

Reliability of SIMALTO Modelling Forecasts **With Sample Size**

John Green, Research For Today Ltd

A county council undertook a SIMALTO Modelling exercise in 2005. The total sample was 500. This scale of investigation allowed for comparisons between 4 separate geographic areas, as well as provision of the total county residents' priorities.

In order to test if smaller sample sizes would suffice for the total survey, the resulting data was split into the first 250 respondents interviewed, then by adding the next 100 and finally using the full 500 sample. The first 3 columns of the table below show the optimum budget allocation of these 3 samples (i.e. how residents would prefer to reallocate the current budget spent on these services so the reallocation had as many of their top priorities and as few of their low priorities as possible). All had exactly the same 7 changes to the status quo, 4 savings in grey 'paying for' 3 improvements in yellow. This shows that had a sample of only 250 been chosen for the survey, it would have given the same optimum budget allocation forecast as the total sample of 500.

In 2009 the county repeated the SIMALTO Modelling survey. 6 of the 22 services changed or were omitted, some to be replaced by other services. But the other 16 remained substantially the same. The fourth column shows the optimum allocation of this 2009 sample where only 400 interviews were conducted in total. The results are similar. Two of the four 2005 savings were repeated and the other two service areas were not included in the 2009 study. Of the two extra savings identified in 2009, both these savings would revert to the current 'as now' allocation (as in 2005) if the total council 2009 budget could be increased by only a very marginal percentage – they were the next items to be improved from the column 4 budget allocation.

Sample size	250 (2005)	350 (2005)	500 (2005)	400 (2009)
1 Youth services centres	As now	As now	As now	As now
2 Road maintenance	As now	As now	As now	As now
3 Pavement maintenance	As now	As now	As now	As now
4 Elderly live at home	5% more	5% more	5% more	350 more
5 Elderly waiting list	As now	As now	As now	-
6 Transport subsidy	As now	As now	As now	Reduce
7 School transport subsidy	Half cost	Half cost	Half cost	Half cost
8 Countryside access	As now	As now	As now	25% less
9 Library service opening	As now	As now	As now	As now
10 Library books	As now	As now	As now	As now
11 Road safety campaigns	2 more	2 more	2 more	As now
12 Rights of way	As now	As now	As now	As now
13 Rogue traders	As now	As now	As now	-
14 Disabled support	As now	As now	As now	As now
15 Community learning	As now	As now	As now	As now
16 A S B assistance	Double	Double	Double	-
17 Voluntary grants	As now	As now	As now	As now
18 Arts, culture grants	50% less	50% less	50% less	10% less
19 Communications	Cease	Cease	Cease	-
20 Access council services	Website	Website	Website	-
21 Amenity sites	As now	As now	As now	As now
22 Fire & rescue services	As now	As now	As now	-

The next table shows the hierarchy of residents 'value preference' index for the right hand member of each pair of adjacent options on the SIMALTO grid, all other things being equal, and recognising the 'better' service level costs a bit more than the 'lesser' option. This is a much more detailed forecast than the overall optimum allocation shown above because the latter is in essence a compilation of these more detailed findings, and so any differences between individual option preferences within the optimum has a greater chance of being 'smoothed out'. Only the 'top half' of the total tabulation is shown - those preferences above 50%.

Again the first 3 columns show the preference for the better option of each change-pair in each of the 250, 350 and 500 samples to be very similar indeed. There is an average difference of less than 2% per service change between the 250 and 500 samples, where the changes themselves vary from 50% to 86%. Thus the county would not have made a different interpretation of relative priorities of discretionary changes on these services if it had used a sample size of only 250 instead of 350 or 500.

Interestingly similar relative preferences from the 2009 study wherever possible show very similar individual service change priorities, indicating robustness of residents' relative priorities over time, and reliability of SIMALTO forecasts.

