Annex 3: Paper B

Research & Insight Working Group

Technical paper supporting Business Plan response

Willingness to Pay, Business Options Testing and Acceptability Testing – what difference did it make?

This paper[1] looks to answer three questions in relation to the quantitative aspects of Willingness to Pay (WtP), Business Options Testing (BOT) and Acceptability Testing (AT) engagement programmes:

- Did WtP, BOT and AT make a telling difference to the Cadent Business Plan?
- Can we see how the way that the research has been constructed leads really clearly and effectively to the choices that Cadent has made?
- What is our view of CEG impact on Cadent's approach to WtP, BOT and AT in particular?

All three types of testing were basically statistically sound and representative – they were large samples. In general, we do not have much comment in relation to the standard survey design and execution issues (sample size etc) – this is covered by Savanta's Assurance paper (Appendix 05.06). This paper is more about the impact of these studies on the Business Plan. One overarching comment, however relates to regionality. All sample sizes were big enough to generate results on a network by network basis. However, while regional variation has impacted the plan in some areas it is not generally prominent as a factor.

Our overarching conclusion was that all this testing was well managed and executed and had a good degree of impact on the plan, with WtP having less impact than the other two types of test. It has fed into cost benefit analyses (CBA's). WtP also (apparently, but not evidenced) informed the development by Cadent of commitment options tested during BOT. In all cases a more thoughtful and strategic approach could have paid dividends in terms of design and impact with this point being most strongly made in relation to WtP. This reflected the fact that Cadent were seeking to turn the handle on a process without having a very clear idea of what they were trying to achieve i.e what it was potentially seeking to propose in the business plan and therefore what it should test. This is probably an inherent weakness in most if not all utility price review engagement programmes, where this sequential order of testing has to be managed significantly ahead of the final Business Plan submission, and may not be specific to Cadent.

[1]Created by Martin Silcock, CEG R&IWG member; final edit and publication Leslie Sopp, Chair R&IWG

Cadent's expertise also grew over the period with the result that it became more competent at thinking through what it wanted and more willing to court and accept challenge from CEG.

Nonetheless this should not take away from the fact that the outcomes have generally been dealt with reasonably (in "triangulation") albeit with some contortion where earlier testing proved to be mis-targeted. The approach to triangulation became more structured and detailed later on in the process.

What follows are some thoughts on each phase of quantitative testing.

Willingness to Pay

A telling difference?

The primary output of WtP is a value per change in a service attribute which gives (part of) the value which can be attributed to the provision of more or less of that service. It is held to include both values placed directly on services by customers (private value) – derived from stated or revealed preference studies - but also elements of social value from benefits transfer studies e.g. carbon costs (public value). This enables the filling in of the benefits side of a CBA (it can also provide a value for non-financial costs (eg disruption) to go on the costs side).

This is clearly a rather technical interpretation of what WtP is: it incorporates information which is not directly derived from customers and where it does directly derive from customers it does so using a method which seeks to remove some of the potential sources of bias which are prevalent when asking a customer "are you willing to pay for this?"

Business plans frequently refer to "what customers are willing to pay for". Quantitative WtP research does not, by itself give a direct answer to that question but it does give a robust take on part of the answer.

Whether WTP makes a telling difference depends on (i) whether the survey outcomes have been actually used and used properly in calculating a CBA and (ii) whether that CBA has then made a difference either because it has affected an investment decision at the margin or because it is the sole or main rationale for proposing an investment.

CEG has become broadly comfortable on (i). CBA's are being used. There is a CBA methodology (which we have confidence in) being used properly for large amounts of the investment programme, particularly noting that WtP for interruptions and carbon costs have been used in the repex and asset health modelling. Cadent also obtained third-party assurance of its CBA by NERA[2] and has addressed the findings of that review.

It is worth noting that for some investments and CVPs an alternative valuation technique based on Social Return on Investment was used. WtP has been used to calculate the benefits relating to three of Cadent's proposed CVPs.

[2]NERA is the consultancy contracted by Cadent to develop the WtP model – framing the questions, and delivering the outputs and reporting. The management of the data collection was handled by other contractors.

