THE FIVE CLASSICAL HOV LANE FALLACIES

✷

HOV lanes are widely supposed to be more effective than regular, mixed-flow lanes at reducing congestion and all its adverse consequences. The overwhelming majority of publications concerning HOV lanes find them to be effective. The shear weight of all these affirmations from independent sources seems itself to present an invulnerable foundation. Nevertheless, legitimate doubts remain. This paper takes a new hard look at those foundations and the "weight of evidence".

✷

Jack Mallinckrodt
AJM Engineering
(714) 544-3200
mally@ieee.org
urbantransport.org
May 1, 2002

THE FIVE CLASSICAL HOV LANE FALLACIES

1. Overview

HOV lanes are widely supposed to be more effective than regular, mixed-flow lanes at reducing congestion and all its adverse consequences. The overwhelming majority of publications concerning HOV lanes find them to be effective. The shear weight of all these affirmations from independent sources seems itself to present an invulnerable foundation. Nevertheless, legitimate doubts remain. This paper takes a new hard look at those foundations and the "weight of evidence".

2. The HOVL Hypothesis

HOV lanes ("HOVL") are generally proposed and adopted in the belief that operating a lane of a freeway as an HOV preference lane:

1. *will afford a relatively free-flowing lane to HOVs,*
2. *that will motivate more people to carpool,*
3. *that will take some cars off the road*[1]*, and*
4. *that will reduce overall congestion, travel time, emissions, and energy consumption relative to the all- mixed-flow operational alternative.*

This may be called the "HOV lane Hypothesis". It's not unreasonable. The premises are plausible and the bottom line goals are certainly significant and fundamental. We take verification of this hypothesis to be the sufficient and *necessary* condition for proof of HOVL effectiveness. Any HOVL finding is relevant to the effectiveness issue only to the extent that it affirms or helps affirm this underlying hypothesis.

But the hypothesis deals with only half the HOVL story – the good half. Due to HOV lane induced traffic diversions in time and space, the choice to operate the lane in the HOV mode rather than Mixed Flow has impacts generally extending well beyond the lane itself, and even beyond the freeway itself to the surrounding surface streets and corridor. Some of those impacts are beneficial and some are adverse. All these impacts must be taken into account in a valid overall net effectiveness assessment. Modern comprehensive regional traffic computer models are capable of doing so. But this is a complex problem, quite beyond anyone's casual intuitive ability.

A careful reading of the hypothesis will reveal that in order to confirm it, a valid effectiveness finding, whether experimental, analytic, or a hybrid of both, must satisfy *all* of the following minimal HOVL effectiveness finding *criteria:*

1. compare the *operational* alternatives of HOV vs Mixed-flow operation of a

[1] This implicitly assumes that the travel demand is to move people, not cars, an assumption with which there should be no significant disagreement.

given designated lane of the same freeway, under the same external conditions.
2. make an *overall* assessment, taking into account aggregate effects on the entire impact corridor, including all lanes, freeways, persons and vehicles affected by the HOV/MF choice.
3. take into account all *negative as well as positive* impacts for an overall *net* benefit assessment.
4. find that the *overall, net* impact on congestion, travel time, emissions, and/or energy consumption is more favorable under the HOVL operational alternative.

But in spite of the many hundreds if not thousands of reports of HOV lane success, that has *never been done*. There has never been a measurement showing an *overall net* benefit in any of these bottom line impacts (congestion, travel-time, emissions, energy or accidents) due to the choice of HOV rather than MF operation of the lane. HOV lane. It may well be impossible to do so experimentally.

3. Direct A-B Comparison Measurements

It may at first seem straightforward to implement such an A-B test of the HOV alternative. Select a Mixed-flow Lane (MFL) and after making complete impact-corridor performance measurements, convert to HOV Lane (HOVL) operation (or vice-versa), re-measure, and compare primary benefits listed above. This has proven to be difficult if not impossible for two reasons.

