KU LEUVEN

CENTER FOR ECONOMIC STUDIES

DISCUSSION PAPER SERIES
DPS15.14

JULY 2015

The myth of traffic-responsive signal control: why common sense does not always make sense

Ruth EVERS & Stef PROOST

Energy, Transport and Environment

Faculty of Economics
And Business

The myth of traffic-responsive signal control: why common sense does not always make sense.

Ruth Evers*

Stef Proost†

Abstract

Intuitively, one is inclined to think that traffic-responsive signal control is the most efficient control policy. In this paper, however, we show that for an intersection of two routes connecting one origin-destination pair where only one route is subject to congestion, anticipatory signal control performs better than traffic-responsive signal control. Furthermore, the unfolded logic behind this result suggests that the superiority of anticipatory signal control also extends to other networks.

Keywords: anticipatory control, Stackelberg game, traffic-responsive control

1 Introduction

A lot of the current signal control systems are based on traffic-responsive control. This type of control allocates green time in proportion to the relative magnitude of the flow. In this paper, we show that, though intuitively superior, this type of control is not necessarily the most efficient. The main reason is that in traffic networks the user equilibrium is often not optimal. Blindly attempting to accommodate to the volume of traffic on a link with congestion, by adding capacity or by giving more green time, is a widespread problem, which can, more generally, be ascribed to the phenomenon of "induced demand".

The theory of induced demand asserts that improvements in the transportation infrastructure attract new traffic. The available literature has largely centered around the demand-inducing and traffic diversion effects of particularly road expansion (Downs (1962), Braess

*KU Leuven, Center for Economic Studies, Naamsestraat 69, 3000 Leuven, Belgium. Email: ruth.evers@kuleuven.be

†KU Leuven, Center for Economic Studies, Naamsestraat 69, 3000 Leuven, Belgium.

(1968), Noland (2001)). This literature has provided a basis for a major rethinking of road-expansion policies. Recognition of the generated traffic effects of traffic-responsive control can influence policy making in the same way.

In many cities, signal control is of the traffic responsive type. The control system SCOOT, for example, has been implemented in more than 250 towns and cities (Hamilton et al. (2013)). The rapid and widespread implementation of traffic-responsive signal control is strongly connected to the intuitive superiority of this control policy, which made it politically more acceptable. An equal intuitive discourse is needed to challenge this inclination towards responsive signal control. The objective of this paper is therefore to provide a clear and accessible comparison of responsive signal control versus anticipatory signal control, which provides insight for policy making.

Already in 1974, the need to take into account the interaction between route choice and signal control was pointed out by Allsop (1974). In most papers nowadays the importance of this interaction is recognized and the interaction is thus included in the model. The way this interaction is modelled differs, however, from paper to paper. Here, two specific ways to model this interaction, i.e. anticipatory signal control and responsive signal control, are considered.

Miller (1963) was the first to introduce the notion of traffic-responsive control. In traffic-responsive control, data collected from vehicle detectors located upstream is used to optimize the signal settings.
The strand in the literature that deals with traffic-responsive signal control focuses on the iterative optimization and assignment procedure. In the iterative optimization and assignment procedure the signal settings and equilibrium flow patterns are updated alternatively, until both flows are at equilibrium and signal settings are optimal given the flows (Allsop and Charlesworth (1977), Cantarella et al. (1991), Gartner et al. (1980), Lee and Hazelton (1996)).

In the case of anticipatory signal control the road authority anticipates the reaction of the drivers to a change in the signal settings and thus optimizes the signal settings taking into account the reaction of the drivers. In the literature, this has been formulated as a bi-level problem, in which the upper level is the signal setting problem and the lower level is the traffic equilibrium assignment problem (Chiou (1999), Yang and Yagar (1995)), and as a Stackelberg game (Fisk (1984)).

A few papers have touched on the shortcomings of responsive signal control. Both Gershwin and Tan (1978) and Dickson (1981) have solved the combined traffic assignment and control problem for a specific numerical example, using on the one hand the iterative optimization and assignment procedure and on the other hand a constrained optimization approach. For their specific examples, both show that the iterative optimization and assignment leads to a worse solution.

