



***U.K. Preference Valuations
of Soft Factors:
How Credible are the Results?***

Mark Wardman

- My impression that SP Valuations of Soft Factors are often not credible
- Are these concerns well founded (and shared)?
- To scope a major literature review on the subject

Many reviews of Value of Time, Congestion and Crowding multipliers, Price Elasticities, Time Elasticities etc.

What drives the values obtained from Soft Factor valuation studies?

- To stimulate discussion on better ways forward (although perhaps not the best way to start a workshop on Soft Factors!)

WHAT ARE SOFT FACTORS



UNIVERSITY OF LEEDS

- Not Timetable nor Travel Time Related (although may influence marginal utility of travel time)
- Not Price Related
- Not the reliability of delivery
- Generally of 'secondary' importance
- Often difficult to measure/represent
- On-board

Seat comfort, ambience, noise, security, information, cleanliness, etc.

- Off-board

Information, shelter, seating, CCTV, lighting/heating etc.

%RP	UK	NON-UK
Time	5%	23%
OVT	14%	30%
Headway	7%	17%
Departure Time Shift	5%	0%
Reliability	0%	10%

SP almost all values of soft factors

Numerous studies undertaken even in UK

WHY DOES SP DOMINATE SOFT FACTORS?



UNIVERSITY OF LEEDS

- Generally small effects
 - Need large sample*
 - Other variables dominate*
- Few real world choice contexts (or costly to explore)
- Some levels do not currently exist
- Lack of variation/difficulty of measuring

- **Bates IATBR 1994**

“In the area where SP might be considered to have the most to offer – that relating to so-called “soft variables” like comfort, cleanliness etc – the direct valuations provided are often unconvincing. This had led to various scalings being applied (to take account of “packing”, for instance) which are difficult to justify theoretically”

“The problems have emerged when the values have been added over a significant number of attributes..... By the time say 10 of these relatively minor aspects are included, it may appear that passengers are prepared for the fare to more than double. Such conclusions do not generally appear plausible.”



- **and also:**

“While “packaging” may be a practical way forward, the theoretical problem remains for SP that the direct valuations between changes in money cost and improvement in soft variables appear to vary according to the number of soft variables included in the experiment. Without a convincing explanation, this must seriously undermine the confidence we can have in such valuations”.

- **Adamowicz, Boxall, Louviere and Swait (1999)**

“Strategic behaviour should be minimal in SP tasks since the choices are made from descriptions of attributes and it will not be clear which choice will over- or under-represent a valuation”



LONG BEEN CONCERNS (III) UNIVERSITY OF LEEDS

Wardman and Whelan (2001) – review of 18 UK studies covering improved rolling stock (overall not component attributes)

“There is considerable variation in the values across studies and indeed an initial impression is that many of the values appear to be large”

Study for Local Authority (2004)

“It has been found necessary to use this upper value to cap, or limit, the benefits obtainable from a bundle of improvements where the sum of the part values may otherwise become infeasibly large”

Steer Davies Gleave (1990) (for London Underground)

SP studies “have in the past appeared to give implausibly high values”

WHY LARGE VALUES?



UNIVERSITY OF LEEDS

Here dealing with:

- 1) Overall Values (new train, new terminus, SP package etc)
- 2) Component Values (seat comfort, ventilation, décor etc)

Sum of Component Effects creates Large Values

- Interactions
- Budget Effect
- Halo effect
- Focus Effect

But could also be:

- Strategic Bias (Also applies to overall values)

- Ratio of Package Value to Sum of Components typically 0.3-0.7
- Little evidence that interactions (but little testing)
- Strategic bias incentive given transparent purpose of study

If component values can be subject to strategic bias, surely the package value can be?

Current convention relies on accuracy of package value

CONTEXT – HOW WELL DOES SP GENERALLY PERFORM?



UNIVERSITY OF LEEDS

- Abrantes and Wardman (2010) 1794 UK values
 - RP IVT 22% larger*
 - RP walk 43% larger*
 - RP wait 45% larger*
- Recent European Meta Analysis – additional 1247 European values from 25 countries
 - RP values 20% larger*
- So some concerns even for hard variables!

