An View of Converting Mixed Flow Lanes to HOV Lanes

David L Mootchnik

1.0 Background

The idea of building High Occupancy Vehicle (HOV) Lanes to relieve traffic congestion has been popular with metropolitan planning agencies for some time. Indeed in Southern California, Caltrans and the various county and regional transportation agencies have been pursuing a vigorous program of adding HOV lanes to existing freeways. In 1994 the Six county region had over 300 miles of HOV lanes and according to the Regional Transportation Plan the planners intend to extend that to 1353 miles of HOV/HOT lanes. This amount would be about 15% of the total freeway lanes (Reference 1). Orange County has been particularly aggressive in this effort and has built carpool lanes on all but one of its freeways. In a survey conducted by the Southern California Association of Governments (SCAG), 79% of Orange County freeway riders indicated that there was an HOV lane on their route to work (Reference 2). Even with this growth in HOV lanes these references indicate a downward trend in carpooling and the lowest percent of carpoolers is in Orange County.

To date HOV projects have generally added lanes rather than converted existing lanes to HOV. The lanes in general are added using existing median and other right of way space for most of the buildout with only limited requirements for additional land. Thus these projects have avoided the major land acquisition issues as were encountered for the I-105 Century Freeway. For the time being there are still freeways with underutilized existing right of way such that these projects require only limited and isolated land additions. These are the easy solutions. Adding a new HOV lane to an existing freeway does reduce congestion if for no other reason that it adds freeway capacity with the added lane. This being so, the public seems satisfied with these improvements.

The issue of whether it is more effective to add an HOV lane or to add a mixed flow lane has been argued for some time and to date the pro HOV position usually persists. Indeed the analysis of HOV versus MF effectiveness is deceptively subtle and complex. The traffic flow on an HOV, easily observable by the public, is actually not a good indicator of the effectiveness or ineffectiveness of an HOV. There have been many papers on the subject, the majority being subjective in nature, with a few being quantitative but using questionable assumptions or methodology. Others have provided reasonable and useful results and but have been ignored for mostly political reasons.

As long as planners can select projects that add new lanes, there is progress in meeting the increasing traffic demands placed on the system by an expanding population. However as time continues the easy solutions, using existing right of way land, will be consumed and lane additions, like new roadways will require heavier doses of new land. At some point adding additional lanes either means going up or out. These are not the easy solutions. Going up is expensive, visually and acoustically disturbing, and raises earthquake safety issues. Going out requires new land acquisition and significant environmental impacts. New road construction, such as the tollways being added in

Orange County are based on routes through under populated land. The route of a new road is a planning choice. Not so with expansion of existing freeways whose routes are already established. Moreover, land around existing urban freeways is already populated. As these issues become more intractable alternatives to new right of way acquisition will likely be conceived and debated. One of these alternatives is to convert existing MF lanes into HOV or HOT lanes.

Unlike the case of adding a new HOV lane, conversion of an existing MF to an HOV lane does not add capacity to the system. In fact, since the lane is attractive to converting travelers to carpoolers only if it stays relatively uncongested, HOV conversion only works by reducing overall capacity of the roadway. It is argued of course that by influencing travelers into higher vehicle occupancy, the vehicle demand will be reduced, with benefits to all parties. This situation represents a complex balance between forces and the outcome; whether congestion will be reduced, is not at all evident.

2) Objective

It is the purpose of this paper to address the utility of converting a MF lane to an HOV lane. An analytic model will be used to study the performance of the lane options to obtain quantitative results. A key issue is the proper selection of performance measures. To often measures are selected that give only partial information or are biased toward some conclusion. This author wishes to avoid such a situation. We desire to select those measures that are quantifiable and equitable such that the reader will accept the consequent conclusions.

One study that resulted in useful and creditable results is that reported in Reference 3, written by Dr. Joy Dahlgren.

2) Methodology

There are few analytic tools available to properly address the issue of HOV effectiveness. Minimal requirements of such a model are that it include a) the ability to model the section of roadway under consideration, b) a model of the traffic demand over some time period, c) a model of the traffic capacity, speed and flow and, d) a model of the mode choices available. In addition an extensive model would also include a) a model of the region surrounding the subject roadway including a matrix of trip origin and destination and b) a route network allowing for optimal trip routing.

In general the large computer models such as SCAG's Regional Transportation Model (reference 4 and 5) and OCTA's OCTAM incorporate mode choice models and methodology capable of providing valid and useful results. These models include all the features described above. Their limitations are in their need for massive data input requirements, long computation times, and need for iteration on a massive scale. Long computation times can severely limit the number of cases investigated and can limit the number of iterations used. Limited iterations can effect the accuracy of the results especially when looking for localized or small changes. In any event this author does not have access to these models or the time to employ them. Another choice was made.

A more tractable model would limit its scope to the roadway in question or perhaps include nearby parallel alternative routes. All models require a traffic flow/ congestion relationship. Two types are available; one based on a flow-speed relationship such as given by the Bureau of Public Roads (BPR) formula, or one based on a cueing model. Dr. Dahlgren in Reference 3 developed a roadway model using a bottleneck and cue representation of the traffic flow and congestion. The Dahlgren cueing model is used in this analysis. The author has programmed the equations given in Reference 3 and provides enough detail of the model here for understanding by the reader. For further details see the reference.