Has WtP made a difference? Yes, in the sense that there are specific areas of the business plan in which CBA informed by WtP research is a factor in the determining the business plan proposals. The specific areas are:

Service attribute in WtP	Impacts on investment in (H/M/L/None)
Interruptions – 3 to 24 hours	Repex and other asset health investment. (**Medium**) WtP is used to calculate CBA either via the AIM model which optimises repex and selected other asset classes (Cadent's extension to Ofgem's NARMS model) or directly into a CBA model. However, interruptions are not a large source of value cf safety and environment. In addition, only a small proportion of the repex programme is CBA driven. MOBS. WtP is used to calculate the CBA of the proposed investment, which is positive, although it is pipeline safety integrity that primarily drives the chosen option. Note that the actual values for short interruptions used were lower-bound values derived from benefits transfer rather than the much higher values from stated preference studies, a prudent approach which NERA supported.
Interruptions - >24 hours	Repex/asset health. (**Medium**) This attribute has been used to calculate CBA as for short-term interruptions either via AIMS or directly into a CBA. MOBS. WtP is used to calculate the CBA of the proposed investment, which is positive, although it is pipeline safety integrity that primarily drives the chosen option. Note that values used have in some cases been derived from stated preference where the benefits of investment relate to the avoidance of very long interruptions. In other cases, lower-bound benefits transfer values have been used.
Duration of short interruptions	Reducing Average Length of short interruptions (**Low**) In theory this could have impacted investment to improve service as there is some WtP for lower levels of average interruptions duration. However, the BOT results demonstrated low ambition to move beyond the status quo and so no there is no proposal to invest.
Leakage reduction	**None** This is the inherent value to customers of shrinkage (not the carbon value). WtP is zero and so has not impacted the plan

Remediating disused sites	**None.** There is no investment associated with this in the plan
Offering timeslots	Timebound Appointments. (**High**) Positive NPV (£24m) against minimal cost contributed alongside strongly positive qualitative feedback towards Cadent including this as a new output commitment. No BOT was conducted on this. The value has also been used in calculating the benefit for the associated CVP.
Reducing number of excavations	**None**. Zero WtP for both domestic and business customers. Therefore it is not a factor in any CBA calculation.
Reducing re-instatement time	Re-instatement time (**None**) Positive WTP from domestic customers so there may be a positive CBA (but it is not exposed). Eventual choice was to maintain current performance at zero cost, based mainly on quantitative and qualitative BOT – most customers voting for the zero cost option. So there is no impact of WtP.
Providing welfare to vulnerable customers	Welfare services to CIVS (**Medium**). WtP is positive as tested but caveated as to applicability of "levels" of service packages tested in WTP versus those offered. Net benefit for the option chosen is quoted at c£17m but not clear whether this is based on SROI or CBA based on WTP. BOT appears to be main driver for choice of most ambitious option. The WtP results have been used in calculating the related CVP.
Measures to improve fuel poverty	Beyond the meter interventions to fuel poor customers. **(Medium)** WtP is positive and the resulting CBA seems strongly positive (£177m for chosen option) – but a more ambitious intervention would have been supported. However, this and strong stakeholder support was moderated by BOT showing least ambitious option was most preferred.
Connecting households in fuel poverty	Fuel poor connections. (**High**) WtP is positive. There is a technical issue in that the range tested in WtP was higher than the eventual target (4000 per annum cf 1250). It is thought that this was because the range was set using past data from the period when eligibility criteria were less stringent. It is mitigated because within the range the marginal benefit is strongly linear (and this is recognised and highlighted by Cadent). BOT did not impact this choice as every choice involved connecting 1250 per year so no other quantitative data was available.

Proportion of gas from green sources	Entry enablement CVP. (**Low**). There is positive WtP by customers to increase the % of total gas from green sources. This has been used to calculate the benefit associated with the CVP for entry capacity enablement for biomethane. (We have concerns about this usage: see our views on the CVP) However, it has not been used in CBA to support output cases. This attribute arguably overlaps with the official valuation of carbon from BEIS. It has been the latter which is used in the calculation of CBA for environmental output cases (carbon neutral ops; lower employee emissions; waste to landfill). It is also the official valuation which has been used in the main repex and asset health models to calculate the CBA – and in all other CBA calcs. No evidence has been seen of this value being triangulated against the values from this attribute.

Overall WtP has had an impact across the plan in some important areas. However much of the testing has ended up not being impactful. In addition, there are proposals in the plan which are not backed by WtP but which arguably could have been (for example CO initiatives and MOBS). Partly this arises because it isn't known what customers value before testing. But in part it is a function of lacking a really sound strategic plan and/or the time to execute one.

