Under diskussionen på seminariet på KTH 25 oktober 2018
refererade Harald Lang till denna skrift

presentation vid

”Seminarium om persontransportmodeller för långväga resor”

Trafikanalys 8 oktober 2014

Harald Lang

On Computing Consumer Surplus in Long Distance Public Transport

Revised and extended 3/4–10

Harald Lang

Long distance public transport

One important property with long distance public transport, as opposed to short distance (or urban) transport, is that departures take place with a low frequency, which implies that it is reasonable to assume that travellers find out the *actual* frequency delay ("generalised" cost for the fact that the departure doesn't take place at the ideal time) rather than the *expected* frequency delay. The implication of this is that the generalised cost (G) for a particular choice of travel alternative (modes, lines) is *endogeneous* in the sense that it depends not only on the ticket price, time schedule etc. for that particular choice, but on the time schedules for all substitute alternatives.

My concern is the assessment of consumer surplus in this situation. There seems to be a rather relaxed attitude among many transport economists on this issue; the view is that standard methods like "logit sums" or "rule of a half" will take care of this once we have calibrated our demand model. I am convinced that this confidence is unwarranted. The danger is that erroneous assessments of consumer surplus will escape detection, since there are rarely directly observable data on consumer surplus (as opposed to data on demand, for example.)

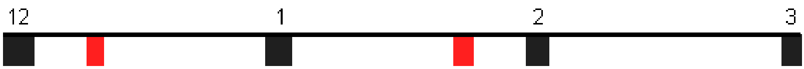
In this report I will test two common ways to assess consumer surplus in the current context (i.e., where travellers consult time schedules prior to deciding on mode or route.) I will perform the test in a similar way that the efficiency of estimation procedures on small samples are conducted. It is done by simulation: the tester decides on a stochastic process with parameter values set by him, then he simulates data from this stochastic process. He then pretends that the parameters are unknown, and estimates them from the simulated data, using the various estimation procedures he wants to evaluate.

I will specify a very simple demand model for public transport between two destination points. This test model is "reality". I will then use the logit model and the "rule of a half" procedure to estimate the (known) shift in consumer surplus when one of the headways are changed. We can then see how well or bad these procedures perform in this context. The important feature of "reality" is again that travellers consult time schedules.

The test model

- Travellers want to go from destination P to destination Q.
- There are two buses, Black and Red, which both go from P to Q,
- they both have the same ride time and fare.
- However, Black has a headway of 60 minutes, Red 90 minutes. Black leaves at 12pm, 1pm, 2pm etc. Red leaves at 12:15pm, 1:45pm, 3:15pm etc.
- There are 300 travellers, each with a personal ideal departure time. These departure times are uniformly distributed between 12pm and 3pm (note that the schedules will repeat in the same manner with a period of 3 hours.)
- Travellers have to be at their destination in time, so they take the bus that leaves at the latest time before (or exactly at) their ideal departure time.
- Travellers consult a time table to decide which line to chose.

The schedules are illustrated below:



Hence, travellers with ideal departure time between 12pm and 12:15pm will chose Black at 12pm, travellers with ideal departure time between 12:15pm and 1pm will chose Red at 12:15pm, etc.

Let us denote the time from actual departure to ideal departure "waste time", which is a kind of generalised cost, G. It is now easy to see, that

– 200 travellers will chose Black, with a mean waste time, i.e. “mean generalised cost” MGC=24.375 minutes.
– 100 travellers will chose Red, with a MGC 18.75 minutes.

Now let us consider a change of headway for the Red buses. Assume that they too leave with an interval of one hour: 12:15pm. 1:15pm, 2:15p,. etc.

Similar calculations show:

– 75 travellers will chose Black, with a mean generalised cost MGC=7.5 minutes.
– 225 travellers will chose Red, with a MGC 22.5 minutes.

Hence,

– the demand for Black goes down from 200 to 75, MGC goes down from 24.375 min. to 7.5 min.
– the demand for Red goes up from 100 to 225, MGC goes up from 18.75 min. to 22.5 min.

