The Interaction Between Signal-Setting Optimization and Reassignment: Background and Preliminary Results

TOM VAN VUREN, MICHAEL J. SMITH, AND DIRCK VAN VLIET

It is well known that signal-setting policies and traffic assignment mutually influence each other. It is not always certain that an equilibrium can be established between both. With the two most commonly used policies, Webster's and a delay-minimizing one, there may be many such equilibria, some of them unstable. A third policy, P_0, has been designed to have substantially better equilibrium behavior. The characteristics of these three policies are discussed as far as their influence on assignment is concerned. An empirical comparison of the behavior of Webster's policy and that of P_0, especially with regard to stability and delays, is presented for a small network. The results for the P_0-policy appear to be promising.

Traffic signals are useful tools in urban traffic control systems. Over the years several signal-setting policies have been developed for isolated junctions, for example, those of Webster *(1)*, Miller *(2)*, and Allsop *(3)*. Usually these policies try to minimize some measure of delay at each junction for the vehicles in the network.

If it is assumed that drivers choose their routes to minimize their own travel time or cost, so that a Wardrop equilibrium results, this kind of signal-setting policy will obviously influence the assignment of traffic over the network, because changes in green times will change costs for the various routes. On the other hand, changes in assigned flows will influence the delays experienced and thus change the optimal signal settings.

This interaction is the basic theme of this paper. The task is to determine a point at which signal settings and assignment are in equilibrium.

NOTATION

- a_{ij} = 1 if turning movement i runs during state j, 0 otherwise;
- C = cycle time;
- d_i = delay at traffic signals for movement i;
- f_i = flow on movement i;
- g_i = signal green time for movement i;
- l_i = link travel times for movement i;
- s_i = saturation flow for movement i;
- T = time period considered; and
- λ_i = green-time proportion for movement $i = g_i/C$.

T. van Vuren and D. Van Vliet, Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, United Kingdom. M. J. Smith, Department of Mathematics, University of York, Heslington, York YO1 5DD, United Kingdom.

TWO WELL-KNOWN POLICIES

Two commonly used signal-setting policies are introduced, both aiming to minimize some measure of delay explicitly.

Webster's Policy

The essence of Webster's policy is

$$\mathrm{Min}_j \ \mathrm{Max}_i \ a_{ij} f_i/(\lambda_i s_i) \tag{1}$$

or, in words, minimize the maximum flow-to-capacity ratio for a movement over all stages by adjusting green times. Unless a boundary is reached (minimum green time, etc.) this policy will try to equalize the flow-to-capacity ratios for the maximally loaded movements in each stage. The effectiveness of this policy lies in the fact that delays increase more than linearly with increasing $f/\lambda s$-values (flow-to-capacity ratios). Therefore it is beneficial to keep the maximum values as low as possible.

Delay-Minimizing Policy

The objective of the delay-minimizing policy is straightforward: minimize the total delay that is experienced at the observed junction *(3)*:

$$\mathrm{Min} \sum_i f_i d_i \tag{2}$$

PROBLEMS

It has been shown theoretically *(4)* that when Webster's policy is used, there may be many equilibria for assignment and signal settings, some of them unstable. Others [Allsop and Charlesworth *(5)*] found empirically that indeed for a certain network there was no unique equilibrium for the delay-minimizing policy: results strongly depend on the initial settings or assignment.

It can easily be explained why these policies do not behave well. Because the policies try to minimize total delay, the most heavily loaded arms of the junction will be awarded the most green time. This, however, will increase delays on other arms, thus "pulling" more traffic to the already heavily loaded arms that received more green. This in turn requires the signal setting to be changed in favor of these same arms, and so forth. So a self-enforcing process results. In this way the objective of the policy (to minimize delay) will not always be met because, in effect, rerouting may even cause the average delays to increase.

This now is the basic deficiency of the traditional delay-minimizing policies: because they do not take into account changes in assignment as a result of the signal-setting policy, their green-time settings are based on outdated flows and as a result are not optimal.