The first reason is political/perceptual. The first HOV lane in California, the famous Santa Monica Diamond lane experiment in 1976, was in fact an A-B test, implemented by converting a regular unrestricted freeway lane to HOV operation (though with inadequate instrumentation for a complete test). The result was an immediate dramatic increase in congestion in the mixed-flow lanes, (and undoubtedly, though not measured, increase also in emissions). A U.S. DOT contracted report on the event said:

> "Energy savings and air quality improvement were insignificant, freeway accidents increased significantly, non-carpoolers lost far more time than carpoolers gained, and a heated public outcry developed" [1]

After 21 weeks of damaging newspaper revelations and a lawsuit, the lanes were restored to mixed-flow by court order. HOV lane establishment was set back for about ten years in California.

That was the first and last A-B HOV lane test in California. Since that time, Caltrans policy has been to "never establish an HOV lane by *taking away* a lane." Instead, the establishment of a new HOV lane is *always bundled* with the creation of a new lane. The addition of the new lane, of course, can only improve congestion, whether operated as HOV or MF.

For the same reason, officials have opposed making such an A-B test as a temporary experiment saying that 'if we ever converted a lane for such a test we would never get it back'(begging the question why should they want to 'get it back' if it were so obviously inferior). The 1998 Caltrans district 7 HOV report stated

> *"The current district 7 program has an important distinction from the (Santa Monica) "Diamond Lane" experiment of the 1970s: no traffic lanes are being taken away. Rather, the new HOV lanes are being added to existing freeways."*

It is important to note that the important distinction in this policy does not address the issue of actual effectiveness, only the issue of *public perceived* effectiveness. Since Santa Monica, because of this policy, there has never been an opportunity in California to observe or evaluate an A-B test of HOV vs. Mixed-Flow operation of a lane; and the question of which works best remains untested and unanswered by this most obvious and direct measurement.

The second, more fundamental reason this A-B experiment has not been done is the technically valid objection that the measurement may fail to capture the fundamental benefit of an HOV lane, the induced carpools formed on account of the lane motivation. The difficulty arises from the fact that the full induced carpooling benefit of an HOV lane takes a long time, possibly several years to develop, during which time, exogenous effects such as population growth, economic trends, business changes, etc. are almost certain to mask the small amount of additional carpooling that may be motivated by an HOV lane[2]. A short-term comparison will underestimate HOVL advantages by missing the induced carpool effect, while a long-term one will almost certainly be dominated by much larger, inseparable exogenous influences.

4. HOVL "*Findings*"

Therefore, in lieu of direct A-B experimental proof, workers have fallen back on several shortcut measurements we may characterize as HOVL "findings". All of these findings are "consistent with", and *suggest* HOV lane effectiveness, but individually do not unambiguously prove it.

In published HOVL findings, the following five prototypes are estimated to represent at least 99% of cases. They have been repeated and reaffirmed in many hundreds of places and times and have assumed a de facto "classical" status. The shear cumulative weight of these hundreds of reaffirmations appears to have bestowed on

[2] A short term (e.g. within 1 month) A-B (HOV vs MF lane operational alternatives) test would capture *most* of the difference between the A-B states, including all the simple diversions of persons and cars, and the carpool time-saving, missing only the second order (10 percent or so) effects due to benefit carpooling induced by the HOVL time saving. That could probably be estimated to adequate accuracy by adapting the modal choice estimation methodology common to most urban traffic planning models. This appears to be a viable candidate for the best possible HOV lane effectiveness "hybrid" measure-analysis.

them a widely accepted status of “proof” which serves as the foundation of widespread acceptance of HOV facilities nationwide. The reasoning seems to be that, like circumstantial evidence, even though any one instance of such finding may not prove the point, the shear weight of hundreds of independent affirmations add up to “proof”. We will return to this “weight of evidence” issue after discussing the five prototype findings individually.

5. The Five Classical HOVL Findings

1. FINDING TYPE A: "Most carpool lanes carry at least 50% more peak-hour person-trips than the adjacent freeway lanes." [2]

In short this may be called the “more persons in the lane” or “lane-to-lane comparison” finding. Common variants include that the lane carries x% more persons, in y% fewer cars, z% faster than an adjacent MF lane. It is without doubt the most widely reaffirmed HOVL finding, included in almost every HOV lane study.