The BPR formula, while used in many models, is recognized as not accurately representing the dynamics of traffic flow. In particular, the BPR equation indicates that traffic flow (vehicles per hour past a screenline) can continue to increase without end accompanied by ever decreasing speed. Actual observations indicate a saturation point of maximum capacity, beyond which speed decreases with reduced flow.

In a cueing model, this issue is bypassed and the traffic is modeled as a) freeflow if the demand is less than capacity and b) a simple first in first out (FIFO) cue, building up in time when the demand is greater than capacity. The cue size (in total vehicles) is;

$$cue = \int (D - C)dt$$

where;
D = traffic demand (vehicles per hour)
C = roadway capacity (vehicles per hour)
t = time.

The time delay (in hours) of a vehicle caught in the cue is given by;

$$delay = cue / C.$$

This formulation cleverly bypasses the issue of the accuracy of the flow equation and provides a valid means of investigating traffic congestion issues. In the real world the flow equation is fully represented by a first order non-linear differential equation in time and position. In the real world, traffic flow has features of both the BPR flow formula and the cueing model.

The cueing model has the added benefit of being able to represent a time varying congestion situation as is the case during peak hour traffic. For our analysis, following the reference, the traffic demand is modeled as indicated in Figures 1 and 2.

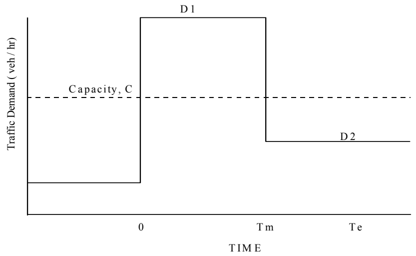


Figure 1 Traffic demand profile

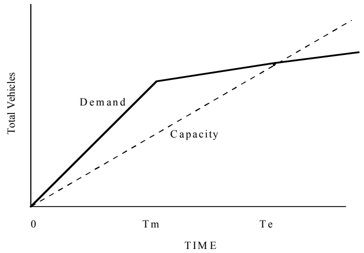


Figure 2 Total vehicles in the system

Before the congestion period, traffic flow demand is at some low level below capacity and traffic is in free flow with no buildup in the cue. Congestion starts at To with an increase in demand rate to D1 > C vehicles per hour and continues to time Tm. At Tm the cue has built up to a maximum and the traffic demand is reduced to a value of D2 < C. Since the cue is not empty, congestion continues until it does empty at Te. Thereafter the cue stays empty and delay is zero. The congestion period is seen to last beyond the time that the high demand rate terminates in order to empty the cue built up along the roadway. This is a real phenomena but is not easily seen on freeway traffic flow charts as these charts show flow at a point on the freeway and not either incoming demand or cue. The cue could be measured by observing flow along many points along the road.

Various relationships are easily derived between D1, D2, C, Tm, Te and max delay. For now suffice it to note that the full situation can be specified by specifying C and any three other parameters. Reference 3 selected Te, Tm/Te and max delay as the independent parameters.

Additional variables needed to specify the problem are the parameters associated with the mode choice model. In this study only two modes are considered; drive alone or non-HOV and HOV. Based on the standard Logit model approach, tow parameters are required;

- The proportion of HOVs in the absence of an HOV lane (HOVo) ,
- The time sensitivity constant (BETA) governing the sensitivity of the choice to the travel time difference between HOV and MF traffic.

Additionally, the average vehicle accupancy (AVO) for an MF and an HOV are needed to be specified.

3) The No-Gain Lane Shuffle

In order to aid the understanding of the results and the utility of HOV lanes it is helpful to understand several basic conditions. There are many incentives that contribute to the formation of carpools. These incentives include, lack of adequate vehicles, cost sharing, alternative to public transit, improved parking priviledges, non driving passengers (such as children), convenience, and companionship. There are also inconvenience factors that limit the formation of carpools such as, wait and formation times, need for travel flexibility, lack of rideshare travelers with same route and schedule, and aversion to HOV lane driving restrictions. The one incentive that HOV lanes provide is the potential for shorter trip times. Even without an HOV lane, there are and will be carpools on the roadway. These carpoolers use the MF lanes as do others. When an HOV lane is added or converted there is normally an in place contingent of carpoolers that will move over to the HOV lane. We label these carpoolers as the initial HOVs (HOVo). The fact that these HOVs move into the HOV lane providing it with instantaneous occupancy with multi passenger vehicles is often sited by proponents as demonstrating HOV effectiveness when in fact no actual benefit to congestion occurred. This is the No-Gain Lane Shuffle. HOVs can benefit the roadway only if they incentivize enough additional travelers to carpool.