Note three important features

1. The MGC for travellers *with Black* goes down by 69%, although the change in schedule is for done for *Red*.
2. MGC and demand goes in the *same direction* i both cases: when price (MGP) goes up, demand also goes up; when price goes down, demand goes down.
3. An apparent *improvement* in the Red schedule—a reduction of headway with 33%—leads to *higher* MGC for that mode.

In this example, we see that the total delay time for all passengers is

$$200 \cdot 24.375 + 100 \cdot 18.75 \text{ minutes} = 6'750 \text{ minutes.}$$

After the frequency change of Red, the total delay time for all passengers is

$$75 \cdot 7.5 + 225 \cdot 18.75 \text{ minutes} = 5'625 \text{ minutes.}$$

The reduction of delay time is thus 6’750−5’625 minutes = 1’125 minutes

The Logit model: Consumer Surplus as Logsums

We specify the deterministic part of the utility of travelling by bus *i* by

$$v_i = u - c \cdot h_i$$

where $c=\ln(2)/30$, h_i=headway, and u is some fixed level of utility for the journey. When h=90 for Red and 60 for Black, this gives a probability for Black equal to

$$P(\text{Black}) = \exp(u - c \cdot 60)/[\exp(u - c \cdot 60) + \exp(u - c \cdot 90)] = 2/3$$

which is in accordance with the example, and, by the same token, P(Red)=1/3. Notice that I have scaled the utility (the parameter c) such that the probabilities are correct, and the assumed Gumbel random terms are normalised (0,1). The model is thus calibrated to the observed demand. The consumer surplus per traveller is now given by the "logsum":

$$CS_0 = \ln[\exp(u - c \cdot 60) + \exp(u - c \cdot 90)] = u - 0.980829$$

After the change of headway for Red, the CS is

$$CS_1 = \ln[\exp(u - c \cdot 60) + \exp(u - c \cdot 60)] = u - 0.693147$$

The increase of CS per traveller is thus $\Delta CS = 0.980829 - 0.693147 = 0.287682$, which corresponds to a total reduction of delay time of

$$300 \cdot 0.287682 / c \text{ minutes} = \underline{3'735 \text{ minutes}}$$

The logit model hence exaggerates the gain in CS by 230 %. The model will also erroneously predict the change in demand. Indeed, after the reduced headway of Red the two alternatives are symmetrical (equal headways,) so the predicted probability for Black, say, is

$$P(\text{Black}) = \exp(u - c \cdot 60)/[\exp(u - c \cdot 60) + \exp(u - c \cdot 60)] = 1/2$$

rather that the true value P(Black) = 1/4.

Employing the "Rule of a Half"

The "traditional" way to compute (changes in) consumer surplus (CS) is to employ the rule that CS is the "area under the demand curve". But in the current context it is unclear what the definition of "demand curve" is. The very notion of "demand curve" is that price is exogenous to the consumer (he is price taker) whereas he chooses the quantity to buy according to a utility maximising scheme. In the current context this is still true, but the actual "price" differs among travellers (because of their heterogeneity as to ideal departure time,) and since they have full information of all alternatives (substitutes,) they will allocate themselves in such a way that on the aggregate, the mean generalised cost is not exogenous, but endogeneous, to the aggregate of travellers.

In the example above, what would the "demand curve" for Red be? What would the "demand curve" for Black be? Wouldn't any candidate for such a curve be *upward sloping* according to the points 1−3 made earlier?

It has been suggested that the MGC (in terms of frequency delay, in this case) for each of the two buses be defined as the mean cost for all 300 travellers, i.e. both those who choose the bus under consideration *and* those who choose the other bus. In this case, the MGC is simply half the headway. Maybe, with this definition the gain in surplus (total reduction in "cost", i.e., frequency delay time) can be computed by comparing areas under a "demand curve", i.e., by the "rule of one half" as an approximation? Let us investigate this possibility.