However, the implication or inference of HOVL effectiveness is erroneous. The implication fallacy may be best illustrated with a simple thought experiment: Suppose we have a long segment of 4-lane mixed-flow freeway with no entrances or exits. It is operating at a moderate congestion level approaching peak flow, at average speed about 45 mph. Traffic is composed of 80% single and 20% two or more occupant vehicles (typical numbers). In the mid-point of this segment of freeway we convert one of the four lanes to HOV2 operation (two or more occupants). At the beginning of the HOV section, traffic redistributes itself so that all single occupant traffic moves out of the HOV lane and all of the multi-occupant vehicles move into it because it now has a smaller vehicle flow and is moving faster. The following statement can easily be confirmed:

> "The HOV lane section is now carrying 50% more people in 25% fewer vehicles than the adjacent mixed-flow lanes.

Yet, in this example, there clearly is no overall benefit from the HOV operation. The freeway is carrying just as many people and no faster. The HOV constraint has just diverted the single- and multi-occupant vehicles to separate lanes. No new carpools have been formed. There has been no reduction of number of vehicles required for the same number of people to get to work. The HOV lane is now carrying

> *more people, in fewer cars, faster*,

but at least equally important, the *other* lanes are now carrying:

fewer people, in more cars, slower

The more people in the HOV lane does not prove the HOVL operation successful any more than the fewer people in the other lanes prove it unsuccessful. Whatever person capacity advantage may appear to accrue to the HOV lane by more people and fewer cars is in this case *exactly* compensated by an equal and opposite disadvantage to the adjacent mixed-flow lanes. There have been no overall benefits and in general there are no benefits on this account. And yet "the HOV lane section is now carrying 50% more people in 25% fewer vehicles". Clearly, the "more people in the lane" finding does not necessarily imply overall improvement.

Ironically, part of the reason the HOV lane is able to carry more people than the other lanes is because the choice of HOVL operation has *caused* the other lanes to carry *fewer* people *slower* than before. In other words, the type A finding is comparing the HOV lane with the mixed-flow lanes, degraded by the HOV lane operation.

In fact *almost all such type A findings are fully explainable, qualitatively and quantitatively with reasonable assumptions as to prior occupant distribution on the basis of traffic redistribution effects alone and without any additional carpools formed on account of the HOV incentive, i.e. without any HOV benefits.*

This fact is most evident in the results of simple HOVL simulations, in which the "more people in the lane" property is almost always satisfied, while the overall traffic impacts are usually adverse.

The one possible exception to the above dismissal may be the several HOVLs that serve also as heavily used busways such as the Shirley Highway, the Houston Katy and Northwest freeways [3], the Lincoln tunnel, and the I-10 in Los Angeles. The very heavy lane occupancy in these cases cannot be explained as simple redistribution effects as above. On the other hand, it also cannot be attributed solely to the HOV lane time-saving incentive since all such instances have also enjoyed significantly increased bus service put in place to take advantage of the lane. The decisive test would be to compare the corridor with bus lane and all its special bus service, against the operational alternative of freeway operated in all MF mode, and the *same amount of provided express bus service* via whatever the best routing. Needless to say this test may never be done, so definitive measures of busway effectiveness may never be obtainable.

The type A argument does not prove that the HOV lanes are effective, nor does the counter-argument here prove that they are not. Ultimately, there may or may not be enough new HOV motivated "benefit" carpools formed to more than make of for the reduced vehicular capacity of the lane, and the adverse effects on the other lanes, but no such effect is indicated by the Type A statement as proven by this example.