In the case of lane conversion, there is a balance point at which it does not matter if there is or is not a conversion. If the ratio of HOVo vehicles matches the ratio of HOV to MF lanes then the flow conditions are the same in the MF and HOV lanes. Travel time, speed and other congestion measures will be equal and there will be no incentive for carpool to form or breakup. This is a balanced steady state condition and no impact of the HOV conversion, good or bad, should be expected. For three and four lane highways, the Balanced No-Gain Lane Shuffle (BNGLS) occurs (assuming an HOV AVO of 2.3) as follows.

LANES	Percent HOV vehicles	Percent HOV people
3	33	53
4	25	43

The BNGLS point is the upper bound of interest for this analysis. From the practical standpoint the levels of carpooling at this limit is well beyond what is actually found in the most HOV area, Especially around Southern California. Note that at any level, the observation that vehicles are using the HOV lane is no proof that HOV lanes have any benefit in reducing congestion.

4) Examination of a Base Case

4.1 Case Description

To begin the investigation we will pick the example case shown in Figure 5.1 of Reference 3. This case assumed the following parameters;

Te = 3 hours
C = 2000 vehicles pr hour per lane
Tm/Te = 0.5
AVO MF = 1.0 people per vehicle
AVO HOV = 2.3 people per vehicle
Total number of lanes = 3

HOVo = variable
BETA = -0.04 per minute 2 way

These parameters are fairly typical for a three lane freeway. Congestion periods usually are considered to vary between 3 and 4 hours long however as congestion worsens this period continues to grow. The AVO for HOVs is typically 2.15 to 2.3. With bussing using the road the value would be higher. The AVO for MF assumes all drive alone. We know from experience that some not insignificant fraction of carpools use the MF lanes. They are not influenced by the time advantage of the HOV lanes. The AVO for the MF is actually higher than 1.0. In the SCAG model documented some time ago the value of BETA was –0.01. The time sensitivity constant derived for various studies as discussed in the reference 1 is in the range of –0.01 to –0.05. The value use here is on the high side and will favor creating carpools. Finally the value of C is normally in the range of 1900 to 2300.

The measure of performance used in the reference was the average delay time. Delay time was averaged over all travelers entering the roadway during the hours of congestion. This was an averaging over time and weighting the MF and HOV vehicles by occupancy. The time slice evaluated was defined by Te.

This definition seems most reasonable and equitable. If one multiplied the result by the total number of travelers entering the system, one would get total hours of delay; a measurement commonly used in congestion studies. We utilize this measure (average delay time, Dtave) for the moment, but will find issue with it shortly.

4.2 Base Case Results

Figure 3 summarizes the initial results of the average delay analysis of the baseline case for values of HOVo below the BPNS noted above. In all these cases the peak delay is held constant. In this set, as HOVo varies, the total initial traffic demand in vehicles is held constant. (For a given HOVo, the vehicle demand will vary timewise as the mode split changes dynamically). Also note that the number of incoming people varies as HOVo varies such that conditions are not exactly the same for the different HOVo cases.

These data match the results as reported in Reference 3 to within reasonable integration accuracies. The no-change option represents the situation with three MF lanes. It can be easily shown that for the conditions of constant demand rates, D1 and D2 that the average delay is simply half the peak delay. Thus for this option the average delay for the 3 hour peak period stays constant at 10 minutes.

Figure 3 Peak Period Average Delay Comparison Baseline Case

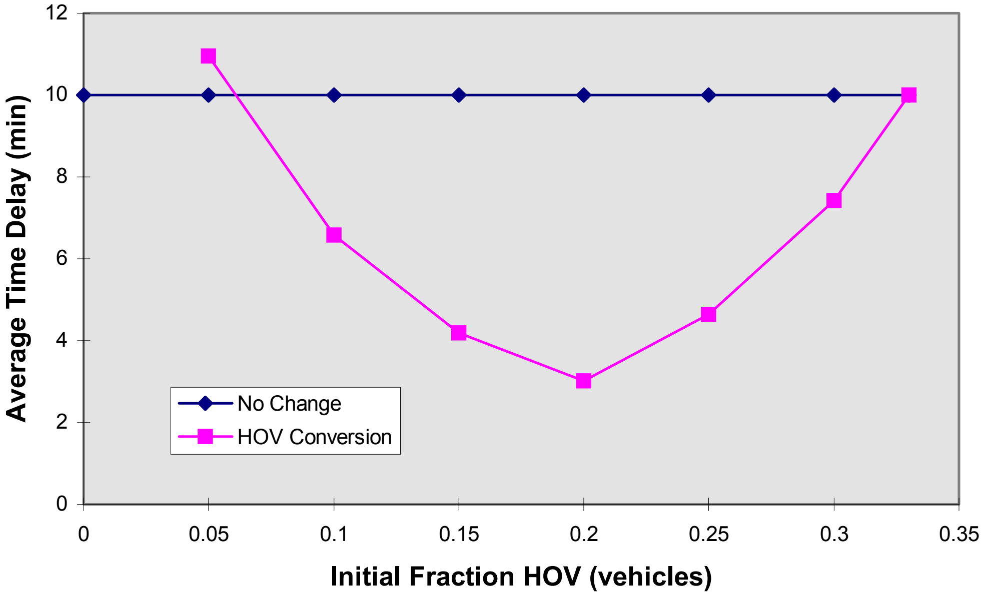


Figure 3 indicates that for levels of initial HOV between about 0.06 and 0.33 there appears to be a significant benefit to converting an MF lane to an HOV lane. To better understand this effect, the time history of delay in both the MF and HOV lanes, the traffic demand curves and the dynamic mode split were determined. The results for a typical case with HOVo=0.1 are shown Figures 4a,b,and c.

