FROM INSIGHT TO INFLUENCE

PR24 Triangulation

(Phase 1 Methodology)

Final Report

Prepared for SSC

Prepared by Impact Research

9 January 2023

Project No. 1345

South Staffs Water

Cambridge Water

Table of contents

Glossary

Industry Terms

AMP7	Seventh Asset Management Period
BAU	Business As Usual
CAC	Cost Adjustment Claim
CAPEX	Capital expenditure
CBA	Cost Benefit Analysis
CCG	Customer Challenge Group
CMEX	Customer measure of experience
CRI	Compliance Risk Index
CTS	Cost to Serve
DMEX	Developer services measure of experience
ERI	Event Risk Index
ESG	Environmental Social Governance
ICG	Independent Challenge Group
IIP	Investors in People
LTDS	Long Term Delivery Strategies
LTF	Looking to the Future
MCDA	Multi-Criteria Decision Analysis
ODI	Outcome delivery incentive
PC	Performance Commitment
PCC	Per Capita Consumption
PCD	Price Control Deliverables
PIC	Public Interest Commitment
PR19	Price Review 2019
PR24	Price Review 2024
WASC	Water and Sewerage Company
WOC	Water Only Company
WRMP	Water Resource Management Plan

Organisations

CAM	Cambridge region
CCW	The Consumer Council for Water
OfWat	The Water Services Regulation Authority
SSC	South Staffs Water & Cambridge Water
SSW	South Staffs Water

Customer Engagement

DCE	Discrete choice Exercise
CVM	Contingency Value Measurement
HH	Household
H2Online	SSC panel
NHH	Non-household
RP	Revealed Preference
SP	Stated Preference
WTP	Willingness to Pay

1. BACKGROUND

Aims

Impact Research was commissioned by SSC to provide the following:

- To deliver a robust triangulation of customers' and stakeholders' priorities that underpins the narrative of SSC's plans
  - Robustly triangulate evidence relating to WRMP to support all key decisions
  - Support the development of SSC's Performance Commitment (PC) package
  - Triangulate WTP values to set central, upper and lower values
- To support the development of SSC plans with triangulated valuations and insights to best deliver 'public value'
- Create an insight matrix from SSC's trackers to assist in the delivery of the PR19 plan and guide PR24
- Enable both SSC challenge panels and board to effectively challenge the approach plus independent review by a third-party expert

In the pursuit of these wide-ranging objectives, our approach will aim to:

- Create a resource that functions effectively as a data interrogation tool that will support SSC throughout the PR24 process and onwards to BAU activities that will follow submission of the plans.
- be transparent to all parties
- allow substantial sensitivity testing
- follow the SIA/CCW recommendations and Ofwat requirements for triangulation as laid out in their latest position paper.

There are two key areas where SSC required support:

- **WRMP and PR24 priorities:** A thematic review of all relevant material reported separately from this document[1].
- **Willingness to Pay (WTP):** Developing a robust and proportionate evidence base on customers' WTP for different areas of investment. The triangulated values are to be used within SSC's investment optimiser tool to undertake Cost Benefit Analysis of investment options and as part of the process of setting ODI rates. **This is the focus of this methodological report**. The calculation and delivery of the WTP values based on this approach is reported in '**SSC09 PR24 Technical triangulation – Phase 2 Results**'.

The best practice framework

SSC has committed to the over-arching recommendations of the triangulation framework put forward by CCW's extensive review of PR19 triangulation work[2], the essential features of their recommended best practice for triangulation are as follows:

Table 1.1: CCW/SIA Framework

2021 Grouping	Key activities
A strategic approach to collecting customer evidence	• Undertaking a phased and iterative approach • Developing a consistent and transparent decision framework • Putting in place assurance of the process • Linking Business as Usual (BAU) insight to strategic goals
Collecting, collating and synthesising customer evidence	• A centralised process within the company • Capturing relevant granular metadata for insight
Weighting and combining customer evidence	• Transparent approach • Use of a standard approach • A clear approach to demonstrating balanced decisions • Defined decision-making framework
Validating outputs	• Using multi-factor validation (internal, external and independent review) • Running sensitivity and scenario testing • Making research findings publicly available • Independent review of the triangulation process
Incorporating validated findings into the decisions	• The key enabler at this final stage is the use of a robust and transparent decision framework