ANOTHER POLICY

Braess's paradox *(6)* shows that, if possible, roads with high marginal costs should be avoided. In this light the signal-setting policy P_0 that Smith *(7)* proposes is very appealing. In essence the objective is

$$\text{Equalize} \sum_i a_{ij} s_i d_i \quad \forall \text{ stage } j \tag{3}$$

For each stage the sum of the experienced delays for all the movements that run during that stage, weighted by their saturation flows, is equalized. Because of this weighting, green time is assigned to the wider roads, even if currently these roads are little used, so that traffic is pushed toward these wider roads. It is proved that under natural but rather severe conditions there will be a single stable equilibrium *(4, 8)*.

The actual goal of this policy is a maximization of the capacity of the network. It is hoped that as a side effect a decrease in delays and travel times will appear, especially at higher levels of congestion. The advantage of P_0 over the traditional delay-minimizing policies is that instead of adjustment of the signals to the current flow situation (without consideration of rerouting) the aim is a future goal, namely, to maximize network capacity by steering drivers toward the wider roads. So rerouting is actually an explicit objective of the policy.

OBJECTIVES OF THE STUDY

Until now, policy P_0 has only been analyzed theoretically, usually with emphasis on stability characteristics. However, another important feature is its influence on the quality of the network; in other words, will it actually cause a decrease in delays?

The aim of this study is to test the various characteristics of policy P_0 by applying it and comparing its results with those of familiar policies. These tests have been made with the simulation and assignment model SATURN *(9)*. The strength of this model lies in the detailed simulation of junctions, which gives more accurate flow and delay curves, together with a Wardrop equilibrium assignment model, as shown in Figure 1.

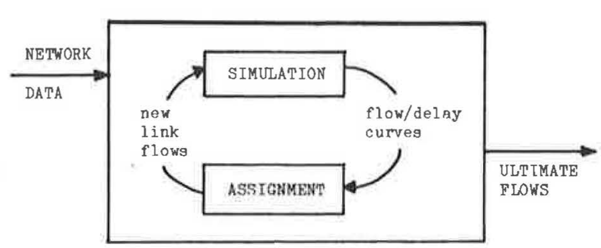

FIGURE 1 Basic structure of SATURN.

The method of testing the various signal-setting policies used here follows naturally from the way in which signals and assigned flows influence each other in reality. After Dickson *(10)* and others, the method is called the iterative optimization reassignment procedure; signal settings and flows are changed alternately until an equilibrium between both is reached (Figure 2).

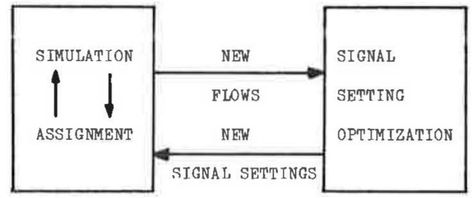

FIGURE 2 Iterative optimization reassignment procedure.

The first tests were carried out on a small network, so the influence of certain network characteristics, such as route lengths, congestion level, and initial signal settings, could be readily distinguished. The effects of policy P_0 compared with the effects of Webster's policy are presented here for the first time, in terms of both the uniqueness and stability of the attained equilibrium and the influence on network delays. The ultimate tests on real-life larger-scale networks are being carried out and will be published later.

RESULTS

The test network (Figure 3) consists of a short, quick route (e.g., through a city center) and a longer but wider route (e.g., a bypass). The main results presented for this network will concern

- Green times at equilibrium,
- Assigned flows at equilibrium, and
- Delays and travel times at equilibrium.

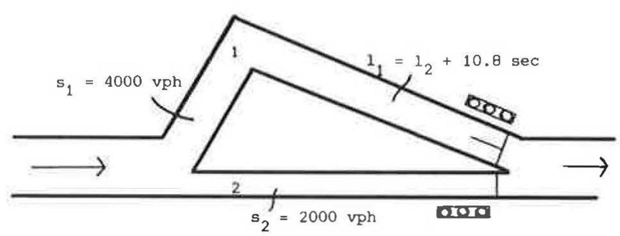

FIGURE 3 Test network.

The major assumptions that were made are as follows:

- Cycle time of 60 sec;
- No intergreen times, so the green times add up to the cycle time;
- Two stages, one for each road;
- Minimum and maximum green times of 0.5 and 59.5 sec, respectively;

- Observed time period T of 30 min or 1,800 sec;
- Delays calculated by a three-term so-called sheared delay formula consisting of geometrical delays, random delays, and queueing delays (above capacity).