FIGURE 4a Instantaneous Delay Time History

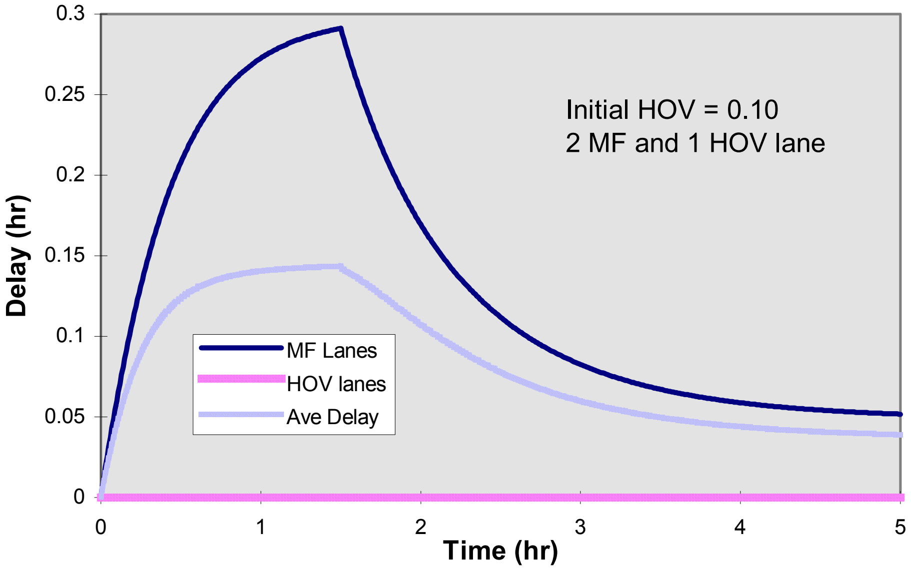
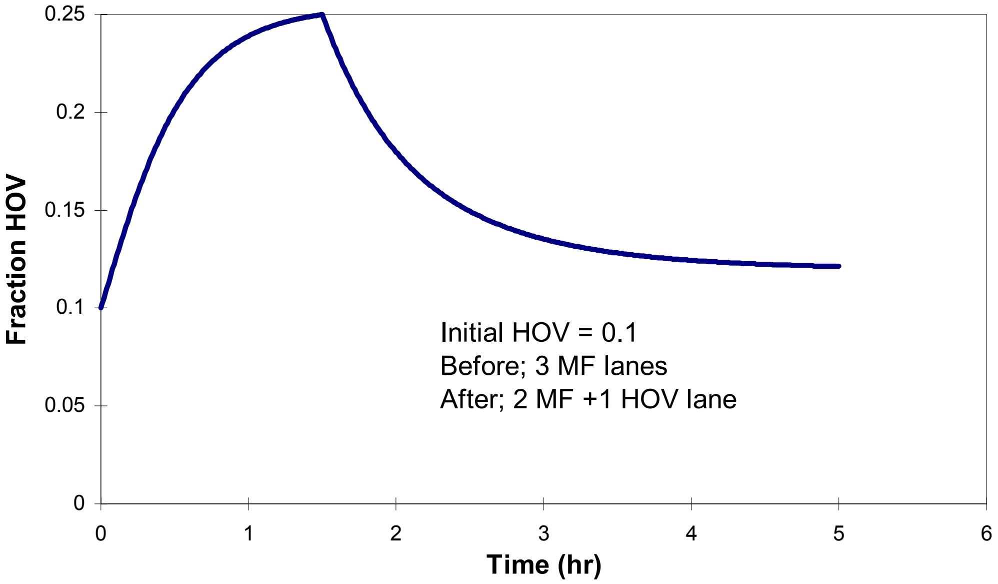


Figure 4a shows the separate vehicle delays for the MF and HOV lanes as a function of the time that the vehicle enters the system. Several points are to be noted. First that the delay in the HOV lanes equals zero, meaning that the traffic in the HOV lane never reaches the lane capacity of 2000 vehicles per hour. For the conditions of the baseline case the total incoming traffic is 7320 vehicles per hour up to Tm and drops to 4680 after. With the initial HOV fraction of 0.1, the traffic demand on the HOV is only 732 vehicles per hour. However as the cue and delay build up in the MF lanes, one would expect that the HOV fraction would build up. This is seen to occur in Figure 4b. The fraction HOV builds up from 0.1 to about .25 at the peak delay time. Also it is seen that the peak delay for the MF lanes builds up to almost 0.3 hours and the average delay for the 3 hour period is 10.85 minutes, greater than the average delay before the conversion.