In this paper we use a game theoretical perspective to model both the anticipatory and the responsive signal setting procedure for a simple network. This results in clear theoretical results, which allow to give insights in the underlying mechanisms. The remainder of this paper is organized as follows: First, the problem at hand is described in Section 2. Subsequently, the outcomes of the traffic-responsive and anticipatory framework are compared and discussed in Section 3. Section 4, finally, offers a conclusion.

2 Problem formulation

The model we use to compare the performance of anticipatory signal control with the performance of traffic-responsive signal control is the simple two-road model represented in Figure 1. We assume that, per time unit, N people want to go from A to B[1]. Furthermore, we limit the model to undersaturated traffic conditions, i.e. queues at the intersection are only created during the red phases and dissolved during the green phases. To go from A to B, drivers can either take a congestible route (Route 1) or an uncongestible route (Route 2)[2]. In this paper, the red phase on Route 2 is represented by 'r', and will be the main control variable. The corresponding green phase on Route 2 will thus be 'c-r', and a reverse scenario holds for Route 1. The duration of the sum of the red and the green phase is the cycle time 'c', which, to simplify matters, is held fixed. Hence, it follows that including intergreen time in the analysis is not relevant and will thus be ignored.

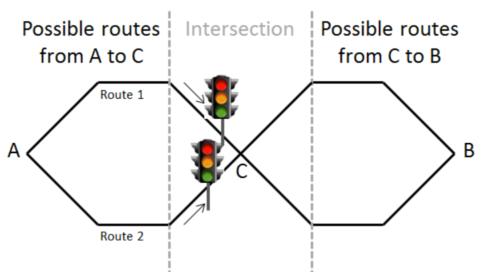

Figure 1: Outline of an intersection of two routes connecting one OD-pair (AB) regulated by traffic lights.

[1]The arrival rate is thus inelastic, static and deterministic.
[2]Up to N drivers per hour, the time cost curve for this route is horizontal.

As there will be alternating red times to avoid collisions at the intersection, drivers on both routes will experience an expected traffic light waiting time cost $(T_1(c,r), T_2(c,r))$. It is clear that the expected traffic light waiting cost functions are increasing in the red time and decreasing in the green time ($\frac{\partial T_1(c,r)}{\partial r} < 0, \frac{\partial T_2(c,r)}{\partial r} > 0$). Assuming undersaturated traffic conditions, the expected traffic light functions take the following form for $0 < r < c$[3]:

$$T_1(c,r) = \frac{(c-r)^2}{2c} \quad (1)$$

$$T_2(c,r) = \frac{r^2}{2c} \quad (2)$$

When it is always red ($r = c$ for Route 2 and $r = 0$ for Route 1) the expected traffic light waiting function jumps to infinity.

We will model both the anticipatory and the traffic-responsive control and assignment problem as a Stackelberg game (Von Stackelberg (1934)). The Stackelberg game is a sequential game in which the leader moves first and the follower acts sequentially. In this paper, the traffic authority is the leader when the signal control is anticipatory and the traffic authority is the follower when the signal control is traffic-responsive. The objective of the traffic authority is to maximize welfare. The drivers in turn represent the follower when the signal control is anticipatory and the leader when the signal control is traffic-responsive. We will assume that all drivers are identical and try to minimize their expected travel cost.
The behaviour of the drivers can also be represented as a game, because the congestion on one road is dependent upon how many users choose to use the same road. In this paper, we will assume that the drivers behave non-cooperatively (Nash (1951)).

3 Results

3.1 Traffic-responsive signal control

When the signal control is traffic-responsive, the signal settings respond to the current traffic conditions measured by a vehicle detector. This situation is represented as a Stackelberg game in Figure 2. The game tree shows all the possible distributions of the drivers over the two routes and all the policies the government can implement in reaction to the drivers' choice.

To predict the outcome of this game, we will first determine the best response of the road authority to every possible distribution of the drivers over the two routes. Assuming that the government wants to maximize welfare, their best response to any distribution over the

[3] We will assume that the saturation flow rate g is very large in comparison to the arrival rate X_i, so that the traffic light waiting time due to departure delay is negligible.