The question arises as to why is SIMALTO Modelling so consistent when 'conventional statistical wisdom' suggests samples of 1000 or more are required for this level of accuracy. The answer lies in the nature of the data. If one asks a simple question – should the council do this or that – or asks on some scale the satisfaction of residents with a given service – then indeed samples of 1000 are required to be sure (i.e. only to have a chance of being wrong one in twenty times) if a difference of 2% between two samples is a real difference or just a facet of 'natural variation' of the question response. SIMALTO data, by its nature, is much more reliable. Firstly the preference for inclusion of a given option is asked four times – so if a respondent makes a 'mistake' once, he/she has three chances to 'correct' it. Secondly the nature of the question provokes more thought on the part of the respondent. It is not a 'top of the mind' response but a considered choice between this and that. More thought is needed to answer the question.

And finally the question is well benchmarked. Response to typical satisfaction scales can reflect the rather 'woolly' nature of the question and the scale provided for the response - 'are you satisfied with the council's performance on street cleaning? - with a 5 or 10 point tick-box choice answer' A SIMALTO question is more likely to be 'do you want the streets cleaned every 8 weeks rather than every 10 weeks, and if so are you prepared to pay a little more to get this benefit?' – a much tighter question demanding a more thoughtful response.

Service	Change	250 (2005)	350 (2005)	500 (2005)	400 (2009)
16 ASB assistance	Cease → As now	86	87	86	-
2 Road maintenance	10% less → 5% less	82	82	80	86
3 Pavement maintenance	10% less → 5% less	80	80	78	77
3 Pavement maintenance	5% less → as now	74	76	74	71
2 Road maintenance	5% less → as now	73	74	72	73
17 Voluntary grants	50% less → 20% less	70	71	72	71
1 Youth centres	10 less → 6 less	71	72	72	71
7 School transport	Close → pay full cost	67	71	71	-
14 Disabled support	Close → as now	71	71	71	68
1 Youth centres	6 less → 4 less	70	71	71	68
7 School transport	Pay full cost → half cost	67	71	71	-
17 Voluntary grants	20% less → as now	65	66	68	67
15 Community learning	Cease → as now ²⁰	64	66	67	-
6 Transport subsidies	Reduce → no w/e	66	67	66	62
1 Youth centres	4 less → as now	63	65	65	57
9 Library service	Reduce → as now	62	64	66	69
4 Elderly live at home	As now → 5% more	67	66	66	68
8 Countryside access	50% less → 25% less	60	63	64	61
16 ASB assistance	As now → double	61	64	63	-
10 Library books	10% less → as now	61	63	62	57
22 Fire stations	Fewer → as now	62	62	61	-
8 Countryside access	25% less → as now	53	57	58	49
11 Road safety campaigns	As now → 2 more	56	57	59	66
6 Transport subsidies	No w/e → as now	57	58	57	54
5 Elderly waiting list	As now → halve	51	50	51	-
20 Access council services	Reduce → encourage web	50	51	49	-

As a final check on SIMALTO forecasting reliability of real life alternatives, 6 different part budgets are described to survey respondents and they were asked their preferences between them. This attempts to reflect preference between 6 actual budget allocations the council could introduce. Of the 6 choices available, A to F, respondents were told budget allocations E and F are slightly more expensive than A to D (in terms of council tax changes required to fund them).

The choices made in the interview and the simulated preferences later calculated from the modelling of the SIMALTO data response is compared below for the 2005 and 2009 surveys.

2005 sample of 500	A	B	C	D	E	F
Questionnaire Preference %	11	10	14	9	23	33
Model Simulation Preference %	10	15	14	10	22	30

2009 sample of 400	A	B	C	D	E	F
Questionnaire Preference %	16	19	10	12	17	35
Model Simulation Preference %	18	14	11	12	15	31

The above similarity between the ‘true’ preferences from the questionnaire, and the simulated ones from the SIMALTO sheet data, gives confidence that the model will correctly pick winners from losers in any alternative budget comparisons the user may desire to make using the SIMALTO sheet data.

The services included in these ‘part budgets’ differed between the 2005 and 2009 surveys and no respondents that took part in the first survey took part in the second. These samples were from a total population of over 1 million, but were quota balanced by age, gender and geographic area.