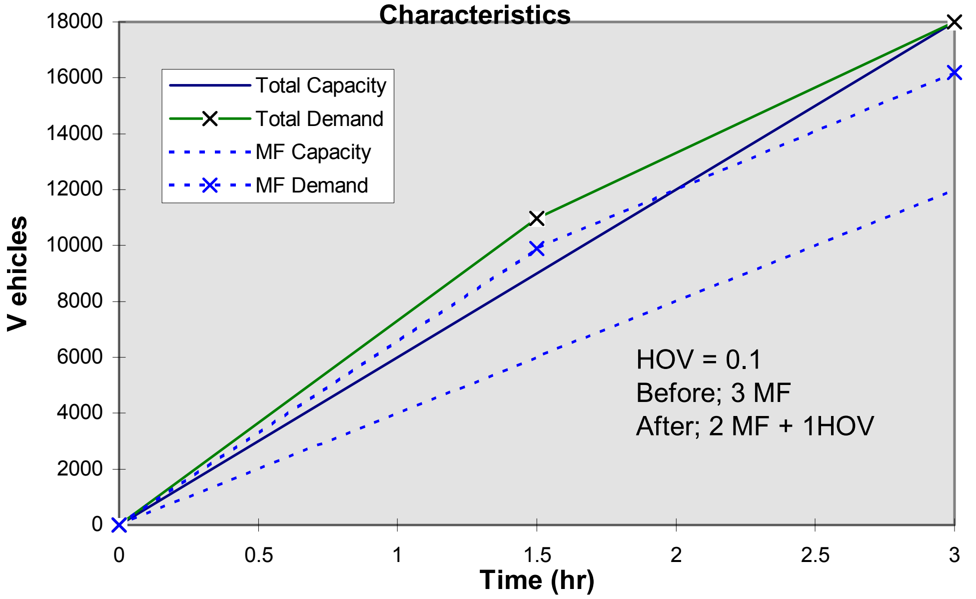
FIGURE 4b Mode Split History



The small increase in average delay for the MF lanes may seem a small price to pay to let those willing to carpool to have no delay. However the price is actually much larger.

The second feature of Figure 4a to notice is that, unlike the No Change case, there is still congestion remaining in the MF lanes after the 3 hour Peak Period is over. The reason for this is seen in Figure 4c which shows the demand and capacity curves for the before condition (No change 3 MF lanes) and the MF lanes in the after condition (2 MF+ 1 HOV). In this figure the after condition is shown when the HOV fraction stays at 0.10. As seen, the demand lines for the 3 MF situation build up to a peak at Tm=1.5 hours and then converge back to the capacity line at Te=3.0 as expected. Thus in this case, congestion will end at 3.0 hours and the average delay is half the peak delay.

Figure 4c Baseline Case Capacity and Demand Characteristics



For the after condition, the capacity in the MF lanes are reduced by 1/3 while, with HOVo = 0.1, the initial demand in the MF is reduced only by 10%. The situation is so bad that in the second phase of the demand curve after Tm (D2) the demand is still higher than the MF capacity. Thus there is never any end to the peak congestion period. What is seen to happen in Figures 4a and 4b, is that the level of HOV adjusts to above the HOVo level because of a continuing level of congestion and congestion asymptotes to a delay of about 0.03 hours (2 minutes). In this condition, since the mode split adjustment needs some level of congestion on the MF, the delay on the MF self regulates to a continuing non zero level.

There are in fact three possible situations regarding the extent of the congestion period

1) The congestion period never ends. This occurs when the value of D2 on the MF lanes is larger than the capacity of the MF lanes after conversion. The carpooling incentive will be to weak to reduce it below the MF capacity

2) The congestion period ends but after the original situation. This occurs when the value of D2 on the MF lanes is slightly less than the capacity of the MF lanes after conversion. This will occur for the Base Case conditions when HOVo is between about 0.15 and0.20 vehicles.

3) The congestion period ends but before the original situation. This occurs when the value of D2 on the MF lanes is less than the capacity of the MF lanes after conversion by a more significant amount. This will occur for the Base Case conditions when HOVo is 0.2 vehicles or greater

4) The congestion period ends at the same time as the original situation. This occurs when the BNGLS is reached.

Figure 5 demonstrates the pattern of congestion for the second condition at HOVo=0.15. Figures 6 and 7 demonstrate the patterns for condition 3 at HOVo= 0.20 and 0.25. The pattern in Figure5 is much like Figure 4a except that the delay time for both lanes ends at about 4.25 hours. The period of congestion was extended in this case by about 1.25 hours above the no change situation.