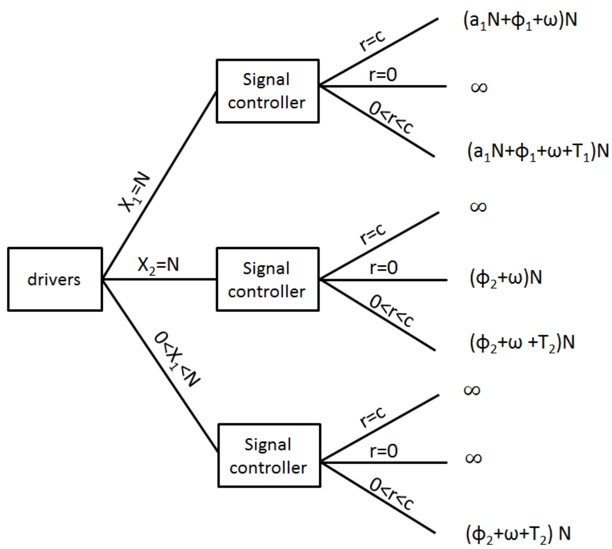

Figure 2: Total costs when signal control is traffic-responsive.

two routes is to maximize welfare taking as given this distribution. From the game tree, it is clear that the traffic authority will always give green to Route 1 when all drivers choose Route 1. When all drivers choose Route 2, the rational decision for the traffic authority is to give always green to Route 2. When the drivers divide themselves over the 2 routes, the optimal response of the traffic authority is the solution of the following optimization problem.

$$\min_{r} \left(a_1 X_1 + \omega + \phi_1 + T_1\left(r\right)\right) X_1 + \left(\omega + \phi_2 + T_2\left(r\right)\right) X_2 \tag{3}$$

$$0 < r < c \tag{4}$$

In this optimization problem a_i represents the sensitivity to congestion of route i, X_i equals the flow on route i[4], ϕ_i stands for the minimal time cost of route i, ω is the resource cost for a trip from A to C on either route and $T_i\left(c, r\right)$ is the expected waiting time cost on route i at the intersection.

Taking the derivative of (3) to r and taking into account that $X_1 + X_2 = N$, we find that the optimal strategy for the government is to implement the following red time:

$$r^* = \frac{X_1 c}{N} \tag{5}$$

[4]This is the flow measured by the vehicle detector.

Making this trip day in day out, the drivers will come to learn the optimization formula of the government, which is easy to understand as the red time is inversely proportional to the flow. As a result, their private cost functions have the following form for $0 < X_1 < N$:

$$a_1 X_1 + \phi_1 + \omega + \frac{cX_2^2}{2N^2} \tag{6}$$

for Route 1 and

$$\phi_2 + \omega + \frac{cX_1^2}{2N^2} \tag{7}$$

for Route 2. Remark that $\frac{cX_2^2}{2N^2}$ is the expected traffic light waiting cost for Route 1 ($\frac{(c-r^*)^2}{2c}$, with $r^* = \frac{X_1 c}{N}$) and $\frac{cX_1^2}{2N^2}$ is the expected traffic light waiting cost for Route 2 ($\frac{r^{*2}}{2c}$, with $r^* = \frac{X_1 c}{N}$).

Drivers will individually seek to minimize their private cost and will consequently change routes until unilaterally changing increases their private cost. At that point the stationary distribution of vehicles in the network, i.e. the equilibrium, is reached. And this stationary distribution will thus be the outcome of the drivers' part of the game which, together with the government's strategy, will determine the total cost of responsive signal control.

Remark that, even though the signal setting policy of the traffic authority provides the drivers with an opportunity to manage the actions of the traffic authority, they can not exploit this advantage as the drivers do not cooperate.