[1] **SSC11 Thematic Analysis report**

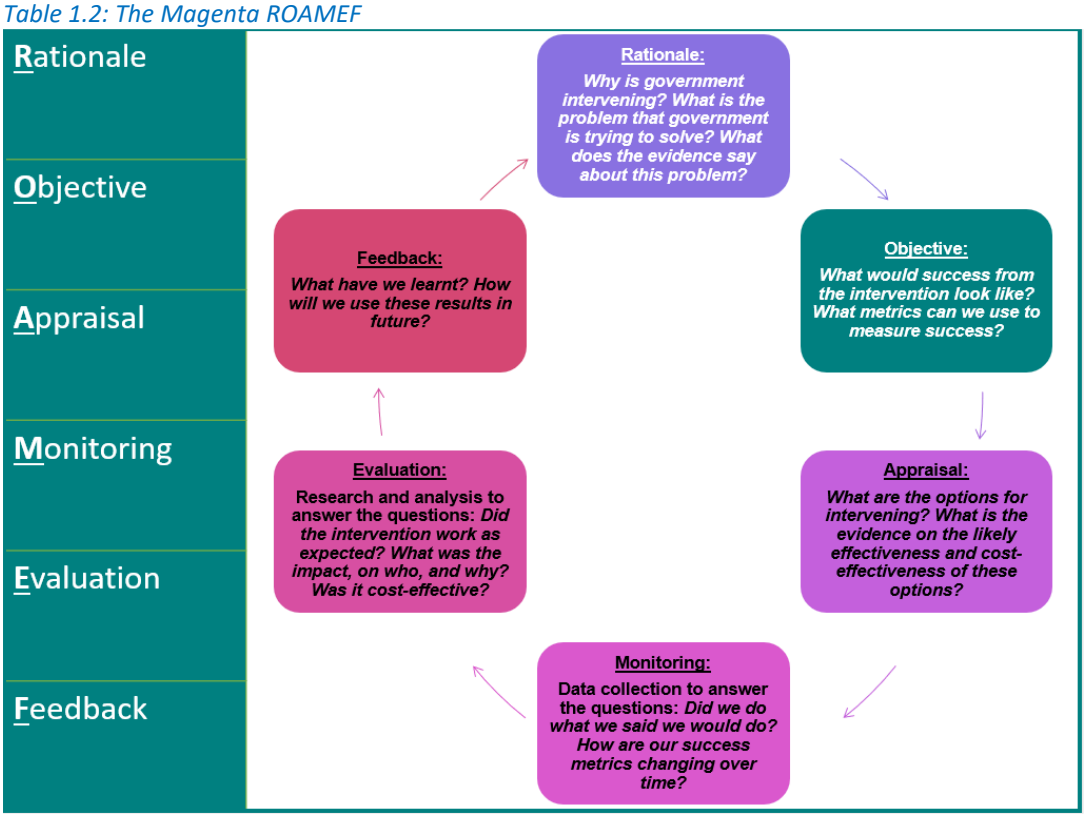
[2] CCW and SIA Partners, April 2021, TRIANGULATION- A REVIEW OF ITS USE AT PR19 AND GOOD PRACTICE

Inputs will also be sought from other sources, such as the ROAMEF structure and ideas on 'theories of change' promoted by the Government's Magenta book[3]. SSC has also been using the UKWIR study[4] and Sustainability First PIN model[5] to help shape the priorities and WRMP local insight programmes, so we will also look to ensure alignment with what SSC are doing in this area and for the triangulation approach to align to these reports, where relevant and appropriate.

The Government's Magenta Book

The Magenta provides a broad structure for evaluating policies / schemes (ROAMEF – see diagram below), of which the (A)ppraisal and (E)valuation elements are the most relevant to the triangulation process. It identifies three types of evaluation: Process, Impact and Value for Money (VfM), of which VfM is the most relevant to the current triangulation. VfM is a key driver of overall satisfaction in SSC's Promises Tracker (customer experience) sat and links closely to the pressing issue of affordability, where Key Driver analysis (KDA) has shown that affordability is by some distance the strongest driver of the customers' VfM score.

Table 1.2: The Magenta ROAMEF



With this in mind, we shall analyse SSC's satisfaction data to examine how measures of VfM and other key metrics have developed since PR19, to see if there is evidence of positive impacts that can be related to major plan initiatives that were developed in part from the WTP / CBA analysis conducted at the time. For example, low pressure and the taste and smell of water have been shown to be key drivers of overall satisfaction and were both included as elements in SSCs PR19 investment plans. To what extent have these changed over time and how much can be attributed to the investment?

Specific monetary-equivalent values can also occasionally be drawn from the Government's 'Green Book', which in our own work for DNOs we have found particularly useful when representing Social ROI (see Appendix 3a and Peer Review comments in Appendix 2).

Another helpful concept in the Magenta Book is a 'Theory of Change', an exercise to help articulate objectives and stress-test potential intervention ideas. Taking the 'Water Hardness' example provided with the project brief, a 'theory of change' can be developed off the back of this. This would put forward the assumption that investment in softer water would benefit customers in terms of lessening the damage to appliances and improving the convenience / offering cost savings to those who currently

[3] HM Treasury, 2020, Magenta Book: Central Government guidance on evaluation
[4] UKWIR, 2021, HOW SHOULD CUSTOMER AND STAKEHOLDER VIEWS BE USED IN REGULATORY DECISIONS? Frontier Economics LTD
[5] Sustainability First, LOOKING TO THE LONG-TERM: hearing the public interest voice in energy and water, eight agendas for change

purchase softening solutions. However, that investment will be most cost effective if it takes the form of advice and localised improvements, rather than a total region softening solution. This is the 'theory' that is then taken forward as a basis for assessing evidence for the value to customers of different softening initiatives.

Other Key Guidance

SSC has also been using the UKWIR study[6] and Sustainability First PIN model[7] to help shape the priorities and WRMP24 local insight programmes, so we will also look to ensure alignment with what SSC are doing in this area and for the triangulation approach to align to these reports, where relevant and appropriate. Summary elements from these studies are given below.