Looking at the details of the triangulation as expressed suggests that the tension between WtP and BOT has not always been fully spelt out, with the outcomes of WtP potentially being under-played relative to BOT. This could be because BOT is more accessible and also because it presented choices which were by then better formed.

Has the way it was constructed led clearly and effectively to choices that Cadent has made?

Up to a point. It has been successfully deployed, in the rather mechanical way in which WtP works. However as can be seen from the above, WtP has not perhaps had the really pervasive impact that it might have done, mainly because it needed to be scoped and initiated well ahead of the time when Cadent had started to think coherently about the specific plan options that it wanted to explore. This was a result of time compression which in turn resulted from not having a really joined up engagement plan. Having said this (and as has been noted earlier) WtP always has a long lead time so it's not an uncommon issue to encounter in regulatory business plan development programmes of this nature.

The execution was competent, and the approach was standard in other utility markets, but the first time that Cadent (or any GDN) had used for revealed preference. We challenged Cadent on the method and were encouraged by the competence of NERA on this. There was also an external independent report by a well-regarded academic in the field (Ken Willis). As this was the first experience for Cadent of this type of work it was perhaps understandable that it appeared to struggle to engage with some of the more technical aspects of the process, particularly at the outset. It was disappointing that a number of innovative ideas

trailed early were not taken forward. On the other hand, it was positive that some revealed preference work around customers' response to disruption was attempted.

The selection of attributes to test seemed sensible and backed up by customer prioritisation. CEG challenged effectively on this. The ranges over which WtP was tested for each service attribute were related to the actual and likely capability of Cadent to deliver those attributes. However, at the time they were selected there was very little evidence that Cadent had very well-developed ideas about its business ambition in relation each attribute (as evidenced by the example of fuel poor connections above). Nonetheless, at the time they seemed a sensible set of attributes tested at sensible levels designed to cover whatever business plan was likely to emerge.

In our opinion, there was an execution weakness in the qualitative materials developed by Cadent's engagement partner which were not particularly innovative and didn't look very compelling. However, cognitive interviews and piloting along with detailed review enabled weaknesses to be corrected.

What was CEG's impact on engagement activity?

It was disappointing that Cadent did not see fit to involve CEG early enough in the process of development and cognitive testing of materials thus reducing our potential impact. This became a common theme of the engagement programme despite multiple attempts to re-direct the Cadent engagement team in this regard. Attempts to "educate" Cadent to a more robust use of WtP were somewhat successful. We raised the need for an expert review of its usage of WtP in CBA calculations which resulted in Cadent commissioning NERA to perform a review, albeit late in the process.

Business Options Testing (BOT)

A telling difference?

It is much clearer that the BOT that was undertaken was directly impactful on the investment proposals in Cadent's business plan. BOT was both quantitative (a large statistically significant survey) and qualitative with a large number of exploratory groups and workshops. CEG do not comment here on the quality and robustness of the qualitative research undertaken. However, there are occasions when in triangulation reference has been made to the quantitative outcomes of surveys of customers attending qualitative workshops without much caveat that these are not statistically robust. There is nothing wrong in quoting them but on occasion they are over-weighted.

One striking example of this is the "balanced benefits" survey which has driven the choice of options applied to repex and other modelling. This choice was surveyed at a small number of workshops and results were that "most" customers preferred balanced benefits. But it was not surveyed in the large quantitative survey (and could have been).

The table (on pages 6 & 7) are all of the choices offered and assessment of impact in the same way as for WtP earlier.