Note that for this simple network the two policies tested reduce to

$$f_1/(\lambda_1 s_1) = f_2/(\lambda_2 s_2) \quad \text{(Webster)}$$

$$s_1 d_1 = s_2 d_2 \quad (P_0)$$

Green Times at Equilibrium

Figure 4 shows the resulting green times at equilibrium for the two policies tested. It is obvious that with the Webster policy the iterative optimization reassignment procedure always causes the signal settings to reach one of the two extremes (i.e., minimum or maximum green times), but more important is the fact that the actual boundary reached is determined by the initial signal settings. There turn out to be three equilibria for the signal settings when Webster's signal-setting policy is applied—the two extremes and an intermediate, which is unstable. Figure 5 shows these equilibria for the various flow levels. Evidently the two equilibria at the boundaries will lead to totally different delays and flows. The two boundary signal settings will be called Upper Webster and Lower Webster.

The P_0-policy gives rise to an equilibrium at a 35/25 setting for low total flows, changing to a 24/36 setting at a flow of 1,073 vph. At that point, green times for the wider route increase until at a 3,910-vph flow level all green time is assigned to this route.

Flows at Equilibrium

The distribution of green times is strongly related to the distribution of flows over the two routes. Figure 6 shows this distribution of flows and again the Webster policy reaches a boundary, which depends on the initial signal settings.

The flows tend naturally to follow the green times: Upper Webster distributes all traffic to the long, wide route until capacity is reached; then some traffic (about 1 percent) is also distributed to the shorter route according to the Wardrop assignment. Lower Webster distributes all traffic to the shorter route until capacity is reached, which in this case is about 2,000 vph; then some redistribution to the longer route also takes place.

Up to 1,073 vph the P_0-policy distributes all traffic to the narrower and shorter route, although at least 40 percent of the green time is given to the other route. This of course causes nonoptimal travel times, as will be seen subsequently.

Beyond 1,073 vph a redistribution to the longer route takes place and as soon as this occurs, the amount of green time for this route also increases (see Figure 4).

From 1,073 to 3,190 vph, traffic uses both routes, following the assignment of green times; at 3,190 vph all traffic is assigned to the wider route. However, not all the green time is shifted to this route until capacity is nearly reached. This is because of the equality condition for the $s \cdot d$ values and so in this range the P_0-policy is inefficient. Above 4,000 vph (the maximum capacity of the network) a small amount of traffic is again assigned to the shorter route to satisfy the Wardrop conditions.

Summarizing, it is seen that although the P_0-policy does not behave efficiently at all flow levels, at least the structure of green time and flow changes is correct. At low levels all traffic