Figure 5 Time Delay History for Condition 2

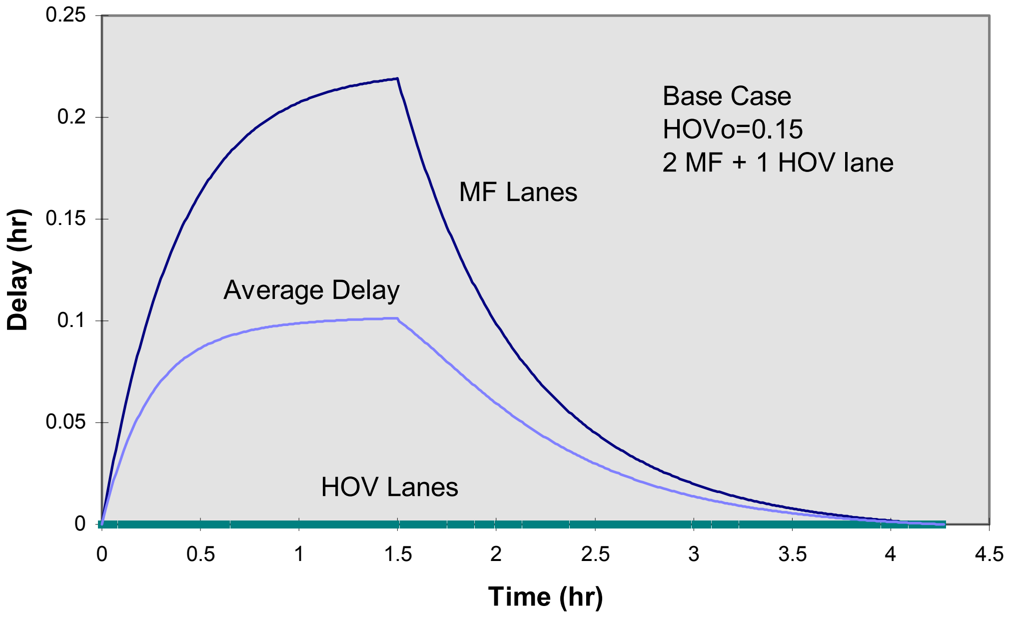
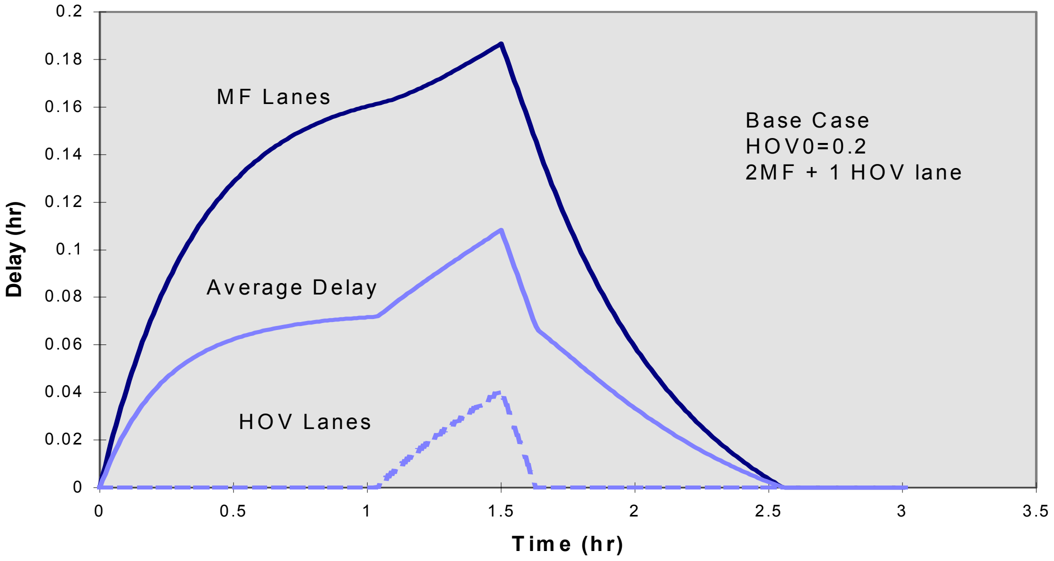


Figure 6 Time Delay History for Condition 3



At HOVo equal to 0.2 it is seen that the HOV lane is beginning to get congested and by HOVo=0.25, the HOV congestion is beginning to approach the MF lanes. At higher HOVo there is even less difference between the two, but enough to maintain an incentive for significant HOV choice. Figure 8 shows the mode split history for the HOVo=0.2 case. As seen the fraction of HOV vehicles is quite high and almost reaches the balance point of 33% for some period of time. At HOVo= 0.25 and greater, the mode split increases to nearly 33% early and stays there for over 2 hours.