THE CONVENTIONAL APPROACH



UNIVERSITY OF LEEDS

The responses to large component valuations has been:

- *Best scenario/package*
- *Assume relativities are ok and rescale (cap) to package value*
- *Cheap Talk and other wording to limit values (but rare in transport)*
- *Little testing of interactions/package effect*
- *Little attempt to make less transparent (can it be for components?)*
- *Little attempt to obtain real world corroboration/values*
- *Little interest SP might be wrong*

SO WHAT IS THE EVIDENCE



UNIVERSITY OF LEEDS

- ‘Laugh Test’ – Individual Attributes **and** Packages
- RP Evidence
- Compare with Elasticity/Demand Evidence
- Other Insights

LAUGH TEST AND INDIVIDUAL ATTRIBUTES (I)



UNIVERSITY OF LEEDS

- Changes in Quality of Motorway Journeys

Inter-Urban Motorists

Mean journey times 2 hours 52 minutes

SP exercises traded-off time versus three soft attributes at a time

- Valuations obtained relative to perfect surface

Once in journey slight rumbling for 2.5 minutes = 12.8 mins

Once in journey substantial litter on verge = 6.1 mins

Once in journey some text on road sign unclear = 9.2 mins

Once in journey you notice some broken glass on road = 18.2 mins

Once in journey notice lane marking missing in places = 7.6 mins

Once in journey water on road splashes on windscreen = 10.4 mins

Once in journey vegetation looks slightly unkempt = 4.7 mins

LAUGH TEST AND

INDIVIDUAL ATTRIBUTES (II)



UNIVERSITY OF LEEDS

3 SEPARATE INTER-URBAN RAIL ROLLING STOCK STUDIES VALUES % OF FARE PAID									
	Study 1			Study 2			Study 3		
Improved	1st	EB	Leis	1st	EB	Leis	1st	EB	Leis
Décor	6	10	3	9	10	6	8	11	6
Information	3	5	3	3	4	3	3	4	3
Audio	2	2	2	2	3	2	2	2	2
Toilets	4	4	3	5	5	4	5	6	4
Vestibule	1	1	2	1	1	1	1	3	1
TOTAL	16	22	13	20	23	16	19	26	16

LAUGH TEST AND PACKAGES (I)



UNIVERSITY OF LEEDS

- Valuation of Quality Bus Attributes (2001)
- Covers security, information at bus stops, reliability, driver manner and whether new buses or not
- Car users value package of attributes at £2.60, exceeds 90th percentile of bus fares
- Bus users value package at £0.32, a third of fare paid

LAUGH TESTS AND PACKAGES (II)



UNIVERSITY OF LEEDS

- Station Improvements 2004
- Bus Package Improvement
 - Waiting: Basic Shelter to Glass Cubicles under Canopy*
 - Security: None to recorded and monitored CCTV*
 - Staff: None to present 7am to 11pm*
 - Information: Paper timetables to paper plus electronic display*
 - Help point: None to present*
 - Toilets: None to provided and cleaned regularly*
- Rail Package Improvement
 - Add provision of free and secure parking*
- Tram Package Improvement
 - As rail but no staff and no toilets*
- Valuations of Package
 - Bus: 42p (46% of fare paid)*
 - Rail: 61p (50% of fare paid)*
 - Tram: 53p (43% of fare paid)*
- Sum of Valuation of Individual Components far higher

- A bit more imagination and effort identifies RP contexts
- High quality buses exist
- Trade-offs between different routes
- Select homes where longer walk to quality bus
- Sample individuals who were aware of trade-off
- Other possible trade-offs (time, egress time, reliability, costs, crowding) and need to accommodate
- Benefits of realism for SP exercise

RP EVIDENCE (II)



UNIVERSITY OF LEEDS

- 606 RP observations and 5388 SP observations
- Index of perceived bus quality (1-5)
- Unit change in index valued in terms of WALK TIME
 - RP** 1.42 ($\pm 65\%$)
 - SP** 2.67 ($\pm 52\%$)
- t stat for difference = 2.2
- Results not materially different for ASC for quality bus
- SP higher and it is package valued

- Choice of Train Departure for London commuters (RP and SP)
- Fast train but often more crowded, less reliable, older trains
- Good choice context for RP and SP valuations
- Train type rated on 10 point scale
- RP

2348 observations

Unit rating = $0.0032T(\pm 26\%)$

Mean difference in rating = 5% journey time change

- SP

7047 observations

Unit rating = $0.00164T (\pm 20\%)$

Mean difference in rating = 2.1% journey time change!!!