To imply or infer that the effectiveness of the freeway is in any way proven or indicated by the "more persons in the lane", TYPE A finding is a serious logical error. Of the above listed minimum criteria for a valid effectiveness proof, it fails to satisfy #2, #3, and #4. Yet, surprisingly, this constitutes by far the most common HOVL finding and the main foundation of belief in HOV lane effectiveness. It is the *only* positive finding offered in the Institute of Transportation Engineers' major report on "The Effectiveness of High-Occupancy Vehicle Facilities" [2]

2. FINDING TYPE B: "Average Vehicle Occupancy (AVO of the freeway as a whole) has increased from 1.17 to 1.26 people per car." [4] or, a variation: **"The percentage of carpools increased by 65% from 15.8 percent to 26.1 percent."** [1]

This again is fully explainable as a redistribution effect. When an HOVL is established the following effects inevitably occur:

Immediately:

- Single-occupant traffic is diverted from the preference lane to the mixed-flow lanes and vice-versa

Soon:

- Some part of the multi-occupant traffic presently using alternative surface streets hears about the HOV lane, tries it, likes it, and migrates to the HOV.

As traffic congestion generally, worsens:

- More and more single occupant users are forced off the increasingly congested mixed-flow lanes of the freeway onto the alternative surface streets.

- More and more multi-occupant vehicles are forced off of the increasingly congested alternative surface streets and attracted to the still nearly freeflow HOVL.

All of these effects will tend to increase the average vehicle occupancy of the HOV lane instantly, of the freeway shortly, and more so over time as general traffic congestion worsens. The latter route diversions may take months or years to fully develop. None of these redistribution effects have any beneficial influence on overall system capacity, in fact slightly the opposite as discussed above. As traffic loads increase generally, the HOV lane itself will generally be, able to accept more multi-occupant traffic while the mixed-flow lanes are already super-saturated, and can only decrease flow under the pressure of increased offered traffic. All of these observed effects are inevitable whether or not any additional new carpools are formed.

As with the TYPE A claim, HOV motivated additional carpooling may or may not occur, but this TYPE B claim in no way proves or measures it; at least not without adjustment for the above effects of selective traffic redistribution, which are certain to dominate TYPE B observations at least initially, and which are of neutral or negative benefit. None of the quoted references attempts such an adjustment. Without such adjustment TYPE B claims are almost totally misleading with respect to HOV effectiveness. They fail to satisfy criteria #2, #3, and #4.

3. FINDING PROTOTYPE C: "Since the opening of the (HOV) lane there has been a 43% increase in the number of vehicles using the freeway ...The reduction in congestion that has occurred in the three existing general use lanes has served to increase the capacity of these lanes as well." [5]

This is a comparison of a 3-lane pre-HOV freeway to a 4-lane post-HOV freeway, (or an n-lane, to an n+1 lane) and hardly needs more comment. The issue is not whether 4 lanes work better than 3. Of course they do, as long as there is any traffic at all in the 4th lane. The effectiveness issue is whether the 4th lane is better operated as HOV or mixed flow. This prototype finding fails to satisfy any of the above listed minimum effectiveness finding criteria.

4. FINDING PROTOTYPE D: "67 percent of car pools on the 55 did not exist prior to opening the (HOV) lane" [6]

The implication that those 67% of carpools on the SR-55 were *motivated* by the HOV lane is totally misleading.

Carpools have a finite average lifetime of several years. Old carpools dissolve and new ones are formed continuously, with or without an HOVL. Dissolution/replenishment processes such as carpool formation are generally modeled as exponential process meaning that the probability, P, that a given carpool will have been formed within the last T years is given by

$$P = 1-\exp(-T/L)$$

where L is the exponential lifetime of the process. Given T=2 years, and P=0.67, one can solve for the implied lifetime which turns out to be 1.8 years, well within the range of various type 4 findings [7]. In other words, the finding can be explained as the expected result of the natural carpool dissolution/formation process with a lifetime of 1.8 years, in no way dependent on, or implying any amount of additional of HOVL induced, or “benefit” carpooling. While this type of statement is *consistent* *with* HOVL induced carpooling, this type of finding does not in any way prove or even indicate it.

This statement fails to satisfy criteria #2, #3, #4.

