is forecast exogenously and does not take into account the impact of the scheme under consideration on the access patterns.

largest impact of induced traffic appears to be in economic assessment, rather than in modelling output.

The economic appraisal of urban road schemes usually involves the processing of aggregate data from a transport model. In the first instance, a matrix of trip movements
is derived from base year data, and then this matrix is factored to some future year, using estimates of sociodemographic data, and any information relating to land-use development expected in the study area. Alternatively, traffic flows are measured on critical links, and the forecast new link flows are estimated using simple factoring techniques. The future-year matrix, or link flow data, is then used with varying road network supply scenarios, in order to derive data for the future year under consideration. The total time spent on the network by travellers (measured in vehicle hours), vehicle operating costs and accidents are then calculated for the do-minimum and do-something future-year scenarios. The difference in appropriately weighted time, operating and accident costs is an estimate of the economic benefit of the scheme. This information is the key input to the economic evaluation process. Note that in the evaluation process the trip matrix used is usually fixed; that is, it does not vary according to road supply (and hence generalised cost conditions) on the network. Given the explanation of induced traffic outlined above, this assumption of a fixed or inelastic demand matrix is erroneous, especially in cases of new road supply in congested urban conditions (Brand, 1992).

Some may argue that induced drivers receive a benefit from using the scheme, and so induced traffic must be advantageous. It is true that, under the fixed-matrix assumption, benefits to induced drivers are ignored. In most circumstances, however, the benefits to induced drivers are far outweighed by the delay costs imposed by induced drivers on the existing drivers on the network. These delay costs (measured in vehicle hours) are significant as, in the congested area of the speed-flow curve, even a small number of additional drivers on the network can impose substantial delay penalties on the network as a whole (Mackie, 1996). Thus, when considering drivers on the network as a whole, induced traffic is not advantageous.

The question which remains is: what practical difference does an assumption of a fixed-demand matrix make to economic evaluation? Coombe (1996) looked at the economic impact of relatively small increases in induced traffic. He found that in West London induced traffic of just 1 per cent led to an erosion of benefits of 30 per cent, and in Norwich induced traffic of 2.3–2.9 per cent led to reductions in benefits of 22–20 per cent. One reason for this is that the small increase in absolute vehicle hours, as a result of induced traffic, has a relatively large impact on the difference in vehicle hours between a do-minimum and do-something case. To understand this, it is important to note that a large proportion of the benefits in economic evaluations come from savings in vehicle hours. Typically the difference in vehicle-hours between a do-minimum scenario and a do-something scenario may be of the order of only a few per cent, depending on the scope of the network considered. An additional 1–3 per cent of induced traffic, in the do-something scenario, thus has a significant impact. Hence, SACTRA was clear in its statements regarding the importance of induced traffic for economic evaluation, stating that ‘the economic value of a scheme can be overestimated by the omission of even a small amount of induced traffic’ (SACTRA, 1994: iii).

In summary, to ignore the impact of induced traffic on economic evaluations can lead to seriously misleading economic evaluation results, and the subsequent allocation of funds to inappropriate road schemes. In particular, SACTRA’s findings indicate that ignoring the induced traffic phenomena would tend to overinflate the economic benefits of schemes where the network is operating at, or near to capacity, or where trips are suppressed by congestion.
5.2 Policy implications

With regard to implications for transport policy formulation, the evidence on induced traffic shows that an urban transport strategy with road capacity improvements at its core cannot bring relief to congestion and car dependency problems in the long term. If increased road capacity induces traffic, it follows that whatever policy for road capacity improvement is implemented, with the caveat of pre-existing congestion, the amount of traffic per unit of road will increase, not decrease. The same has been argued to be true of ‘advanced transport telematics’ or ‘intelligent transport systems’ aimed at the more efficient use of road capacity through driver guidance systems and improved traffic control (Bell, 1995). Mogridge (1997) goes one step further and suggests that improving urban road capacity can in fact make congestion worse, due to the impact that shifts in mode have on the viability of public transport, which in turn leads to further mode shifts. This implies, as Goodwin (1998) puts it somewhat bitingly, that policies based on road capacity improvements differ only with respect to ‘the speed at which congestion would get worse’.