Depending on the parameter values and the amount of drivers on Route 1, the average cost curve is either downward sloping ($\frac{dAC_1}{dX_1} < 0$) or upward sloping ($\frac{dAC_1}{dX_1} > 0$). The average cost on Route 1 can be downward sloping because an increase in volume on Route 1 implies a decrease in volume on Route 2. With responsive signals this implies a longer green phase on Route 1 which can outweight the increased congestion on Route 1. In this paper, we will determine the equilibria for only one instance: the waiting cost always outweights the congestion cost ($\forall X_1 : \frac{dAC_1}{dX_1} < 0$). We can restrict ourselves to this instance, as it suffices to show the superiority of anticipatory control.

When the waiting cost outweights the congestion cost for all possible distributions of vehicles over the two routes, we can distinguish between three cases: either the average cost of Route 1 is always larger than the average cost of Route 2 or the other way around, or the average cost curves intersect. We will determine the equilibria for each of these cases separately.

Case 1. If for every distribution of vehicles over the two routes AC_1 is larger than AC_2, then the only equilibrium is $X_2 = N$.

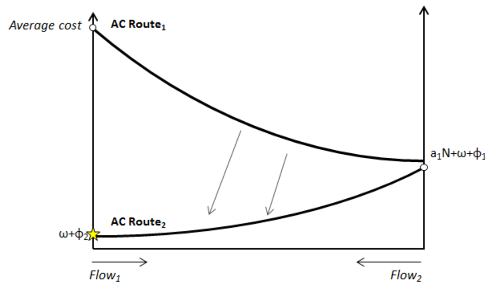

Figure 3: The only stationary distribution is $X_2 = N$

The search for potential equilibria in the set of possible distributions of vehicles over the 2 routes is greatly simplified by Wardrop's first principle (Wardrop (1952)): In a user equilibrium, all used routes for an OD-pair should have equal generalized prices, and there are no unused routes with lower generalized prices. For this case, the first part of the principle, eliminates all distributions in which both routes are used. The second part of the principle, eliminates the outcome where all drivers are on Route 1. The only distribution left, $X_2 = N$, satisfies the definition of a Nash equilibrium. The total cost in this case equals $(\omega + \phi_2)\, N$.

Case 2. If there exists a distribution of vehicles over the two routes for which $AC_1 = AC_2$, then there are two potential equilibria: $X_2 = N$ or $X_1 = N$.

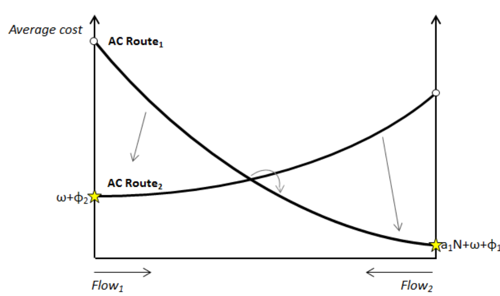

Figure 4: There are two stationary distributions: $X_2 = N$ and $X_1 = N$

In this case, Wardop's first principle leaves us with three potential equilibria: $X_2 = N$, $X_1 = N$, and $X_1 = \frac{\left(\frac{c}{2}+\phi_1-\phi_2\right)N}{c-a_1 N}$. When the distribution of vehicles is such that the average cost of both routes is the same, then a driver on Route 1 could lower his cost by unilaterally changing to Route 2 (or a driver on Route 2 could lower his cost by unilaterally changing to Route 1). So depending on the initial distribution of vehicles over the two routes, either $X_2 = N$ or $X_1 = N$ will be the stationary distribution. The corresponding total cost equals $(\omega + \phi_2) N$ when the equilibrium is $X_2 = N$ and $(a_1 N + \omega + \phi_1) N$ when $X_1 = N$.

Case 3. If for every distribution of vehicles over the two routes AC_2 is larger than AC_1, then the only equilibrium is $X_1 = N$.