Figure 7 Time Delay History for Condition 3

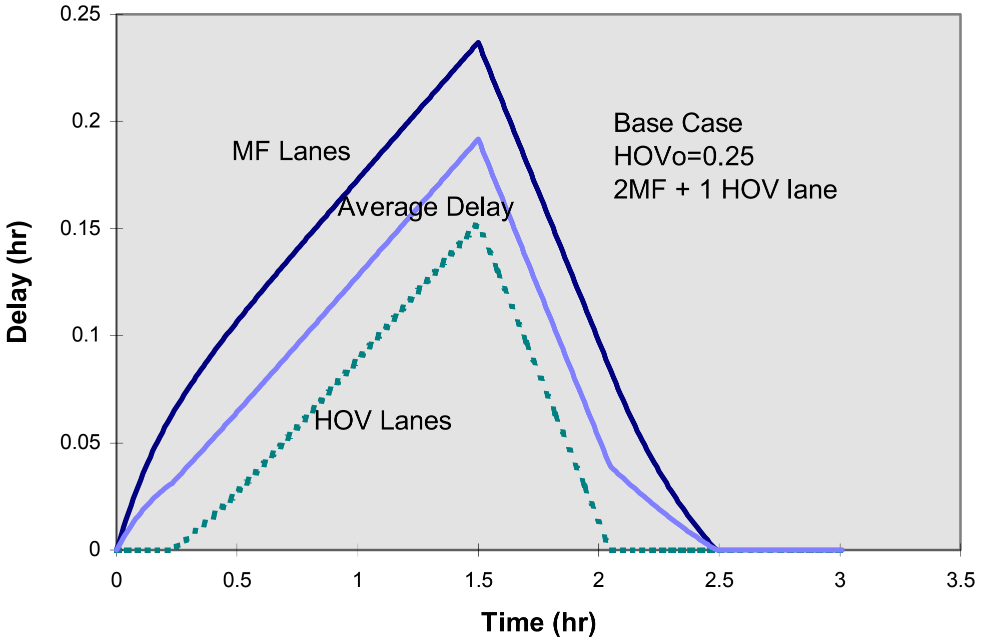
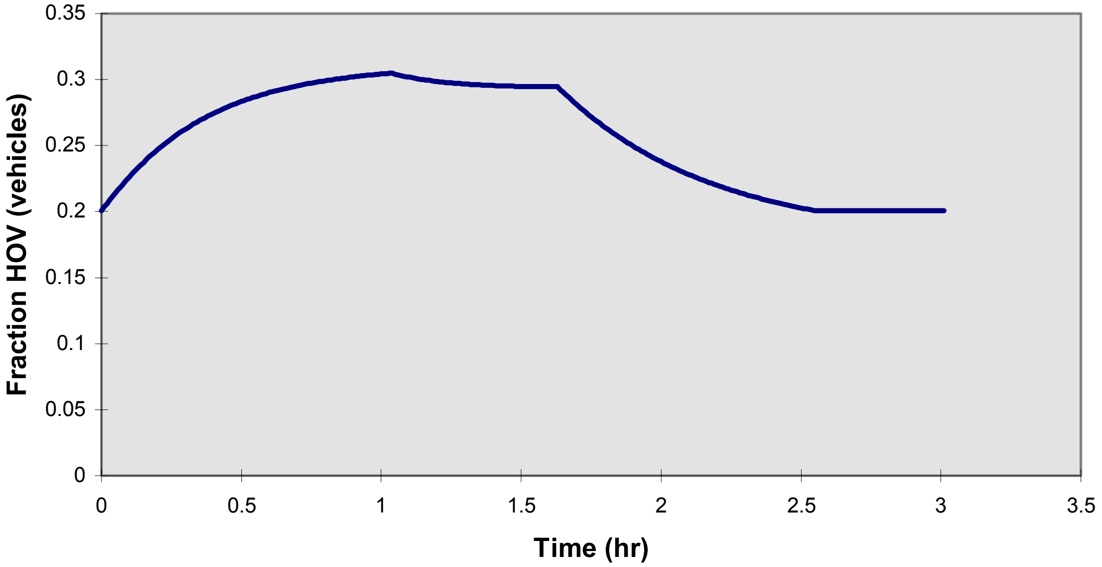


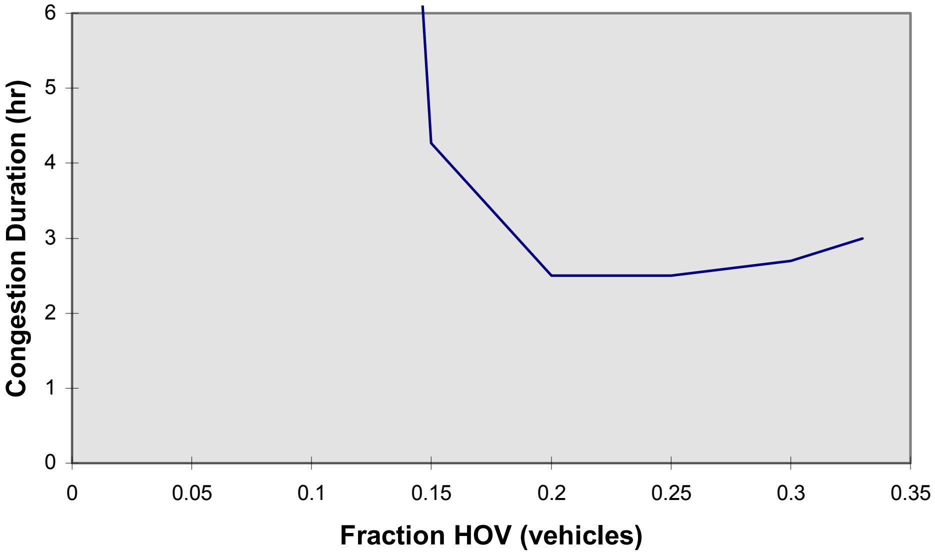
Figure 9 summarizes the trend of congestion time after a conversion to HOV as a

Figure 8 Mode Split History
HOVo=0.2



function of the initial HOV fraction. For initial fractions less than 0.14 the congestion period extends indefinitely but rapidly decreases to less than the original Te value of 3.0 hours.