Table 1.3: UKWIR Customer Types

Type of customer	
Transactional	Bill-paying customers
	Non bill-paying customers
	Large users
	Retailers
Inter-generational	Current customers
	Future customers
Societal	Citizens
	Communities

The UKWIR customer definitions emphasise the importance of representing a broad range of customer types and other stakeholders. We note that in addition to these, environmental groups/NGOs are a key stakeholder audience of interest that will be also included.

Much of this relates to the Thematic Review being conducted in a separate report, as purpose-built WTP values are confined to Household and non-household customers. Nevertheless, we will encourage participants in the Delphi Approach to consider how adequately the needs of these different groups are represented in the range of WTP values and to reflect this in their assessments.

Table 1.4: Sustainability First (PIN)

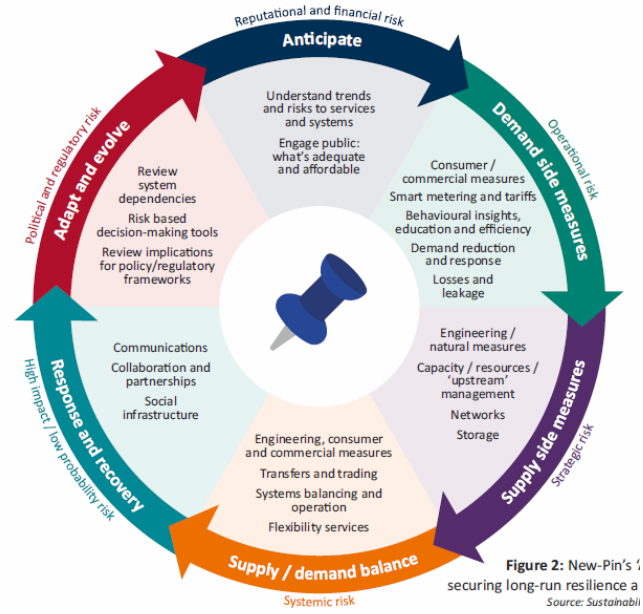


Figure 2: New-Pin's '
securing long-run resilience a
Source: Sustainabili

The PIN model portrays plan development as an ongoing, continuous cycle. The elements that relate most to the WTP triangulation process are 'anticipate' (engage the public), 'supply / demand' balance (consumer measures) and 'adapt and evolve' (risk-based decision-making tools).

The framework we are following for the current triangulation work is part of the 'adapt and evolve' element, contributing as it does to identifying the range of WTP values to consider for use in the investment calculations. This in turn feeds the 'anticipate' and supply / demand balance' elements, representing the consumer presence in the planning process from the outset.

[6] UKWIR, 2021, HOW SHOULD CUSTOMER AND STAKEHOLDER VIEWS BE USED IN REGULATORY DECISIONS? Frontier Economics LTD

[7] Sustainability First, LOOKING TO THE LONG-TERM: hearing the public interest voice in energy and water, eight agendas for change

2. THE TRIANGULATION FRAMEWORK

The main challenges

Taking the five stages of the CCW/SIA framework to structure the discussion, we set out below what we think are the key challenges to address and our suggested approaches to meeting those challenges:

Stage 1: A strategic approach to collecting customer evidence

It is key to recognise that triangulation will always be an ongoing and evolving process ('phased and iterative'). It is not only an approach for drawing together and effectively combining evidence to guide WOC plans, but also for directing subsequent customer engagement over the whole PR24 process and beyond. At all times the approach must be transparent and intuitive to all users. The approach will incorporate regular feedback and oversight from external parties (ICGs, reviewers) and provide SSC with the tools to efficiently incorporate future evidence as it is planned and delivered. An example of this will be a transparent table of weights which can be readily applied to each wave of tracking data and other planned sources of engagement.

Stage 2: Collecting, collating and synthesising customer evidence

All the key summary information will be stored in Excel workbooks using a standardised approach as developed in consultation with SSC. This is an open and flexible format that can be readily adapted as inputs to structured reporting systems, automated to produce standard summary reports in Word, PowerPoint, or PDF, as well as serving as stand-alone data reports in their own right.

Stage 3: Weighting and combining customer evidence into preferences and values

SSC's SMARTS approach developed by Accent and PJM Economics in PR19 provides a sound basis for the triangulation framework. The Red-Amber-Green (RAG) rating system is a pragmatic solution that has the merits transparency and relative simplicity, but is capable of further enhancements, as described in the next section. As with the PR19 approach, the final ratings will rely primarily on 'expert judgment'. With the aim of strengthening this important aspect of the triangulation approach, the 'Delphi method' will be introduced, a long establish method for drawing together expert opinion in a focussed and independent way.

Stage 4: Validating outputs

The different layers of validation (internal challenge, external challenge, and independent review) can all be incorporated in the Delphi approach, where expert contributors will be required to use the RAG model for assessing validity and bias of a range of sources.

Sensitivity testing will also be essential, both in the formation of suitable inputs to Copperleaf investment modelling, as per the previous PR19 work, but also in the testing of a wider range of values in these final analyses, as requested by SSC, for the use of low, central, and high values in the modelling.