Tested in BOT	Impacts on investment in (H/M/L/None)
Carbon monoxide safety	Output commitments on CO education, alarms, partnership.(**High**). Strong support for the most ambitious option tested in BOT, backing up qualitative research. No WtP but SROI calculated.
Responding to CO incidents	Appliance isolations (**Low**). While BOT was strongly in favour of the most ambitious offering, subsequent work cast doubt on the scale and deliverability of the commitment and it was removed from the plan. This appears to cast doubt on how well considered BP options were prior to BOT. No WtP and no CBA calculated.
CO repair and replace faulty appliances	Appliance repair/replace. **(High)**. BOT strongly supported the most ambitious option which was then included in the BP albeit at slightly higher cost than exposed in testing. No WtP but a positive CBA was calculated based on SROI.
Helping vulnerable customers without gas	Welfare services to CIVS. (**High**) BOT was strongly in favour of the most ambitious welfare package for CIVS and this supports the proposed option to provide an ambitious level of welfare to CIVS.
Helping vulnerable customers without gas	Welfare services to all customers. (**High**). BOT showed no strong support for extending welfare provisions to non-vulnerable customers so this commitment was excluded. Specific eligibility criteria for qualifying for welfare were developed.
Getting customers back on gas	Reducing average length of interruptions. (**Low**). BOT only tested improvements on current service levels (a narrower range cf to that covered in WtP and it is not clear what drove this test design) but this showed no strong support for an improvement. Therefore, no commitment to reduce interruptions was factored into the plan beyond what can be achieved at no incremental cost to customers. Positive WtP was not telling (see above).
Carrying out safety checks	Pro-active safety checks (**High**). The proposed commitments was removed based on lack of support in BOT for more ambitious options, backed up by qualitative research.
Minimising disruption	Improving reinstatement times (**High**). BOT showed strong support for retaining the current level of performance on re-instatement time and this was influential in stepping down the

	ambition that had arisen after earlier qualitative and WtP work (See above on WtP).
Tackling Fuel Poverty	Beyond the meter interventions (**Medium**). BOT tested packages of advice and whole house interventions. Results favouring the least ambitious option drove a moderation of ambition to the "middle" option, in contrast to qualitative research and WtP which would have supported more ambition. The actual number of connections was not tested by BOT.
PSR awareness	Conversations and partnerships (**Low)**. BOT suggested the least ambitious option for raising PSR awareness. Stakeholder feedback and subsequent qualitative feedback (and non-significant quant surveys of workshop attendees) led to Cadent proposing a commitment with maximum ambition on conversations and partnerships but low ambition on "innovation". The implication was that the BOT packages should have been disaggregated (but there may have been more detailed conversations later).
PSR Training	Training our people. (**High**) BOT revealed preference for moderating the ambition on training to customer-facing staff and this fed through to the output commitment.
Becoming carbon neutral	Carbon neutral business (**Low**). Despite BOT suggesting that the least ambitious option should be chosen, Cadent chose to go for the most ambitious and hence show leadership – in line with some customers' views. Qualitative engagement showed customers struggled with some of the concepts on carbon neutral, suggesting a better survey design could have elicited more useful information.
Communities not on gas	Off-gas communities. (**High**). BOT showed that there was majority support for some trialling of this (even though the most popular choice was not to trial), in contrast to the views of some stakeholders. This led to Cadent proposing the lowest level of trialling.
Keeping the energy flowing reliably and safety	Repex programme (**Medium**) BOT tested whether customers would prefer Cadent to replace more km of main and hence deliver more and earlier benefits for slightly higher bills. Results were split but with a majority in favour of some additional investment, Cadent proposes a middle level of CBA-driven repex: an additional 50km p.a. cf the most ambitious option of 100km p.a. (which would also have been cost

	beneficial) but against a total repex programme of over 1500km p.a.

The above shows that BOT did appear to have a substantive impact on the plan with it confirming or restraining ambition in several areas. However, it was also impacted a little by design issues in that a more thought-through approach may have resulted in more useful questions, and indeed testing of different elements of the plan. It is noteworthy also that a second, apparently innovative, element of the quantitative BOT test, the "maxdiff" analysis, which comes up with a ranking of the attributes tested (from the same responses) does not appear to have been used at all in triangulating the outcomes.

Has the way it was constructed led clearly and effectively to choices that Cadent has made?

It can be seen from the above that BOT has been impactful and on balance it has turned out to be an important driver of the business plan. Moreover, Cadent have taken account of it and have recognised the need to triangulate with other sources and have done this reasonably. However, a better design would have been possible at the time that it was designed, given more careful thought about the insight gained at that time and what Cadent actually intended to propose in the plan.

CEG has at times been uncomfortable with Cadent's own understanding of the methodological rationale for this quantitative work and particularly how it fits with the WtP. It is important to note that while BOT choices were associated with specific bill impacts, they do not provide direct evidence of willingness to pay – they are simply expressions of preference between options in which the bill might be one, but not the only, factor. This is valuable information. However, there was often no choice which left the bill unchanged and BOT does not have an experimental design which aims to eliminate the well-known biases which arise when customers are asked what they would pay for hypothetically. There are reasons why the two methods can complement one another but these have not been drawn out well in any documentation that has been seen in preparing this paper. Repeated requests to Cadent to provide such explanations were generally unsuccessful. However, in the end the body of evidence produced by BOT has been useful and if nothing else has in itself prompted a greater degree of thought about what customers actually want and what Cadent wants to offer them.