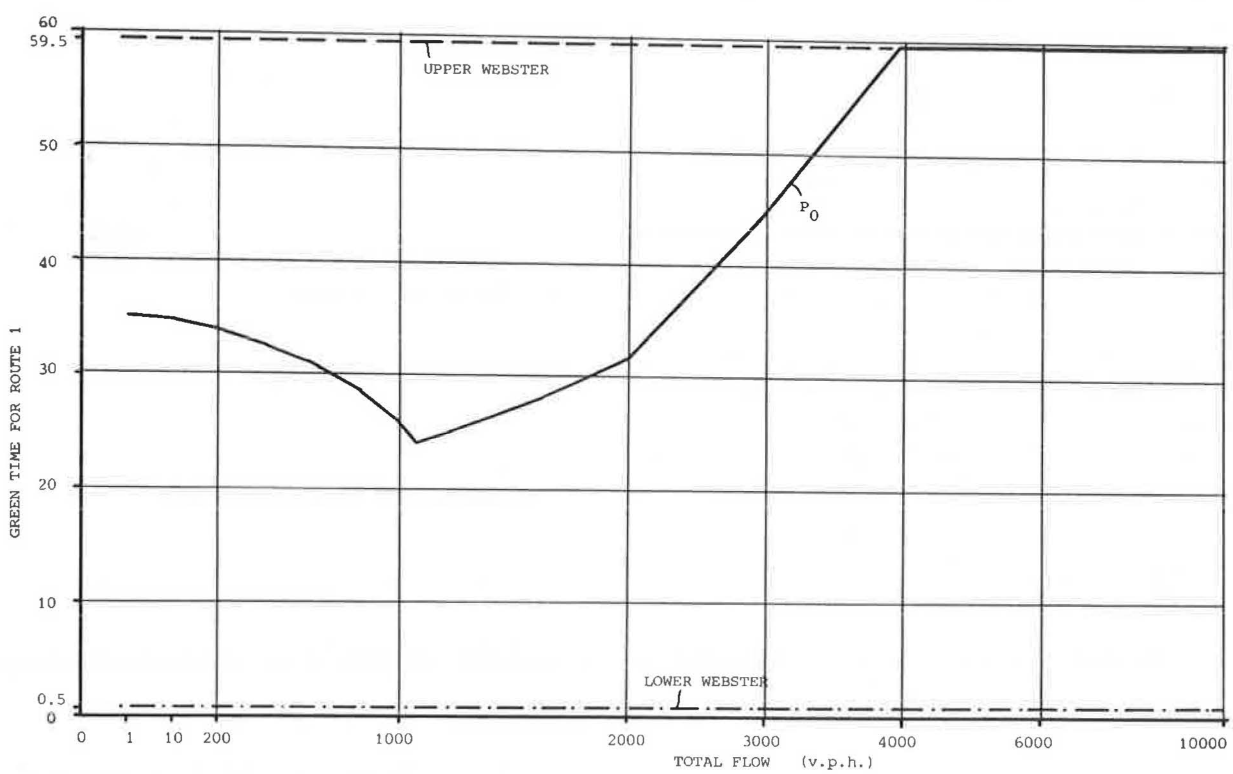

FIGURE 4 Green times at equilibrium.

is assigned to the shorter route, but as flow levels increase toward capacity, a redistribution to the wider route takes place together with a corresponding shift of green time. This is exactly the behavior one would expect from a sound signal-setting policy.

Average Delay and Average Travel Times at Equilibrium

The ultimate test for the performance of the policies is by comparison of their influence on delays experienced and total travel times (the sum of link travel times l_i and delays d_i). It can