Stage 5: Link to business plan

To ensure that the triangulation work plays its full role in enabling SSC to formulate its plans, open and accessible reporting will be essential. This must be implicit not only in the design and explanation of the triangulation approach, but in the intuitive and informative nature of the final outputs presented to all participants in the planning process. To achieve this, the following steps have been included in the process:

- Regular weekly meetings with SSC accompanied by a regularly updated action log
- Meetings with and submission of reports to CCW to update them on the process and invite comment
- Preparation of a Thematic Report in parallel with the triangulation work designed to provide SSC with a comprehensive overview of all customer engagement relevant to the preparation of the business plan

This iterative process ensures that SSC are informed early in the planning process and have opportunities throughout to evaluate their draft plans against the available customer engagement.

3. TECHNICAL TRIANGULATION DETAILS

This section outlines in detail the proposed approach to the PR24 triangulation framework.

SMARTS

We will adopt the SMARTS approach used in PR19:

- SCREEN data sources to identify those with potentially comparable measures
- MAP non-core evidence to core measures where possible to enable comparison
- ASSESS theoretical and statistical validity of the resulting measures
- RATE measures as Red/Amber/Green (RAG) depending on how well they perform with respect to the validity measures
- TRIANGULATE to conclude on the values to take forward based on applying RAG weights to obtain central values and ranges.
- SENSITIVITY TEST the results based on amending the weights to conform with alternative reasonable perspectives.

For PR24, it was believed that the approach could be enhanced in several ways, notably by expanding the 'Assess' part of the process and enhancing the Red-Amber-Green (RAG) approach.

Screen

Studies that contain potential evidence for triangulation will include:

- WRMP
    - Community Research: Findings from the WRAP's (Water Resources Advisory Panel) Theme: Strategic Decisions, 2021
    - Community Research,2021, Findings from the WRAP DEEP DIVES on universal metering and water transfers
    - Community Research,2021, Findings from the WRAP FOCUS GROUPS on options relating to metering, tariffs and water transfers
    - SSC, 2021, Stakeholder Roundtable feedback
    - Explain, 2021-2022, H_2Online - Summary of activities relevant to WRMP engagement
    - Accent/PJM, 2022, MCDA Quantitative Insights
    - Accent/PJM, 2022, SSC WRMP: Themes 1&3 Managing Droughts / Leakage Ambition / Universal Metering / Environmental Ambition
- Tracking Research
    - Turquoise, Report 2021/22, Customer Tracking Research
    - Accent/PJM, 2020 - 2022, Priorities Research Qualitative and Quantitative Insights
    - Customer journey satisfaction (Qualtrics survey)
- WTP
    - NERA, 2022, New Stated Preference research for PR24
    - Review of PR19 outcomes (SSC and other companies)
    - WTP values from the energy sector ('benefit / value transfer' for any translatable attributes, such as water quality[8])
    - National ODI rates study being run centrally across all water companies, led by Ofwat and CCW
- Other BAU insight sources
    - Regular sector benchmarking (customer experience; cross-sector insights, collaboration through WASCs)
    - Contact data from HH customer database covering all channels (online and offline) – complaints, unwanted contacts, wider contact trends
    - Third party supplier/open data sources: e.g., CACI ACORN data, engagement platforms (e.g. Get Water Fit service)
    - Summaries of key points from stakeholder meetings and forums – e.g., MPs, environmental groups
    - Online communities (H_2Online, run by Explain Research)
    - SSC Staff Pulse Surveys
- Other sources
    - There are a range of reports from other water companies and national bodies, such as CCW, which will be added.

In the PR19 approach, triangulation focussed on those sources that offered quantitative values for relevant service attributes. For PR24, we will aim to broaden this approach to make more explicit use of qualitative and more generic studies. This will mainly relate to determining whether the values given from the main quantitative sources are low, high or representative of customer priorities as expressed in the more in-depth qualitative work.

[8] See Appendix 3a

Map

The evidence will be assessed in relation to the specific service attributes under consideration. The final set of service attributes to be triangulated for WTP purposes are as follows:

Table 3.1: Areas to be tested in 2022

SSC LTDS* outcome	SSC's Technical Description of Service (the issue)
Customer Service	To provide excellent levels of service when customers get in touch with queries – by phone, email, online, letter, or face-to-face. In 2021/22 (TEXT SUB: South Staffs Water / Cambridge Water) customer satisfaction was rated 3rd out of all 17 water & sewerage companies in England and Wales.
Risk of a temporary "do not drink" notice	Occasionally, water companies have to send customers a notice saying not to drink the tap water because of an issue with the water quality. Usually this would last about 2-3 days, and (TEXT SUB: South Staffs Water / Cambridge Water) would provide safe drinking water near your property at temporary water stations and would deliver bottled water directly to vulnerable households.
Installing 'smart' water meters	(TEXT SUB: South Staffs Water / Cambridge Water) needs to carefully manage demand for water to ensure there is enough for the future. 'Smart' water meters automatically send regular readings. Having more information helps the water company and customers to understand where and when water is being used or lost to leaks.
Hard water supply	(TEXT SUB: South Staffs Water / Cambridge Water) has a hard water supply. Hard water is not harmful to human health, but it can lead to limescale damage on taps, showerheads and appliances (e.g. washing machines).
Lead pipes	Some properties in your area are served by a lead supply pipe. Most of these pipes are owned by the customer and not your water company. (TEXT SUB: South Staffs Water / Cambridge Water) treats the water supply to ensure lead levels in the water are safe, but there are some circumstances where it can become unsafe (e.g. if lead pipes are badly damaged). Over time, lead exposure can be damaging to health.
Water lost to leakage from pipes	Every day treated water is lost to leakage from the (TEXT SUB: South Staffs Water / Cambridge Water) pipe network as pipes age or are damaged. The majority of the water lost to leaks is from the water company's pipes (70%) and the rest is from customer pipes. The company aims to fix the largest and most disruptive leaks first.
Issues with tap water colour, taste, or smell	Every year, some (TEXT SUB: South Staffs Water / Cambridge Water) customers suddenly experience a temporary issue with the look, taste or smell of their tap water. The water is still safe to drink. The most common issues are the water turning a light brown colour or a chlorine smell, typically lasting up to 24 hours.
Chance of property flooding from a burst pipe	Sometimes the main water supply pipe owned by the water company can burst and flood the ground floor of a customer's home or business. When this happens, (TEXT SUB: South Staffs Water / Cambridge Water) covers the cost of the repair through its insurance to get the property put back as it was.
Low water pressure	Every year some properties experience temporary periods of low water pressure, normally lasting less than 6 hours. These periods of low pressure are usually caused by problems with the pipe network.
Supporting nature and wildlife	(TEXT SUB: South Staffs Water / Cambridge Water) has a legal duty to protect and enhance nature and wildlife and ensure there is no permanent damage to the areas where it operates. The company aims to ensure rivers, (TEXT SUB IF CAM: chalk) streams, reservoirs and underground water stores are healthy.
Unplanned short interruptions to water supply	Every year some customers will experience a short interruption to their property's water supply, where it suddenly stops working without warning for 3-6 hours. During this type of interruption, (TEXT SUB: South Staffs Water / Cambridge Water) would deliver bottled water directly to the homes of vulnerable people.
Risk of temporary use ban, including hosepipes	To protect essential water supplies during extended periods of dry weather, (TEXT SUB: South Staffs Water / Cambridge Water) may send you a notice saying you must not use a hosepipe or sprinkler, or use water for other non-essential uses. The length of temporary use bans can vary, but are usually issued for five months, between May and September.

Although these will also feed into the WRMP thematic reviews, there will also be the requirement to review wider themes, notably the price control deliverables, such as universal metering / smart metering and changes to water companies' operating licenses, such as financial support and PSR support.

Key to the mapping stage is the conversion of the local measures used (e.g., ratings, importance index values, WTP, mentions / interactions from tracking surveys) to a common score based on values apportioned to a total of 100%.[9] While this is a relatively simplistic way to compare the values, it has the merit of being transparent and practical to implement. The potential downside of this approach is that it can give the appearance of equivalence between different data sources which vary in robustness. For example, calculating proportional values from rating scales has less theoretical validity than apportioning utility values from SP trade-off exercises. Such caveats will be considered when comparing across the different sources as part of the RAG approach.

Assess

The PR19 approach to assessing evidence began with defining two types of validity: 'theoretical' and 'statistical'[10]. Theoretical validity covered the robustness of the questioning method (e.g., use of established rating response scales), the accuracy of how the service attributes were communicated with customers and the potential biases that may be present. Statistical validity related to the specific features that affect statistical robustness: sample size, sample representativeness, statistical accuracy of models estimated from the data (model fit) and use of best practice techniques.

To make the approach more inclusive of non-numeric outputs, we will add a third dimension of validity: 'depth'. This relates to the *quality and detail* of information given to survey participants and the level of discussion and education that contributed to participants' views. The intention is to encourage greater consideration of qualitative sources. These will not normally provide numeric values comparable to those provided by quantitative sources but may give more confidence that issues have been covered in sufficient depth for customers/citizens to express an appreciably different opinion.

Rate

For each of the dimensions of validity (Theoretical, Statistical and Depth), Red/Amber/Green ('RAG') ratings can be applied. The table below shows examples from the PR19 approach (where only Theoretical and Statistical validity were included).

Table 3.2: Examples of Validity Criteria

Validity measures	Main WTP 'Core Research' - Key considerations	RAG rating
Theoretical validity	i. WTP surveys involved customers in the design of survey and service measures for inclusion in the main survey hence the core measures are based on more informed choices	GREEN
	ii. Well-designed and implemented WTP studies	GREEN
	iii. Relative values in the WTP core DCE are affected by differences in scope of change offered, but this study is the benchmark	GREEN
	iv. WTP core Max Diff measures are however weaker on validity since there is no scope sensitivity and no explicit controlling for bill impacts	AMBER
Statistical validity	i. Robust sample frame to ensure representation of all customer types (including hard to reach)	GREEN
	ii. WTP measures based on large sample sizes	GREEN
	iii. Results based on recent survey (2017 and 2018)	GREEN
	iv. Due to the very small sample sizes obtained for non-households in the Cambridge area in the Wave 2 research, we down-weight this segment of results in the triangulation analysis.	RED
Overall validity	i. WTP Wave 1 DCE: Well-designed and implemented study with large representative samples of SSW/CAM customers	GREEN
	ii. WTP Wave 2 DCE: All segments except CAM NHH	GREEN
	iii. WTP Wave 2 DCE: CAM NHH	RED
	iv. WTP core Max Diff: Same statistical properties as WTP Core: DCE, but somewhat weaker on validity	GREEN / AMBER