5. FINDING PROTOTYPE E: "Over 90% of the carpoolers cite use of the carpool lanes as the primary reason for carpooling on Route 55" [8]

This finding is the result of a mailout questionnaire. In the particular case quoted, a number of other questions were asked, including one that established that 43% of the carpool respondents had been carpooling more than two years prior which was when the carpool on the route 55 first opened. In other words, at most 57 % of carpoolers could *possibly* have been influenced by the carpool lane in their choice to carpool. The 33% (=90-57) of carpoolers who had carpooled since before the lane was opened, and answered that the lane was their primary reason for the carpooling contradicted themselves. They were either mistaken or told a little fib.

One might expect that normally intelligent respondents would understand that as carpool lane users, it would be in their best interest with respect to future possible carpool lanes, to say that "use of the lane was a primary reason for my carpooling on Route 55". It would be naive to expect that the answers to such a question to be unbiased.

This finding fails to satisfy minimal effectiveness criteria #2, #3, #4.

6. Summary of the Five Classical HOVL Findings

All of the above classical finding prototypes fail at least three of the four minimum essential requirements for a valid effectiveness finding. All fail to

2. Make an *overall* assessment including the entire HOVL impact corridor.
3. Account for positive and negative components for a *net* impact assessment.
4. Show net impact on congestion, travel-time, emissions, or energy consumption.

7. Weight of Evidence Issue

While none of the above finding types satisfy the minimum requirements for a whole effectiveness proof, there is a widespread notion that the weight of evidence of at least five different types of finding, reaffirmed in hundreds of different places and times, may add up to a whole proof of the issue.

This concept probably has its roots in a basic statistical theorem, that in the present context could be stated as follows:

> Suppose one makes a non-conclusive yes/no measurement indicating whether his HOV lane is effective or not and finds that it is effective ("yes").
>
> But suppose further there is a significant error probability, say, p =1/3 of error in that finding. Then taking account of only this one measurement, the probability that the finding is in error is 1/3
>
> Suppose another party makes the same kind of measurement and also finds "yes" again with error probability, p = 1/3.
>
> *Under certain conditions*, taking account of *both* measurements, the probability of error, is reduced to, p = 1/9.
>
> If under those same *certain conditions* this measurement were repeated many, many times with the same result, the overall probability of error would shrink to approaching zero. We could be confident that "yes" was the correct answer, HOV lanes would be proven effective.

It appears that the instinctive feeling everyone has for this theorem is responsible for the "weight of evidence" concept. The problem lies with the "*certain conditions"* qualification. That condition is that the errors in the various measurement replications must be "statistically independent" or loosely speaking, uncorrelated. If the error is due to a common but unknown bias, then no amount of replication of the experiment can improve the probability of error.

An example may illustrate the point. Suppose one wants to measure the direction to a given marker with a compass. If the compass is perfect but the measurer has shaky hands resulting in unbiased random errors of +/- 10 degrees, that random error can be reduced to as small as desired by averaging replicated measurements. On the other hand, if the hands are steady but the compass has a 10 degree *bias* error of unknown sign, no amount of replication will improve the measurement.

The error, if any, of the classical HOV findings discussed above are strictly of the latter, constant but unknown bias type. Essentially, they almost all seem to indicate "yes", the HOV lanes are effective. That bias is essentially zero if the lanes are in fact almost always effective. On the other hand it is a large positive bias if they are not. And there is no way to be sure there is not such a large positive bias, other than to determine by a *real* proof, whether the HOVL really is effective. There is no way to use the possibly biased measurements themselves to determine their own common bias.

Findings such as these do not accumulate "weight of evidence" regardless of how many times they may be replicated.

8. Summary

1. We have looked at the five classical types of HOVL findings.
2. We have found that all of them fail to
   - take into account effects on all persons and vehicle impacted by the choice to operate the lane in HOV mode.
   - take into account all the negative as well as positive impacts of the choice to operate the lane in HOV mode
   - show that any of the quantifiable benefit measures, congestion, travel-time, emissions, energy consumption , or accidents are actually improved by the choice to operate the lane in HOV mode.
3. We have found that the nature of the error if any in these types of HOVL findings would be that of a constant but unknown bias, that does not enjoy any reduction in error by replications. There is no added weight of evidence.