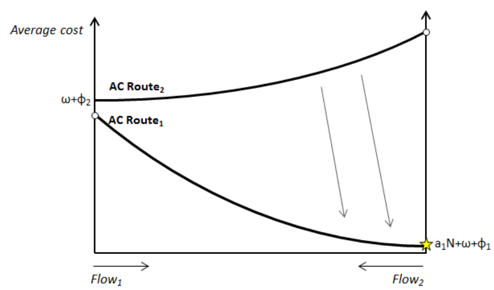

Figure 5: The only stationary distribution is $X_1 = N$

If for every possible distribution of vehicles over the two routes, the private cost when taking Route 2 is higher than the private cost of taking Route 1, then all drivers will take Route 1. This dominant strategy leads us to the equilibrium distribution $X_1 = N$. The private cost every driver will incur equals $a_1 N + \omega + \phi_1$, and the total cost thus amounts to $(a_1 N + \omega + \phi_1) N$.

3.2 Anticipatory signal control

When signal control is anticipatory, the traffic authority moves first and bases its decision on the expected reaction of the drivers. Figure 6 shows the different options for the government, and the possible reactions of the drivers.

The traffic authority has three possible policies: to grant always green to Route 1, to grant always green to Route 2, or to implement an alternating signal setting[5]. The users in turn can react in three different ways to any chosen policy: to only take Route 1, to only use Route 2, or to use both routes.

[5]The third branch, representing the decision of the government to implement an alternating signal setting $(0 < r < c)$, is a clubbing of all red times between zero and c.

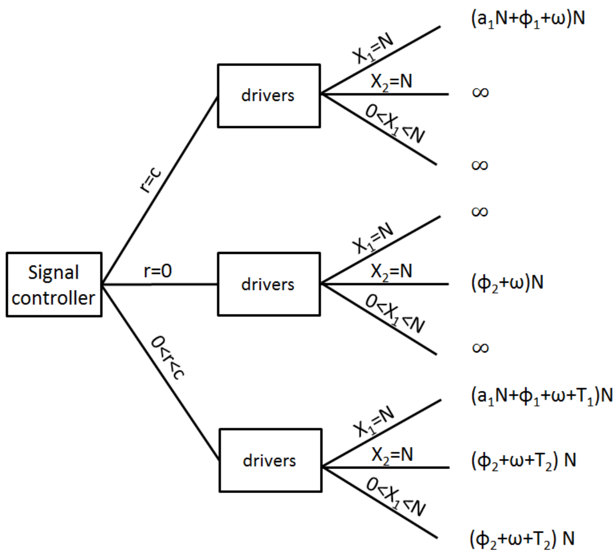

Figure 6: Total costs when signal control is anticipatory.

If the traffic authority decides to implement always red for Route 1, then all drivers will take Route 2 and the total cost will equal $(\omega + \phi_2) N$. If, on the other hand, the traffic authority grants always green to Route 1, then the user equilibrium will be $X_1 = N$ and the total cost will be $(a_1 N + \omega + \phi_1) N$. If, however, the government decides to implement an alternating signal setting, the equilibrium reaction of the drivers will be $X_1 = N$ if $a_1 N + \phi_1 + T_1(c, r) < \phi_2 + T_2(c, r)$, and $X_2 = N$ if $\phi_1 + T_1(c, r) > \phi_2 + T_2(c, r)$. If $\phi_1 + T_1(c, r) < \phi_2 + T_2(c, r) < a_1 N + \phi_1 + T_1(c, r)$, the drivers will use both routes, and the Wardrop equilibrium implies $(\omega + \phi_2 + T_2(c, r)) N$ as total cost.

A glance at Figure 6 reveals that a rational government will never decide on an alternating signal setting when both routes are substitutes. Indeed, let $T_i(c, r)$ be the expected waiting time cost on route i at the intersection. As $T_i(c, r)$ is positive when $0 < r < c$, the total travel cost for an alternating signal setting will always be higher than for $r = c$ or $r = 0$. Which of the two non-alternating signal settings will be optimal depends on the values of the parameters a_1, N, ϕ_1 and ϕ_2. Whenever $a_1 N + \phi_1 < \phi_2$, $r = c$ is the optimal solution and whenever $a_1 N + \phi_1 \geq \phi_2$, $r = 0$ will be implemented.