[9] Accent and PJM Economics, 2018, PR19 Data Triangulation, SSC:
The various measures are combined as follows:
- The sum of the priority indices derived from each source is reweighted to account for the fact that not all the data sources include sufficient information to derive a priority ordering for all the WRMP options.
- The priority indices from each of the sources are then rescaled with respect to the reweighted sums obtained in the above step.
- Finally, a Combined Priority Index is obtained for each of the WRMP options by taking a weighted average of the priority indices derived from each of the data sources and then rescaling them to ensure that they sum to 100. Further, the range for the priority index for each of the WRMP options is defined as the difference between their minimum and maximum values.

[10] South Staffs Water – Water Resources Management Plan 2019: Appendix A7 - PR19 data triangulation study – SSW WRMP, p9

A further aspect of statistical validity is the implicit assumption that 'best practice' applications of stated preference methods have been used (for example in the efficiency of experimental designs and robustness of the statistical models), though this is not cited specifically (see peer review query in Appendix 3).

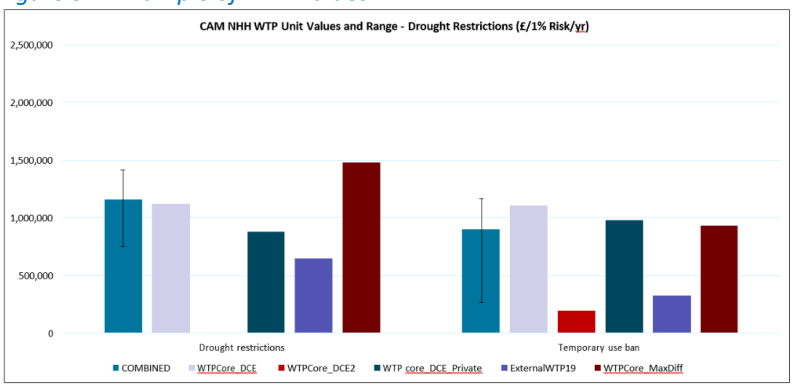
Weights were then associated with the PR19 RAG ratings to represent the relative strength of validity, as follows:

Table 3.3: RAG weightings

Overall RAG rating	Weight used in PR19
Green	100%
Green / Amber	50%
Amber	25%
Amber / Red	10%
Red	0%

Finally, these weights attached to the RAG ratings were used to derive an *average value* for each service attribute across the different sources, as illustrated in the example below, taken from the PR19 triangulation[11]. Here, the Max Diff value (brown bar) and DCE 2 value (re bar) make a 50% and 0% contribution respectively to the Combined value, reflecting the Green/Amber and Red classifications given to them.

Figure 3.1: Example of WTP values



Accent and PJM Economics, 2018, PR19 Data Triangulation, SSC

The RAG weights are somewhat arbitrary (they effectively halve in value with each reduction in level) and sensitivity testing will be conducted to test their influence on the final recommended central WTP values, in the same way as carried out for PR19, where a pure linear scale (100,75,50,25,0) was used instead. The point has also been made by our peer reviewer that if a study is rated 'Red' for theoretical validity, statistical validity becomes irrelevant. A hierarchy should be assumed by which a minimum level of theoretical validity is sufficiently established (eg rated amber or better) for a study to be included.

Our original intention was to ask Delphi panel participants to review the detailed RAG ratings and give overall *rankings* of the data sources. The purpose of this was to encourage greater differentiation in the way the different sources are assessed – useful, for example, in cases where several sources are classified with the same RAG rating. However, during development of the Delphi questionnaire, it became apparent that this would be too onerous a task for participants. Instead, they will be asked to consider the overall WTP values (low – central – high) for each service attribute that arises from the RAG ratings and assess how well these represent the views of customers, given all the evidence made available to them.

An example of the weighting tool, which will be used by all Delphi participants to rate the evidence and form the basis for apply the RAG weightings to the available data, is shown in Appendix 1.

Triangulation

In the same manner as for PR19, the final WTP triangulation results were derived from a process of applying weights to each of the data sources based on their overall RAG ratings and then combining these measures to derive central values and ranges for the core WTP and customer preference service measures subsequently used in investment modelling.

[11] Accent and PJM Economics, 2018, PR19 Data Triangulation, SSC

The triangulated WTP values reflect customers' willingness to pay for each of the WTP options and will be used within CBA as part of the process of setting any SSC be-spoke PC levels or price control deliverables. In conjunction with outputs from the central Ofwat research and in keeping with the PR19 approach, the household and business WTP are added together for each region and the regional WTP totals are weighted by the size of each region (using property counts) to produce a final, weighted, combined WTP. The 'Combined SSC' WTP triangulated values are calculated as a weighted average of the South Staffs and Cambridge area results.