9. Is It Possible to Determine HOV Lane Effectiveness?

So far here we have considered only purely *experimental* determinations. I'm convinced there is no way to do so by purely experimental means. On the other hand, there have been a number of *analytic* estimates that do satisfy the above "minimal criteria" using the methodology of traffic planning network modeling. These computer models have the following advantages over experiment:

- inherently able to take all important effects into account simultaneously
- inherently look at negatives as well as positives
- quantify time-saving, emissions, etc.
- no exogenous disturbance problem
- effectively unlimited measurement resolution (though not unlimited accuracy).

The down side is the dependence on the accuracy of model calibration. This, however, is mitigated to a large extent by the fact that HOV lane studies are looking for the performance response to relatively small incremental system perturbations. Models can be quite accurate incrementally even though the absolute accuracy is a good deal less. The nine known instances of modeling studies of HOV vs. Mixed-flow operation of a facility are summarized in "The Best Evidence on HOV Lane Effectiveness"[9] on line at <www.urbantransport.org/bestev.pdf>. These studies are in reasonably close agreement as to HOV effectiveness. Any modeling errors are further mitigated to some extent by the fact that such modeling errors, using mostly independently developed models can probably be regarded as statistically independent so that "weight of evidence" *does* apply.

10. A Hybrid Experimental-Analytic Measure

A hybrid of experimental and analytic procedure may afford the best determination of

HOVL performance satisfying all the minimal effectiveness criteria. This would consist of a short term (e.g. within 1 month) A-B (HOV vs. MF lane operational alternatives) with extensive corridor measurements of flow and speed in both phases of test. This would capture *most* of the difference between the A-B states, including all the simple diversions of persons and cars, speeds, and the carpool time-saving, missing only the second-order effects of benefit carpooling induced by the HOVL time saving. As a small, second-order perturbation effect of a single factor, it can be estimated quite well by a linear expansion (sometimes in this context called pivot-point expansion) of the locally calibrated modal choice estimation model, a part of the local traffic planning model available in most metropolitan areas. That is,

$$FPH_{hovl} = FPH_{mf} \times (1-C_t \times TS)$$

where

FPH_{hovl} , FPH_{mf} are the fraction of persons carpooling in HOVL or MF modes

TS is the time saving of HOVL over the other lanes in HOVL mode11.5

C_t is the modal choice modeling utility coefficient for travel time saving for all trip purposes, a number typically in the range of – 0.01 to – 0.03 per minute, the elasticity coefficient of fraction of persons carpooling with respect to travel time saving.

This is suggested as a viable candidate for the best possible HOV lane effectiveness "hybrid" measure-analysis of HOV lane effectiveness.

REFERENCES

1. The Santa Monica Diamond Lanes", SYSTANS Inc., Report No. D148-8, for U. S. Dept. of Transportation, Catalog No. PB286567, September, 1977.
2. "The Effectiveness of High-Occupancy Vehicle Facilities", Institute of Transportation Engineers, Wash. D. C., ITE Publ. No. IR-050, 1988.
3. "An Evaluation of High-Occupancy Vehicle Lane in Texas", Research Rpt. 1353-3. Texas Transportation Institute, November, 1995.
4. "Route 55 Freeway, One Year Report", CALTRANS, December 29, 1986.
5. "Route 55 Commuter Lane Demonstration Project: 90-day Evaluation Report. OCTC, February, 1986.
6. "Route 55 Commuter Study, Report on the 1987 Mailback Travel Survey Results", Orange County Transit District, February 1988.
7. Shu, Jarvia and Glazer, Jesse, "Carpool Evaluation Program", CTS (Commuter Computer) 1979.
8. "Commuter Attitude Survey - 1984", Orange County Transit District, 1984.
9. Mallinckrodt, A. J., "The Best Evidence on HOV Lane Effectiveness", AJM Engineering, August 23, 1998, on-line at <www.urbantransport.org/bestev.pdf>