3.3 Anticipatory versus traffic-responsive signal control

From the analysis in Section 3.2, we know that when $a_1 N + \phi_1 < \phi_2$, the total cost amounts to $(a_1 N + \omega + \phi_1) N$ and when $a_1 N + \phi_1 \geq \phi_2$, the total cost equals $(\omega + \phi_2) N$ when traffic control is anticipatory. The outcome in each of the two scenarios is not so straight-forward in case traffic-responsive control is implemented[6], so it is rather cumbersome to directly compare total costs. However, the following line of reasoning allows to asses the relative performance of anticipatory and traffic-responsive control in an indirect way.

Comparing the outcomes of anticipatory signal control (Figure 6) to the possible outcomes of traffic responsive signal control (Figure 2), it is clear that the performance of traffic responsive signal control can only be equally well or worse than anticipatory signal control. Consequently, if there exists one case for which the performance of traffic responsive signal control is worse, we can conclude that the overall expected perfomance of traffic-responsive signal control is worse than anticipatory signal control.

Take the scenario in which $a_1N + \phi_1 < \phi_2$, then out of the three cases we have dealt with in section 3.1, Case 2 and Case 3 can occur. If Case 2 occurs the total cost is either $(a_1N + \omega + \phi_1)\,N$ or $(\omega + \phi_2)\,N$. If Case 3 occurs, the total cost is $(a_1N + \omega + \phi_1)\,N$.

Remember that in this case the total cost with anticipatory signal control equals $(a_1N + \omega + \phi_1)\,N$. The possible outcome of the stationary distribution $X_2 = N$, resulting in a signal setting $r = 0$ and total cost $(\omega + \phi_2)\,N$, when $a_1N + \phi_1 < \phi_2$ thus proves that there exists at least one case for which the performance of traffic responsive signal control is worse than the performance of anticipatory signal control. As a result, we can assert that for our model, the performance of anticipatory signal control is superior to the performance of traffic-responsive signal control.

This result can be explained by recognizing the presence of externalities and the first mover advantage. Because of externalities, the drivers' individual choices are not socially optimal. Or, putting it differently, every driver minimizes his own cost, but this does not necessarily minimize the cost of all drivers. The traffic authority's objective is to minimize the cost of all drivers, so its unconstrained decisions are socially optimal. However, in both the traffic responsive control problem and the anticipatory control problem the traffic authority's optimization problem is constrained, to a greater or lesser extent, by the behaviour of the drivers. When the traffic authority is the leader, she can act so as to elicit the most favorable response of the driver. However, when the traffic authority is the follower, her influence on the drivers' behaviour is more restricted. Analytically, when signal control is anticipatory the leader's optimization problem is constrained by the drivers' reaction function, while when signal control is traffic-responsive the leader's optimization problem is constrained by the drivers' individually optimized distribution over the different routes. It is clear that the traffic authority's constraint is much more restricting when signal control is traffic-responsive. This also becomes apparent in Figure 7 below, in which the total cost

[6] As it depends on the relative values of some of the parameters.

is compared in case the traffic authority is the leader and when the traffic authority is the follower.

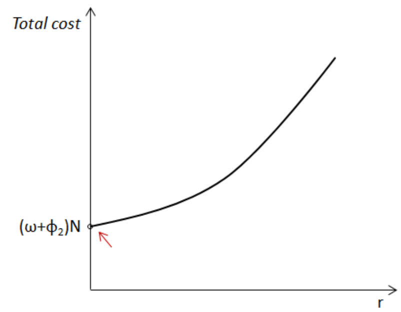

Figure 7: Comparison of the total cost when traffic control is anticipatory and traffic-responsive for an interior equilibrium.

When both routes are used in equilibrium, the anticipatory signal setting will always be the lowest red time possible (on Route 2), resulting in the lowest costs (see red arrow in Figure 7). However, when signal control is responsive, the equilibrium signal setting equals $r = (\frac{\phi_2 - \phi_1 - \frac{c}{2}}{a_1 N - c})c$, which, depending on the parameter values, results in a total cost that is at least as high as the lowest cost.

Our discussion allows to extend the obtained results to other networks. Indeed, the above reasoning can be applied as well to other situations in which signal settings can influence route choice. If there is no route choice, then the traffic-responsive control coincides with anticipatory control.