The purpose of the study was not evidently rolling stock

DEMAND IMPACT (I)



UNIVERSITY OF LEEDS

- 1990 Lancashire New Stations Study
- Based on 10 recently refurbished stations
- Improvements – surfaces improved, waiting/shelter facilities provided, general repairs
- SP – time, cost and the actual improvement package
- Local journeys dominate
- Value of improvement – 11% of return fare (7 minutes)
- Should detect an effect on demand
- No discernible difference between demand changes at refurbished and unrefurbished stations

Bus Soft Factors Study (2009) for Department of Transport

- Case study of 10 operators where quality bus (QB) introduced
- Range of Improvements (always with new buses)
- Comparison with other routes/previous year

Arrested decline in demand

Generally small growth (but confounded with improvements in hard variables)

- The RP evidence indicates relatively small effect
- BUT the SP evidence would imply a large effect

DEMAND IMPACT (III)



UNIVERSITY OF LEEDS

Attribute	Value in mins	
	Bus User	Car User
Audio Announcements	1.22	
CCTV at Bus Stops	3.70	2.49
CCTV on Buses	1.66	3.18
Climate Control	1.24	
New Bus Shelters	1.08	
New Bus with Low Floor	1.19	2.23
New Interchange Facilities	1.27	
On-Screen Displays	1.90	0.89
Real Time Information	1.47	1.74
Simplified Ticketing	0.84	2.06
Trained Drivers	2.46	2.78

No package effect

Average journey time 20 minutes

DEMAND IMPACT (IV)



UNIVERSITY OF LEEDS

- Same study conducted mode choice SP on PACKAGE
- Grouped data to estimate 'Direct Demand' models
- 1146 bus users and 729 observations in model
- Time values of bus improvements entered (**no package effect**)

Bus Fare Elasticity: -0.70 ($\pm 18\%$)

Bus Time Elasticity: -0.22 ($\pm 57\%$)

Bus Headway Elasticity: -0.11 ($\pm 38\%$)

Bus Average Lateness Elasticity: -0.05 ($\pm 59\%$)

Car Time and Cost Cross Elasticities – not significant

Quality Bus Term -0.015 ($\pm 29\%$)

Quality Bus valued at 9.4 minutes - removal forecast to reduce demand by 13%!

- Just because other elasticities plausible DOES NOT mean that the soft factors demand effect is plausible



- Econometric Analysis of introduction of new trains
- Even though large data sets and good quality ticket sales data, rare new train effects significant – **despite high SP values**
- Significant effects for new Leeds - London trains
- No timetable changes but demand increased:
 - To London +4% ($\pm 2\%$) From London +3% ($\pm 2\%$)*
- *Prior SP study of these new trains valued at 10% of fare*
- *Deemed reasonable at the time*
- *But price elasticity of -1 would imply larger demand effect*
- *Consistent with findings below that 3 times too high*

Wardman and Whelan (2001) Review of Rolling Stock Values

- 40 values (expressed as % of fare for new trains)
- Regression to understand variations
- Largest effect – values 3.0 times higher when purpose of study was transparently valuing rolling stock
- Consistent with findings on demand impact versus valuations

Lu et al. (2008) Study of Improved Local Rail Services

- Used 'Cheap talk'
- Originates in experimental economics/psychology
- Essentially warning script to reduce incentives to bias

“Previous surveys have sometimes found that people say they would be happy to pay extra for improved trains but when the fare is raised and the improved trains are provided, people say they would prefer the cheaper fare with the old trains. Bearing this in mind, as you read through the following choices, please imagine you would actually have to pay the fare stated”

- This could limit biasing towards new trains – but could itself cause strategic bias towards fare!

OTHER EVIDENCE (III)



UNIVERSITY OF LEEDS

- Unlabelled Pairwise SP exercise (Train A and Train B)
- Time, cost, headway, crowding, punctuality, new train
- 1222 individuals and 10885 observations
- Significant effects from whether respondent received Cheap Talk

Cost coefficient increased $\approx 20\%$ (respondent reminded of cost increase)

Time coefficient increased and crowding coefficient fell

No effect on the rolling stock term

SUMMARY



UNIVERSITY OF LEEDS

- No doubt lots of evidence I've not been able to cover
- SP evidence too high even where package
- If such high values, would expect to pick up demand effects
- Is the assumed decision making process valid - soft factors nice to have but not fundamental drivers of demand?
- Transparency provides incentive to bias responses
- If we imply something is important, we recover it as important!
- More testing interactions

- Lot of advances design, modelling, data collection.
- Also advances in how interpret responses (difficulty, relevance, ignoring). But what about incentive to bias which was the initial concern about SP?
- More concerted efforts to identify RP contexts
- More validation against implied demand response (widely done in using SP for forecasting)
- Less transparent studies
- Debriefing questions (perceived purpose, attitudes) but not cheap talk