Figure 9 Congestion Period Trend

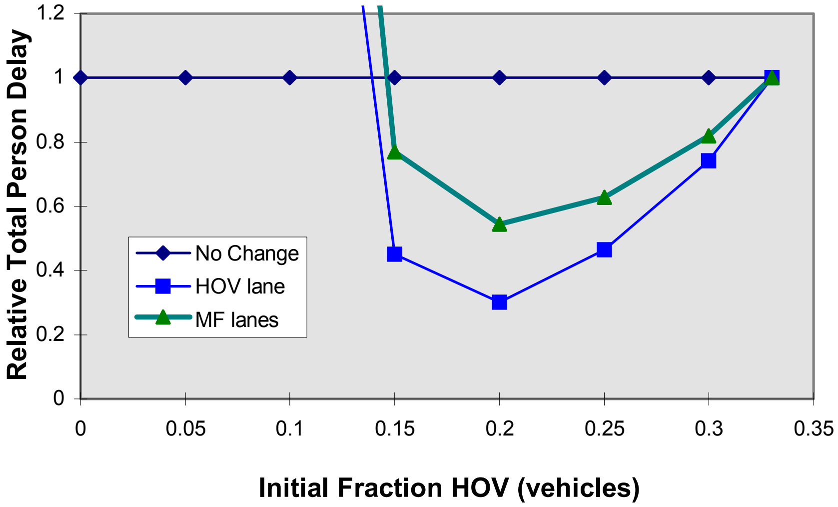


In the situation in which congestion extends beyond the original peak period, the average delay in this period is no longer a proper measure of performance. An alternative measure of performance proposed here is the total person delay (TPD) over the entire extended congestion period. TPD is the sum of delay time for all riders entering the system from T=0 to Te or the end of congestion whichever is larger. This measure reflects the entire impact of the conversion on traffic while continuing to weight MF and HOV according to their proportion. For those cases with congestion time less than the value of Te, the total delay equals the average delay times the total riders entering the system during the peak period. That is the total delay is proportional to average delay in those cases. However as the congestion period increases the total delay will reflect any delay to those riders entering during that period.

Figure 10 shows the TPD for the base case. TPD is plotted in relative terms (TPDrel= TPD/TPDo) compared to the value before conversion. A value of 1.0 indicates no impact of the conversion. When the period of congestion extends indefinitely, as it did for HOVo < 0.14, the value of TPD will be infinity. However for HOVo> 0.5 to 0.3 the TPD is significantly less than the before case.

Figure 10 also shows the relative TPD for the MF lanes. While the total TPD is considered to be the best equitable measure, if the drive alone travelers experienced a drastically poorer situation then, subjectively one would question the veracity of the conversion. However Figure 10 shows a less but still significant improvement for the MF lanes.

Figure 10 Relative Delay for MF and HOV Traffic



4.3 Base Case Conclusions

The Base case was selected as the starting point for analysis based on the work in reference 3. In that report, Dr. Dahlgren reported that, while for the vast majority of cases analysed (91%) it is better to add MF lane, in 74% of the cases analysed it was better to convert a lane from MF to HOV if no additional lanes are contemplated. This author became curious of this apparent contradiction and decided to investigate further. The Base Case is that described in Figure 5-1 of Dahlgrens report. I found that there were hidden factors in some of those results that made some results misleading. One can however indeed conclude that for the base case there are situations where an HOV conversion will significantly benefit the travelers along the roadway. This benefit requires that the initial share of carpoolers be above the critical value of 0.15 and that the travel time elasticity of the mode split equation reflect the ability of up to 53% the riders to convert to carpools. However it is also seen that if the initial share of carpoolers is below the critical point then the result may be a much deteriorated situation in which congestion is an all day and night affair. A very close examination of the roadway situation is therefore warranted before any conversion is attempted. The remainder of this report will examine the other factors and conditions affecting the conclusions.

Because the period of congestion can be extended by the conversion, the average delay during the initial congestion period alone does not reflect the system performance properly. To correct for this it is recommended that the total person delay extended over the entire ensuing congestion period be used. This measure properly includes the effect of extended congestion. While it will not change any results, trends or conclusions, this author used a simpler way of representing this measure is as a ratio, showing it relative to the original total person delay.

5.0 Reasonableness of the Base Case

This section will discuss some of the parameters used in the base case and compare them to local OC conditions.

6.0 Examination of a potential real situation

7.0 Examining sensitivities

8.0 Summary and conclusions

REFERENCES

1) Southern California Association of Governments, Draft 98n Regional Transportation Plan, Appendix Table B-1, November 1997

2) Southern California Association of Governments, State of the Commute Report, 1996

3) Dr. Joy W Dahlgren, An Analysis of the Effectiveness of High Occupancy Vehicle Lanes, Univ. of Berkeley, Dec 1994, UCB-ITS-DS-94-2