What was CEG's impact on engagement activity?

Despite attempts to become involved at early stages in design, CEG were unable to contribute as much as they might have liked. Cadent clarifying their thinking earlier on could have made this more effective. We have had more exposure to the latter stages and particularly to the triangulation of results: our attendance at one key meeting was helpful and re-assuring and resulted in Cadent making improvements. Earlier reluctance to engage may have been due to lack of confidence about the proposed approach which was only really assuaged once sensible results started to come in.

Acceptability testing

A telling difference?

The last major tranche of quantitative work was in some ways the easiest to grasp in terms of its impact. The design of the survey was reasonable and CEG was able to be involved from an earlier stage.

Essentially quantitative acceptability testing resulted in a positive outcome across the plan in that what was proposed was generally acceptable to customers both on an informed and an uninformed basis. The quantitative survey was disaggregated in the sense it asked specific questions separately about the three key outcome areas, all of which were seen as largely acceptable.

The outcomes of quantitative acceptability testing have been used by Cadent to offer substantive support for its plan, and it has particularly relied upon it when it is seeking to support elements of the plan for which there is a lack of prior engagement or where there has been conflict. In this sense it has been telling.

Has the way it was constructed led clearly and effectively to choices that Cadent has made?

It has confirmed choices rather than led to them. No areas have been identified where acceptability testing has resulted in a substantive change.

Although much was expected of acceptability testing, going into it, to pick up specific gaps or outstanding questions, in fact the quantitative work was framed very generally. So, for example, we were told that engagement on steel pipes was to be covered but we did not see this in practice in qualitative or quantitative testing. In addition, customers in qualitative workshops raised concerns about concluding as to acceptability of network resilience proposals (e.g. km of main per annum in repex) due to lack of framing, but this was picked up as feedback and addressed for the remaining qualitative fieldwork. This may have also influenced quantitative results, although it is hard to say which way any bias might have worked.

One notable aspect of the test design was Implicit Acceptance testing, in which customers are asked to give intuitive impressions of the business plan by choosing from a series of word-pairs describing the plan on various dimensions (e.g. innovative/old-fashioned, straightforward/complicated, etc) This was interesting, but it was hard to see how it would inform the plan and in practice does not appear to have been used.

CEG has specific concerns about the qualitative testing on risk and uncertainty (which are covered in the Chapter on Uncertainty Mechanisms).

Having said this we conclude that the survey was done reasonably, and it was good to have the opportunity to input and influence the approach at an early stage.

What was CEG's impact on engagement activity?

CEG was reasonably well involved and at an earlier stage then for the other testing. CEG members critiqued and made suggestions to the quantitative survey instrument which were generally acted on (e.g. on bill presentation). CEG members attended cognitive testing and a

number of qualitative engagement events and made suggestions for improvement which were generally acted on.

Overall summary

All three programmes were executed to largely standard utility pricing, options and acceptability testing protocols. WtP had three components which made it stand out; BOT utilised a Max-Diff approach, and AT an Implicit Association testing component, neither of which appear to have been used.

It is possible that improvements could have been made to the WtP research as long as they could be keyed directly into the final plan proposals. Cadent (and their partners) could have been more engaged and appreciative of the issues in framing qualitative and quantitative studies for these purposes from a conceptual through to actual use perspective (rather than the focus on process, design and management), and involved CEG technical specialists earlier.

The tension between WtP and BOT has not always been fully spelt out, with the outcomes of WtP potentially being under-played relative to BOT. This could be because BOT is more accessible and also because it presented choices which were by then better formed. These issues do not appear to have been reconciled effectively by the business owners, nor checked or commented on by Cadent and partners.

There are specific areas of the business plan in which CBA informed by WtP research is a factor in the determining the business plan proposals

It is much clearer that the BOT that was undertaken was directly impactful on the investment proposals in Cadent's business plan.

Both WtP and BOT however, were impacted by design issues in that a more thought-through approach may have resulted in more useful questions, and indeed testing of different elements of the plan.

The outcomes of quantitative acceptability testing have been used by Cadent to offer substantive support for its plan, and it has particularly relied upon it when it is seeking to support elements of the plan which are for reasons of lack of prior engagement or where there has been conflict. In this sense it has been telling.