Estimating the Contribution of Various Factors to Variations in Bus Passenger Loads at a Point

ROBERT M. SHANTEAU

A procedure for estimating the relative contribution of various factors to the variation of passenger loads on buses at a point is described. These factors include unequal bus headways and the uncertainty in the number of passenger loads on buses at a point than all other factors combined. It is assumed that bus headways contribute far more to variations of passenger loads than all other factors combined. It is also assumed that bus headway does not change much over a typical passenger-trip length. Data are presented to show that, in the typical case, unequal bus headways contribute far more to variations of passenger loads on buses at a point than all other factors combined. It is rare that headways are so well controlled that their contribution becomes comparable with that of other factors. Where headways are poorly controlled, the public would most likely benefit from investments in headway-control strategies. In principle, the cost of controlling headways can be comparable with that of other factors.

The number of buses an urban transit agency must provide on a busy line is usually determined by the peak passenger flow past the maximum load point. Most often, the objective of the bus company is to use the fewest number of buses while still providing an acceptable level of service, where level of service is defined in terms of overcrowding and passenger waiting time.

Overcrowding is undesirable because (a) it causes discomfort and inconvenience to the passengers and (b) it makes circulation within the bus difficult and thus causes the bus to spend more time loading and unloading. In this paper, passenger waiting time is not used as a measure of performance. Instead, passenger loads are used because the bus company itself is mostly concerned with overcrowding.

Most bus companies try to provide some excess capacity to compensate for variations in loads that occur from day to day. These variations are caused by the stochastic nature of passenger arrivals and by unequal headways (elapsed time between buses). It was recognized as early as 1916 (1, p. 156) that unequal headways can be a major cause of variations in passenger loads.

Since fewer buses are needed on a line for which variations in loads are small than on one for which they are large, it is usually to the economic advantage of the bus company to use a control strategy that attempts to equalize headways (2). However, because traffic conditions and passenger stops cause variations in the travel times of buses, it is difficult to prevent deterioration in the regularity of headways (3). It is useful for the bus company to know what improvements in variations of loads it can expect from a range of control strategies so it can properly allocate its resources between vehicles and control.

PROBLEM STATEMENT

The practitioner is faced with the problem of how to evaluate various strategies without actually implementing them. In this paper, it is assumed that the future is sufficiently like the past that we can predict what would happen under different strategies in the future by analyzing data collected in the past. Basically, we want to relate the load on a bus to its headway in such a way that we can predict its load for different headways but for the same general passenger arrival pattern. First, however, we must set up a structure for the problem in which we can carefully state our assumptions.

ANALYSIS

Bus Trajectories and Passenger Arrivals

Suppose, for a particular day in the past, we plotted the time of arrival, origin stop, and destination stop of every passenger on the line on a time-space diagram, as shown in Figure 1. Here, the x's represent passenger arrivals and the subscript the destination stop. Multiple subscribers mean that several people arrived at a stop at the same time. On the same diagram, we also plot the trajectory of each bus.

Usually, when a bus stops, it picks up all passengers who have arrived since the last bus left. For instance, in Figure 1, bus 3 picks up two passengers at stop 2, one each destined for stops 4 and 5. Assume that each bus always has room for all passengers waiting for it. Then the load on a bus at a particular point is the number of people who have arrived between that bus's trajectory and the previous one's and who desire to travel beyond that point. Passengers who got off prior to that point or get on after it are irrelevant. For instance, bus 3 is carrying seven passengers as it leaves stop 5. Note that, where bus drivers refuse passengers because their buses are already too full, the actual load may not match precisely the load that we have constructed. Note also that only departure times from stops are relevant and that the details of the trajectories between stops can be ignored.

Assumptions

Suppose that passengers' arrival times, origins, and destinations in no way depend on buses' departure times or on the strategy being used. That is, in Figure 1, if the buses had arrived at different times, the location of the x's and the subscripts would not change. This would happen, for instance, if buses arrived either so frequently or so unpredictably that passengers did not check the time before leaving for their stop. Then if we were to hypothesize a different set of trajectories for the buses, we could still easily construct the load each bus would have carried.