If the evidence on induced traffic illustrates the futility of supply-side policies as a means of solving problems of urban congestion, the evidence on suppressed traffic perhaps provides a pointer to an appropriate policy alternative. There would, however, appear to be consensus in debates around the policy implications of induced and suppressed traffic that to argue simply for no future road capacity increases for private vehicles in urban areas would be absurd. This seems particularly true in a developing country with comparatively poor infrastructure and a need for rapid economic growth. Rather, what is called for in the literature is a balance between supply-side and demand-side strategies within an integrated transport policy framework (Bell, 1995; Goodwin, 1998). Within such a balanced framework, it is argued that new road construction would need to be justified in terms of its contribution to the implementation of a larger multi-modal transport plan, rather than simply in terms of its estimated congestion benefits for motorists. Some parts of the transport network may require road capacity increases, while other parts – particularly in city centres and residential neighbourhoods – may require selective reductions in road capacity and the reallocation of road space to non-motorised and public transport modes. In the context of growing travel needs, it is argued that selective reductions in road capacity for private vehicles would need to be accompanied by increases in the capacity and quality of the public transport network. Other aspects of such an integrated policy are seen to include a variety of measures for travel demand management that involve neither the increase nor decrease of capacity, but changes in the generalised cost of travel, such as congestion pricing and parking tariffs.

6. CONCLUSION

What, then, is the relevance of the empirical findings reviewed in this article, and what is the relevance of their practical and policy implications to the South African context? It is probable that choice passengers in South Africa perceive public transport to be less safe and less reliable than do their counterparts in countries like the United Kingdom. It is therefore assumed that, in the absence of significant safety and reliability improvements, South African choice passengers are less likely to switch from private car to public transport modes. Consequently, it is likely that induced traffic resulting from shifts from public transport modes to private car use will be considerably less in South Africa than in the United Kingdom, as a relatively larger proportion of choice
passengers already utilise private cars. Similarly, suppressed traffic resulting from shifts from private car to public transport, due to reduced generalised costs for the former, would also be considerably less. The other behavioural responses to capacity change discussed in Section 3 (i.e. changes in route, vehicle occupancy, trip frequency and destination) should, however, remain reasonably constant across the different contexts, and the direction of impacts should therefore remain true in South African cities even if the scale of impact is different. Meta-analysis of South African before-and-after studies will, however, be necessary to test the assumptions embedded in this argument.

From the particular interpretation of the relevance of the evidence reviewed in this article, it is therefore concluded that strategies for addressing traffic congestion problems in South African cities, based primarily on supply-side road capacity improvements (in the form of either new infrastructure provision or measures for transport systems management aimed at achieving better utilisation of existing capacity), will be unable to bring relief in the longer term. Balanced supply- and demand-side strategies that incorporate new infrastructure provision, transport system management and travel demand management, integrated with land-use strategies that reduce the need for travel, offer the only realistic longer-term prospects for relieving congestion and improving accessibility for the populations of South African cities.

In the authors’ view, however, it is necessary to reframe the urban passenger transport problem from one focused primarily on inefficiency or traffic congestion, to one focused equally on inefficiency, inequity and non-sustainability. The former problem framing leads to the allocation of road space resources to alleviate congestion for general traffic. The latter leads to a more disaggregated and dedicated allocation of road space resources to benefit all passengers equally and to encourage the use of high-occupancy and low-polluting travel modes. From this perspective, given the historical injustices imposed on passengers captive to public transport and non-motorised modes in South African cities, the authors believe it to be entirely justifiable and appropriate to reallocate proportionate shares of road space, in the form of footways, cycleways and dedicated bus lanes, for exclusive use by pedestrians, cyclists and public transport users, even if this imposes greater congestion costs on private motorists. The evidence reviewed here suggests that, provided policy commitments to improving the quality and safety of public transport services are followed through, this would not lead to the general traffic chaos many would expect.

REFERENCES