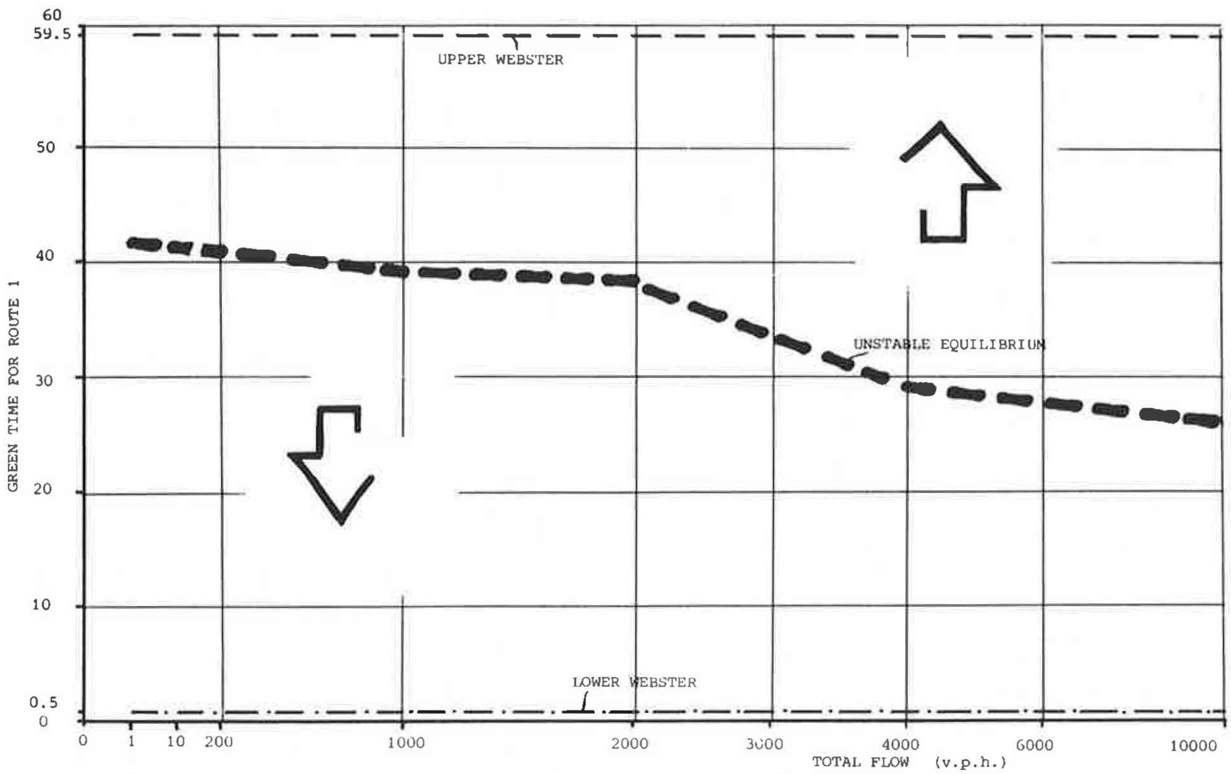

FIGURE 5 Stable and unstable equilibria for the Webster policy.

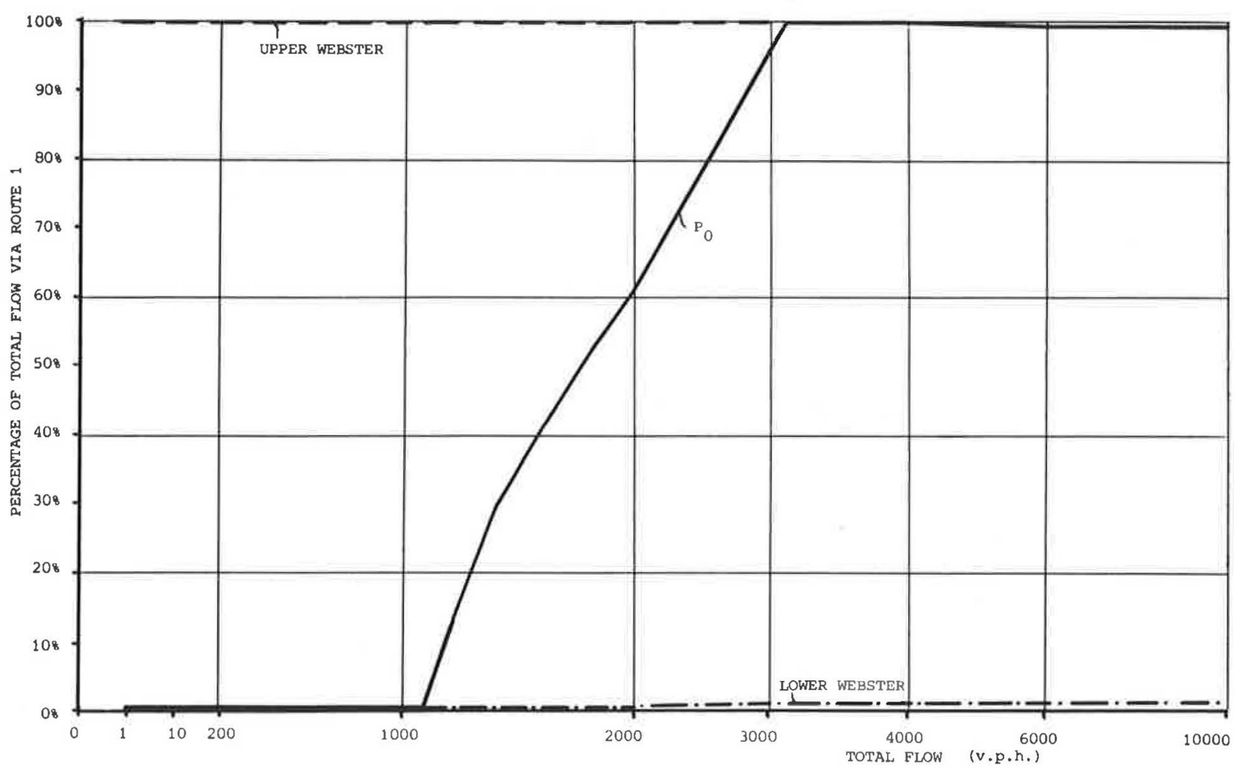

FIGURE 6 Assigned flows at equilibrium.