One difficulty with the procedure presented so far is that we need to know the passenger arrival times and the bus departure times at every stop. Suppose that bus trajectories are parallel or, in particular, that the time between a bus's departure at one stop and its departure at the next is the same for every bus (but not necessarily the same at every stop). If we now observe the departure times of all buses at one stop, we know the departure times of all buses at all stops. Essentially, we have assumed that a bus's headway does not change significantly over a distance comparable to the length of an average passenger trip. This allows the headway to change slowly. On a well-controlled bus line the headways do in fact change slowly.
This assumption also makes the description of passenger arrivals easier. In particular, it implies that the bus travel time from any point \( x \) to the point in question, \( x_0 \), is fixed. Let us call this travel time \( \tau(x, x_0) \). Suppose we observe two passengers, one who arrives at stop \( x \) at time \( t \) and another who arrives at stop \( x_0 \) at time \( t + \tau(x, x_0) \). Because bus trajectories are parallel and the arrival time of both passengers is the same relative to a given trajectory, both passengers catch the same bus. For our purposes, the two passengers are equivalent, and we can replace the original arrival process by an equivalent one at \( x_0 \) for which the arrival time of a passenger at \( x \) is shifted by \( \tau(x, x_0) \). (Note that passengers who alight before the bus reaches \( x_0 \) are ignored.)

If we know the passenger arrival times for the equivalent process, then for any set of bus departure times from \( x_0 \) we could construct the loads on the buses at \( x_0 \). This procedure, however, still depends on the exact bus departure times (as opposed to simply the headways), since, for instance, the average rate of passenger arrivals might change over time and the loads on two buses with the same headway might be different. Let us assume that passenger arrivals are stationary, so that their average arrival rate is constant. (The extension to non-stationary arrivals requires only that a headway be replaced with the time necessary to pick up a given number of passengers.) Then the expected load on a bus is simply proportional to its headway. That is, if \( L/H = h \) is the random variable associated with the load on buses that have a headway \( h \), then the expected value (or average) of the loads on these buses should be

\[
E(L|H = h) = mh
\]

where \( m \) is the average passenger arrival rate. Because of the uncertainty in the number of passengers to arrive during a headway \( h \) (which is here called the uncertainty in passenger arrivals), the load on a bus with headway \( h \) will vary about \( mh \).

Similarly, the average load over all buses, regardless of headway, is

\[
E(L) = mE[H]
\]

Because of unequal headways, uncertainty in passenger arrivals, and perhaps other factors, the load on an individual bus will vary about this mean. Suppose for the moment that only unequal headways contribute to the variability of loads on buses—i.e., passengers arrive in a steady stream. Then a bus with headway \( h \) would carry exactly \( mh \) passengers. If we were to plot the distribution of loads under this supposition and compare it with the distribution of actual observed loads, then any difference observed would be caused by factors other than unequal headways.

**Example**

Figure 2 shows a comparison such as that proposed above for data gathered for two different time periods on a bus line in Oakland and Berkeley, California. (The data were collected in November 1977 at the maximum load point of the northbound 51 line of Alameda-Contra Costa County (AC) Transit. This load point is downstream from a transfer point of the Bay Area Rapid Transit system.) The details of the data collection are given elsewhere (4). Basically, loads and headways of buses at the maximum load point were observed during two time periods when the passenger arrival rate was nearly constant. In this case, the maximum load point was downstream from a rapid transit station and any difference between the two distributions represents the effect not only of the uncertainty in the arrival of bus passengers but also of the uncertainty in the arrival of train passengers upstream of the transfer point and the uncertainty as to whether a bus picks up a batch of transfers from the rapid transit station at all.

That the distributions nearly coincide for each of the two time periods shows that the contribution of factors other than unequal bus headways to the variability of loads on buses is unimportant. In particular, uncertainty in passenger arrivals contributes little. Only if the headway distribution were considerably narrower than it was for these buses would uncertainty in passenger arrivals become significant. That is, unequal headways cause most of the variation in loads.

**Variance Calculations**

The phenomenon cited above can be quantified. If buses arrive either so frequently [i.e., at headways less than about 10 min (5)] or so unpredictably that passengers arrive independently of buses, then the variance of the load on a bus \( \text{Var}(L) \) is related to the variance of the headway of a bus \( \text{Var}(H) \) and a function characteristic of the passenger arrival process \( r(h) \), as follows (4):

\[
\text{Var}(L) = m^2 \text{Var}(H) + mE[H^2]r(h)
\]