What would the "rule of one half" give? The MGC for Red has decreased from 45 to 30 minutes. The demand has gone up from 100 to 225. The reduction of cost would then be, according to this rule:

$$0.5\cdot(225+100)\cdot(45-30) \text{ minutes} = \underline{2'437.5 \text{ minutes}}$$

This is an exaggeration by 117 %.

A Composite Good

It is indeed possible to employ the "rule of a half" to assess the CS if we consider the two bus modes as a composite good. It is imperative, though, that we then assess the correct generalised price of this good. First we compute the total cost in terms of frequency delay. It is equal to 6'750 minutes, i.e., 22.5 minutes per passenger. After the reduction of headway of Red, the cost is 5'625/300 = 18.75 minutes. If we define 22.5 minutes and 18.75 minutes as "generalised prices", then the "rule of a half" gives the gain in CS:

$$0.5 \cdot (300+300) \cdot (22.5-18.75) = 1'125 \text{ minutes}$$

which is the correct value. This comes as no surprise, of course, since we have just made a circular computation. In order to assess the "correct" generalised prices, we have to *first* calculate the CS, then use these prices to compute the very same CS. However, this procedure could be commendable in the case when total demand is elastic. Assume, for instance, that the reduction of headway of Red attracts another 25 travellers. Then a reasonable approximation of the gain in total CS would be

$$0.5 \cdot (300+325) \cdot (22.5-18.75) = 1'172 \text{ minutes.}$$

However, this correct way of using "rule of a half" is infeasible for computing CS, since it assumes that we *already know* the CS in order to calculate the "correct" generalised prices.

"Integrating over all incidents"

A possible problem with the "true" model is that the analyst needs to know the time schedules of the two buses. In a realistic situation with many O-D pairs, this may require an infeasible amount of information. A common approach is to "integrate over all incidents" which means that the analyst knows the headways, but not the actual phasing of departure times. There is hence a continuum of possible schedules pertaining to these headways. The procedure is then to take an average over all possible schedules, with a "diffuse" prior, i.e., all possible schedules are given the same weight. If h_1 and

h_2 are the headways of Black and Red, respectively, then the proportion of travellers choosing Black is calculated as

$$\text{P(Black)} = \frac{1}{h_1 h_2}\int_0^{h_2}\int_0^{h_1} \text{step}(y-x)\,dx\,dy$$

("step" is the Heaviside function, defined as step(x) = 1 if x>0 and = 0 if x<0) and so on. In the current example the demand for Black, prior to the change in headway for Red, is 200, and the demand for Red is 100. The total delay time for those choosing Black is 5'000 minutes, and for those choosing Red 2'000 minutes. The initial total cost is thus 7'000 minutes.

After the reduction of Red's headway, the demand for either bus is 150 passengers, and the average delay time (per passenger) is 20 minutes. Note that this is 1/3 of the headway, *not* 1/2, as would be the case if passengers did not consult time schedules! Hence the total delay time is 300· 20 = 6'000 minutes. The gain in CS is thus

$$\Delta\text{CS} = 7'000-6'000 = 1'000 \text{ minutes}$$

This is an underestimate of the true value by 11 %. It is not perfect, but at least in the same ballpark. If the phasing of departure times varies over the day or week, then this averaging procedure could be expected to perform even better.

Conclusions

When travellers are assumed to consult time schedules, computing consumer surplus is considerably more demanding than when they make their decision based only on headways (and other characteristics unrelated to departure / arrival times). The often overlooked feature is that when the headway is changed for one mode, the average delay time will be affected also for other modes. A problem is that erroneous estimates of consumer surplus may pass undetected. A plain vanilla logit model can not handle the situation, nor can a simple "rule of a half". The RDT (Random Departure Time) model seems to be the best suited for the task. However, in its simplest form that model does not take individual preferences into account, but Odd Larsen has shown how the RDT model can reasonably easily be combined with a “discrete choice” specification so as to remedy this limitation; see his report *“A note on discrete choice and assignment models”*