be seen that the Lower Webster settings give both minimum delays and minimum travel times up to near capacity of the short route (2,000 vph). However, above this (when the flow-to-capacity ratio exceeds 1), queueing increases delays substantially. From a flow of

$$f = \lambda_{max} \cdot s = (59.5/60) \times 2{,}000 = 1{,}983 \text{ vph}$$

The addition of one extra vehicle per hour will cause an increase in delays of

$$0.5T/f = 0.5 \times 1{,}800/1{,}983 = 0.45 \text{ sec}$$

Extra average delay equals half the observed time period (which is the average time a vehicle has to wait if queueing) divided by the total flow. So this lower branch of Webster's policy becomes rapidly worse than either the upper branch or P_0. The high-initiated Upper Webster setting gives all traffic to the wider and longer route and causes minimal delays up to $f =$ 3,967 vph. Delays then also increase rapidly but only at about 0.23 sec per extra vehicle, which is half the rate calculated earlier for the lower branch. Because the wide route is 10.8 sec longer, average travel times will be l_2 + 10.8 sec higher than average delays, as shown in Figures 7 and 8.

Finally, the P_0-policy again shows the most interesting graph. It is no surprise that break points appear at the same places as they do in Figures 4 and 6. Average travel time and average delay are increasing (via the same curve) up to 1,073 vph. Up to about 200 vph average travel times are lower than for the Upper Webster settings because all traffic is assigned to the shorter route (l_2). Above 200 vph delays for the P_0-policy (induced by the "unfavorable" signal settings) are higher than 10.8 sec (which is the extra travel time via the longer route) so that average travel times for the P_0-policy are higher than those for the Webster policy.

At 1,073 vph a redistribution of traffic over both routes takes place, thus decreasing average delays, but because of a redistribution to the longer route, average travel times keep increasing. The gap between delays and travel times keeps widening until at 3,190 vph all traffic is assigned to the longer route (l_1), and the gap is l_2 + 10.8 sec.

To describe the behavior of the P_0-policy at varying flow levels, it can be said that at low flow levels (below 2,000 vph) the policy does not behave efficiently because of a nonoptimal combination of flows and green times. However, the differences with the other policy are limited to some 10 to 20 sec. Above about 2,000 vph the policy performs better than the Lower Webster settings, although still average travel times are some 10 sec higher than those for the Upper Webster settings.

Above capacity (about 4,000 vph) delays increase rapidly. At this stage both the Upper Webster and the P_0-policy perform alike and optimally.

DISCUSSION AND CONCLUSIONS

Conclusions that can be deduced from the test runs on this simple network are as follows:

1. The P_0-policy indeed gives a unique and stable equilibrium for the combined signal-setting optimization and reassignment process.
2. The Webster policy has more than one equilibrium solution; final signal settings and the corresponding flows and delays depend strongly on initial settings.
3. The P_0-policy performs tolerably well with respect to delays and travel times at low flow levels. With increasing flow