Note that the variance and expected value are taken with respect to the headway distribution. In the second term, the expected value is taken of the function \( hr(h) \), where \( r(h) \) is the variance-to-mean
and the standard deviation of the load is about
buses might be dispatched so that the average load
of r (h) as well as the complete headway distribu­
individual bus quite easily could carry as many as 27
is equal to the number of seats on a bus, an indi­
load variance of unequal headways, and the second
again that unequal headways, represented by
is not). Let this constant be r. Then mE[Hr(H)]
for a bus line, E(L) = 50, so C(H) would have to be
insized bus, E(L) = 50, so C(H) would have to be
such a small vari­
was 4.05 and 4.58. As seen in Figure 3, r(h) is about 2-3 for headways of 2-6 min (the aver­
more to the vari­
for either time pe­
then headways and passengers contribute about
If r (h) is a constant, 1. It has
been shown elsewhere (4) that, for values of h near
r(h) is not constant. Its shape is typically as
if passengers arrived in a Poisson process, for
were 0.63 and 0.61, and E(L) was 41.4 and
that this statement is also true if we take square
rule is this: If C(H) E(L) is comparable to
Then (2.5) 1/2 = 1.2 is
sions for the various contributions.
for instance, when C(H) is about 0.2 and E(L) is
er ratio of the number of passengers to arrive in a
taken as a function of h. If passengers arrived in a Poisson process, for
were 0.63 and 0.81, and E(L) was 41.4 and
for instance, then r(h) would be a constant, 1. It has
been shown elsewhere (4) that, for values of h near
would be zero; if there were no variability in the
arrival of passengers, r(H) would be zero. Thus,
the first term [m² Var(H)] is the contribution to
we will let r be 2.5. Th e n (2.5) 1/2 • 1.2 is
have to be about 0.05 for headways and passengers to contribute equally to
Var(L) = Var(L)/E(L) (5)
Now the analysis of whether headways contribute more
or less than passengers to the variation in loads
comes down to whether C(H) E(L) is large or small relative to r.

Rule of Thumb for Buses and Rapid Transit
The rule is this: If C(H) E(L) is comparable to r, then headways and passengers contribute about
equally to variations in the load on a bus. Note
that this statement is also true if we take square
roots: If C(H) [E(L)] 1/2 is comparable to r1/2,
then headways and passengers contribute about
equally. For the two time periods analyzed above,
C(H) was about 0.63 and 0.61, and E(L) was 41.4 and
and is given elsewhere (4). Table 1 gives the nu­
unequal headways contribute far more to the varia­
the bus company to invest in control strategies that
starts to dominate.
For the two time periods analyzed above,
three years later, it was found to be 0.20.
The same analysis can be used for systems that
have larger or smaller average passenger loads per
vehicle. For instance, a rapid transit system might
schedule its trains so that they carry about 1000
passengers. Then C(H) would have to be about 0.05
for headways and passengers to contribute equally to
Var(L).

CONCLUSIONS
In effect, C(H) is a measure of the variability of
the headway (specifically, the standard deviation)
on a scale of the mean. This analysis shows that,
for a bus line, if C(H) is above about 0.30, then
unequal headways contribute almost exclusively to
the variability of loads on buses. In this case, we
would say that headways are poorly controlled. On
the other hand, when C(H) is below about 0.3, un­
equal bus headways and uncertainty of passenger ar­
vivals contribute about equally. In this case, head­
ways are well controlled. In fact, it does
little good to reduce C(H) below about 0.2 because
under that value the uncertainty in passenger arriv­
als over which the bus company has little control,
starts to dominate.

When headways are poorly controlled, it might pay
the bus company to invest in control strategies that
reduce the variance in headways. If it can thus re­
duce the variance in loads, it can either use fewer
buses and tolerate the same amount of overcrowding
or it can reduce overcrowding and use the same num­
ber of buses. In the first case the bus company
saves money, and in the second case the public re­
ceives better service. In either case, the public
benefits (2).
This paper does not discuss the cause of unequal
headways or the cost of controlling headways. It
does illustrate that, once unequal headways occur,
they are the dominant cause of variations in loads.
A strategy to control headways, of course, must be

![Figure 3. Typical shape of r(h) for AC Transit bus line in Oakland and Berkeley, California.](image-url)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Value of Contribution (passengers²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unequal bus headways</td>
<td>665</td>
</tr>
<tr>
<td>Uncertainty in the arrival of</td>
<td>63</td>
</tr>
<tr>
<td>nontransferring passengers</td>
<td>1</td>
</tr>
<tr>
<td>Unequal rapid transit-train</td>
<td>7</td>
</tr>
<tr>
<td>headways</td>
<td>2</td>
</tr>
<tr>
<td>Uncertainty in the arrival of</td>
<td>14</td>
</tr>
<tr>
<td>transferring passengers</td>
<td>1</td>
</tr>
<tr>
<td>Uncertainty that a bus picks</td>
<td>47</td>
</tr>
<tr>
<td>up a batch</td>
<td>41</td>
</tr>
<tr>
<td>Total</td>
<td>796</td>
</tr>
</tbody>
</table>

Table 1. Contributions of various factors to variance of load on a bus.
The basic goal of all transit agencies is to provide transit service in the most equitable and cost-effective manner. The development of criteria defining the number, size, and location of fixed facilities constitutes a key element in the realization of this goal. A mislocated or improperly sized facility can, over a few years, account for millions of dollars in wasted funds. Conversely, the dollars saved by optimally locating and sizing these facilities can be more effectively used in other areas of system operations.