4 Conclusion

In this paper we have shown that for an intersection of two routes connecting one O-D pair where only one route is subject to congestion (1) traffic responsive signal control can only perform just as well or worse than anticipatory signal control and that (2) the expected performance of traffic responsive signal control is worse than the performance of anticipatory signal control. The game theoretic perspective taken in this paper furthermore suggests that these results can also be extended to larger instances.

These results have important implications for policy. The counter-intuitiveness of these results indicates that great attention needs to be given to the accuracy of the appraisal of signal control investments. Since both the costs of road transportation infrastructure and user costs are large, policy should be based on careful analysis rather than on intuition alone. The allocation of public money to the intuitively superior traffic-responsive signal control, may actually make society worse off as the money could be more efficiently spent on anticipatory signal control. This paper furthermore intends to raise awareness that policies based on intuition alone can have unintended consequences in the hope that these can be recognized and avoided.

A final note on the results in this paper concerns the deterministic nature of demand in this paper. An interesting extension would deal with stochastic demand as the flexible nature of traffic-responsive signal control could mitigate the advantage of anticipatory control in this case.

Acknowledgements

This work was supported financially by the OT/11/068-project of the KU Leuven. The authors wish to thank C. Tampère and M. Smith for helpful suggestions on a previous draft.

References

References

Allsop, R. (1974). Some possibilities for using traffic control to influence trip distribution and route choice. In *Transportation and Traffic Theory, Proceedings*, volume 6.

Allsop, R. and Charlesworth, J. (1977). Traffic in a signal-controlled road network: An example of different signal timings including different routeing. *Traffic Engineering and Control*, 18(Analytic).

Braess, P.-D. D. D. (1968). Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, 12(1):258–268.

Cantarella, G. E., Improta, G., and Sforza, A. (1991). Iterative procedure for equilibrium network traffic signal setting. *Transportation Research Part A: General*, 25(5):241–249.

Chiou, S.-W. (1999). Optimization of area traffic control for equilibrium network flows. *Transportation Science*, 33(3):279.

Dickson, T. J. (1981). A note on traffic assignment and signal timings in a signal-controlled road network. *Transportation Research Part B: Methodological*, 15(4):267–271.

Downs, A. (1962). The law of peak-hour expressway congestion. *Traffic Quarterly*, 16(3).

Fisk, C. (1984). Game theory and transportation systems modelling. *Transportation Research Part B: Methodological*, 18:301 – 313.

Gartner, N. H., Gershwin, S. B., Little, J. D., and Ross, P. (1980). Pilot study of computer-based urban traffic management. *Transportation Research Part B: Methodological*, 14(1):203–217.

Gershwin, S. and Tan, H.-N. (1978). Hybrid optimization: optimal static traffic control constrained by drivers route choice behavior. In *Decision and Control including the 17th Symposium on Adaptive Processes, 1978 IEEE Conference on*, volume 17, pages 1042–1047. IEEE.

Hamilton, A., Waterson, B., Cherrett, T., Robinson, A., and Snell, I. (2013). The evolution of urban traffic control: changing policy and technology. *Transportation Planning and Technology*, 36(1):24–43.

Lee, S. and Hazelton, M. (1996). Stochastic optimisation of combined traffic assignment and signal control junction modelling. In *Transportation and traffic theory. Proceedings of the 13th international symposium on transportation and traffic theory, Lyon, France.*

Miller, A. J. (1963). A computer control system for traffic networks.

Nash, J. (1951). Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295.

Noland, R. B. (2001). Relationships between highway capacity and induced vehicle travel. *Transportation Research Part A: Policy and Practice*, 35(1):47–72.

Von Stackelberg, H. (1934). *Marktform und Gleichgewicht.-Wien & Berlin: Springer 1934. VI, 138 S. 8.* J. Springer.

Wardrop, J. (1952). Some theoretical aspects of road traffic research. In *Proceedings of Institution of Civil Engineers*, volume 1, pages 325–378.

Yang, H. and Yagar, S. (1995). Traffic assignment and signal control in saturated road networks. *Transportation Research Part A: Policy and Practice*, 29(2):125 – 139.