Sensitivity

We will derive low, central, and upper values for the WTP metrics. With the introduction of the low/central/upper values, there will be a requirement to work with the modelling and planning teams to test the full potential variation in values. For example, if there are 10 PCs to test, the number of potential combinations to model will increase from 10 (single values) to 10^3 or 1,000 combinations of low, central and high. Clearly this is likely to be impractical, so we would anticipate, as part of any ongoing support to the end of AMP7, a process of identifying the main combinations to test, perhaps focussing on those items that are highest in impact and /or have the widest value range. Also, we suggest using a statistical (fractional factorial / orthogonal) design[12] that tests only 'main effects' of the individual PCs/ODIs, and thus keeps the amount of testing more manageable level. For example, such a design would allow 12 service attributes, each of 3 levels (lower, central, upper values), to be represented by just 27 combinations (from a total full factorial of 3^12, more than half a million combinations)

[12] ".. a designed experiment is orthogonal if the effects of any factor balance out (sum to zero) across the effects of the other factors. Orthogonality guarantees that the effect of one factor or interaction can be estimated separately from the effect of any other factor or interaction in the model." (https://support.minitab.com/en-us/minitab/18/help-and-how-to/modeling-statistics/doe/supporting-topics/basics/orthogonal-designs/)

4. VALIDATION

Internal Validation

The Delphi Method

A key innovation compared to PR19 will be the expansion of the number of people who will be involved in the rating and ranking of evidence, using the Delphi method to encourage movement to a consensus view (or, in cases where consensus cannot be reached, a clear set of arguments for different outcomes which SSC can use to choose a final set of ratings and rankings). The Delphi method is a flexible tool for drawing together the opinions of several independent assessors with the aim of moving towards a consensus, or at the very least a well-documented and clear basis for disagreement. It is used for decision-making and forecasting, with the former being the more relevant to this current application. It is an established process used in numerous fields of study, by which a panel of people with relevant expertise (be it technical or commercial) are invited to take part in an iterative process run by a single organiser. All participation is anonymous, which overcomes several potential biases that are present in any qualitative group exercise: anchoring bias (the impact made on opinions of the first speaker in a discussion), halo / authority bias (the undue influence of an individual because of personality / status / specific expertise) and 'loss of face' (the natural unwillingness of participants to be seen to change their mind).

In their account of using the Delphi Method, Amos and Pearse[13], suggest that the Delphi technique is typified by five main characteristics:

- its focus on researching the future or things about which little is known
- reliance on the use of expert opinion
- utilising remote group processes
- the adoption of an iterative research process, and
- the creation of a consensus.

Our Approach

Compared to typical applications of the method, our intended approach is quite narrowly defined when considered in terms of the first point above and equates more to 'things about which little is known'. It can be captured in the core question: "What is the appropriate WTP value to use for each service attribute in SSC's investment appraisal for PR24?' This is accompanied by the secondary question of 'What is the appropriate range of values to test around each of these central WTP values?' Not only will it require assessment of the validity of WTP values derived from research, but also evidence from wider sources that could indicate whether these values are low or high in comparison. There is also an issue relating to what these values represent: use or non-use. See Appendix 3b for further discussion.

How will the Delphi survey instrument be constructed and who is involved in this process?

Impact Research is responsible for constructing and implementing the method, but our intention is that every step is open to scrutiny both internally (SSC) and externally (peer review and SSC Customer Stakeholder Group). This will include the way we summarise panel member's feedback and select their open-ended responses to support their assessments.

The typical steps of the Delphi method can be summarised as follows[14]:

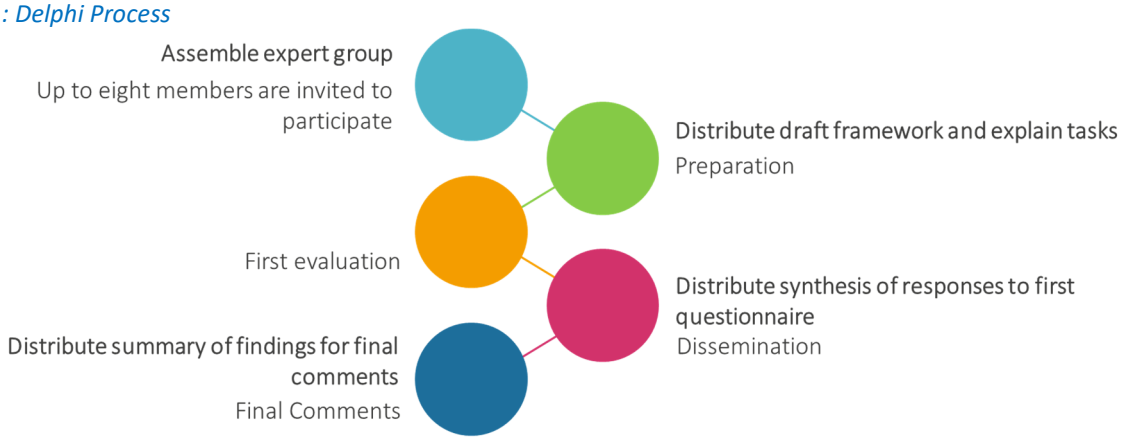
1. Identify your experts
2. Identify other roles (facilitator, analyst, etc)
3. Craft the first set of questions
4. Tabulate and analyse results
5. Prepare the next set of questions
6. Continue gather-analyse-ask
7. Stop at convergence
8. Report

[13] Amos and Pearse, 2008, Pragmatic Research Design: an Illustration of the Use of the Delphi Technique, EJBRM, Volume 6 Issue 2.