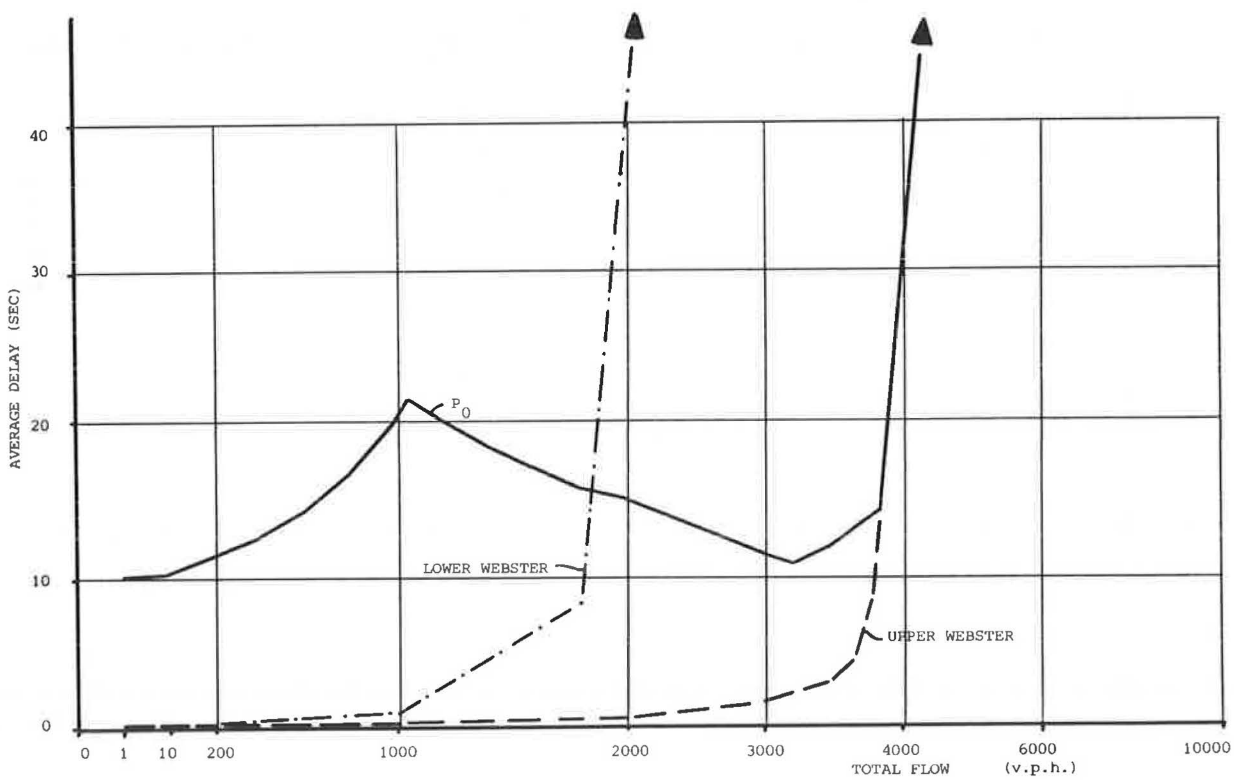

FIGURE 7 Average delays at equilibrium.

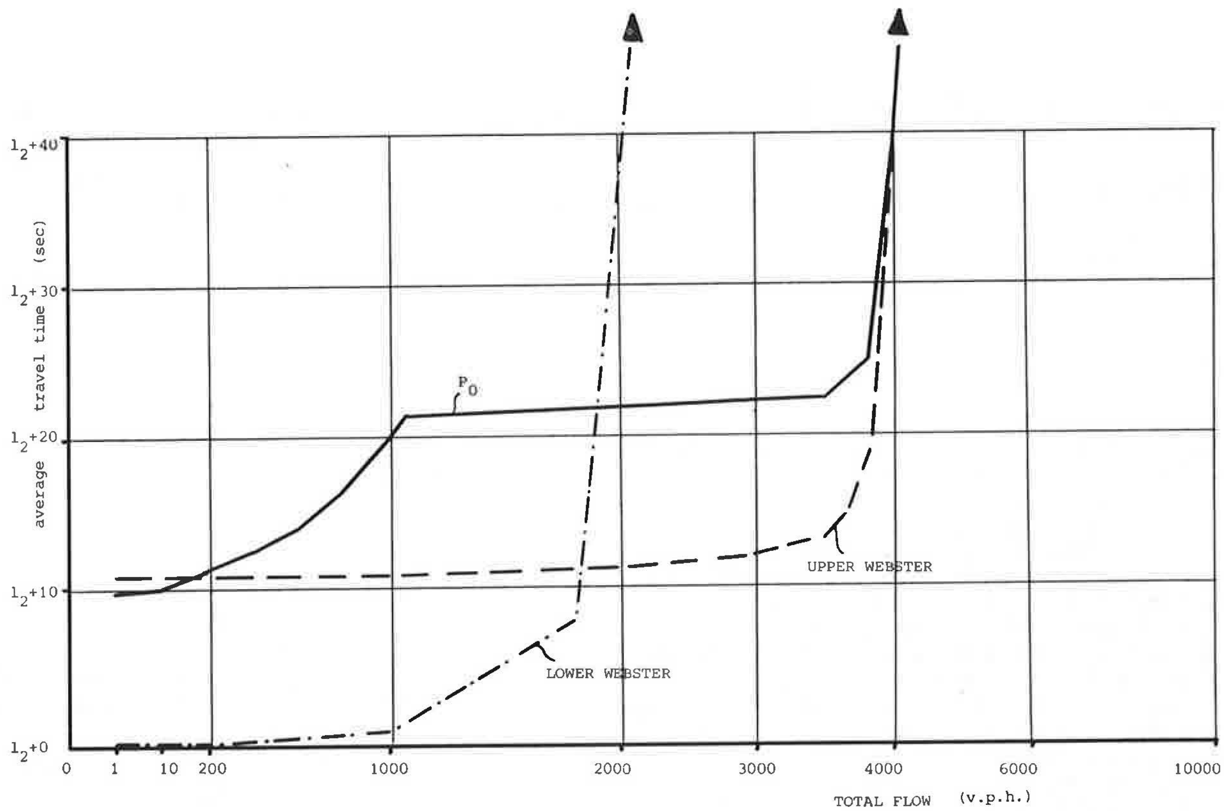

FIGURE 8 Average travel times at equilibrium.

levels, the policy performs better. The policy performs very well, especially above capacity, a confirmation of expectations.

The less-than-efficient performance of the policy at low flow levels is not that disastrous, because delays are small then. Good performance at high flow levels is more important, together with a unique and stable equilibrium. The multiple equilibria that arise with the Webster policy mean that unfavorable initial settings can give very poor results.

The promising results for this simple network may not appear in general. Further tests on larger and more complex networks, to show all the characteristics of the P_0-policy, are being carried out. Some of the first results of these tests were detailed by Smith et al. *(11)*. They appear to show that also on larger networks P_0 performs favorably in comparison with more traditional policies at higher congestion levels. More information can be obtained from the authors.

ACKNOWLEDGMENT

This research project was financed by the Science and Engineering Research Council of Great Britain.

REFERENCES

1. F. V. Webster. *Traffic Signal Settings*. Technical Paper 3. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1958.
2. A. J. Miller. Settings for Fixed-Cycle Traffic Signals. *Operations Research Quarterly*, Vol. 14, No. 4, 1963, pp. 373–386.
3. R. E. Allsop. Delay Minimizing Settings for Fixed Time Traffic Signals at a Single Road Junction. *Journal of the Institute of Mathematics and Its Applications*, Vol. 8, 1971, pp. 164–185.
4. M. J. Smith. Traffic Control and Route Choice: A Simple Example. *Transportation Research*, Vol. 13B, No. 4, 1979, pp. 289–294.
5. R. E. Allsop and J. A. Charlesworth. Traffic in a Signal-Controlled Road Network: An Example of Different Signal Timings Inducing Different Routing. *Traffic Engineering and Control*, Vol. 18, No. 5, 1977, pp. 262–264.
6. D. Braess. Über ein Paradoxen aus der Verkehrsplanung. *Unternehmungsforschung*, Vol. 12, 1968, pp. 258–268.
7. M. J. Smith. A Local Control Policy Which Automatically Maximises the Overall Travel Capacity of an Urban Road Network. *Traffic Engineering and Control*, Vol. 21, No. 6, 1980, pp. 289–302.
8. J. J. Smith. Properties of a Traffic Control Policy Which Ensure the Existence of a Traffic Equilibrium Consistent with the Policy. *Transportation Research*, Vol. 15B, No. 6, 1981, pp. 453–462.
9. D. Van Vliet. SATURN—A Modern Assignment Model. *Traffic Engineering and Control*, Vol. 23, No. 12, 1982, pp. 578–581.
10. A Note on Traffic Assignment and Signal Timings in a Signal-Controlled Road Network. *Transportation Research*, Vol. 15B, No. 4, 1981, pp. 267–271.
11. M. J. Smith, T. van Vuren, B. G. Heydecker, and D. Van Vliet. The Interaction Between Signal Control Policies and Route Choice. Presented at 10th International Symposium on Transportation and Traffic Theory, Boston, Massachusetts, July 1987.

Publication of this paper sponsored by Committee on Traffic Signal Systems.