[14] Clayton, 2021

With these steps in mind, the planned structure and content of the Delphi Method for this current work is:

Table 4.1: Delphi Process



It is suggested in the literature[15] that panels of 10 or more participants are typical, as are three iterations or more. In comparison, our approach is necessarily more modest in scope. The time and expense in recruiting a large panel of experts is an importance practical consideration, so we planned to recruit a smaller group, which results in 4 participants. While resources are also a factor in limiting the number of iterations to two, it also reflects the focussed nature of the study scope. Typically, Delphi studies include a wide-ranging set of open-ended questions at the first stage, reflecting the complexity of the issues related to a topic such as forecasting future economic performance or agreeing the development steps for a new health treatment. In our case, the sources of information will already be defined (all available WTP studies and related research) and the task narrowly defined (agree the most appropriate WTP values to use).

Our approach will use interactions between expert group members via anonymised questionnaires rather than face-to-face communication. This encourages focussed and open expression of opinion, where potential conflict of interests, issues of confidentiality or simply the influence of personality are minimised. Core to the method is the use of a multi-stage self-completion questionnaire with individual feedback encouraged.

A summary of the steps in the Delphi approach for this research is shown overpage. The initial briefing note to participants and a draft first-stage questionnaire is shown in '**SSC10 PR24 Technical Triangulation – Application of the Delphi Method**' Appendices.

The Expert Panel

The expert participants were recruited from different disciplines:

- An expert in the field of customer engagement from the energy industry
- A technical expert in the field of stated preference research and WTP estimation
- A representative from the ICG
- A representative from Sustainability First.

Anonymity is key to the process, so the views of panel members on the validity of contributions from their fellow panel participants will be confined to the answers that are given and reported after the first stage of the process.

First Round

The first-round questionnaire presents a series of statements that each participant is asked to evaluate, using the RAG rating + ranking approach, supplemented by open-ended justifications for each rating. Participants are asked both to rate the item and to write free-text comments that, for example, explain their rating or express disagreement with the statement's relevance. Participants will be asked to focus on WTP values; these best represent customers' priorities and avoids discussion of technical issues related to how best to represent these in CBA. Rankings of the WTP values will also be requested, as recommended by our peer reviewer.

For clarity, and ease of comparison, we will split WTP separately by HH/NHH, given that the values and priorities of these groups are different.

Second Round

[15] Fink-Hafner, Dagen, Douˇsak, Novak and Hafner-Fink, 2019, Delphi Method: Strengths and Weaknesses, Metodoloˇski zvezki, Vol. 16, No. 2

The responses to the first-round questionnaires are collated and used to create the second-round questionnaire. The latter presents the same statements as before, together with both the individual respondent's rating and the (anonymised) ratings from the each other member of the panel. A selection of the free-text responses is given, to represent the breadth of opinion. After considering the group ratings and free-text comments, respondents re-rate the statements, by either giving the same rating as before or an amended rating. Respondents may give further comments about the statements if they wish.

Impact sent an outline of the process to participants in November. Each participant was expected to participate on three occasions (first assessment, follow-up assessment, final comments), each separated by a gap of at least two-weeks. Participants were advised that some three to four days in total would be required over the whole process to complete the work.

The Role of the Facilitator

The Impact team will be led by Dr David Pearmain, who will be supported by a dedicated research team to develop the approach, facilitate the research process, summarise the responses and report the final recommended WTP values with explanatory notes to justify this recommendation. The team will draw on their extensive knowledge and experience of deriving WTP values for assessing consumer valuations of Water and other Utility service investments. However, the nature of the Delphi approach means that the actions of the facilitators will have an influence on the final outcomes. This is reflected particularly in the way that the open-ended commentary from experts is summarised for use in the second iteration.

We will address this potential for bias in the following ways:

- All quotations used in the summary of the response to the first stage questionnaire will be selected only in so far as they explain the RAG ratings that participants have given to the WTP values.
- The summary of the responses will be shared with internal and external parties for scrutiny before it is sent out to the panel for the second stage

Reaching a Consensus

In some applications, the Delphi method continues through however many iterations are required until a consensus is reached. Our narrowly defined approach, limited to two iterations, requires that this aim is relaxed. A more applicable definition of the outcome is “for each WTP value established from this process, what is the level of consensus *or otherwise* around it?” By exposing each panel participant first to a summary of the research related to the WTP values and then the responses of the rest of the panel, we will establish a series of data points (WTP values) that may or may not converge closely around one value for each service attribute. It is entirely acceptable for some of these values to still have a wide range of uncertainty at the end of the process, because this will reflect the diversity of opinion. Such instances will simply place more emphasis on the importance of sensitivity testing in the CBA process.

Our peer reviewer suggests that a "consensus" can be thought of as a point estimate, whereas we are expecting a range of values, such that the consensus explicitly reflects the uncertainty inherent in the WTP values. This is credible given that WTP estimates are typically subject to large standard deviations and that focussing on a point estimate can be very misleading. As part of the Delphi approach, we will ask participants to express their degree of uncertainty about the WTP values.

The output from this process will be a more considered set of arguments supporting or questioning each WTP value. In the final use of these values in the investment process, SSC may well choose to alight on one particular WTP value that is not necessarily backed by a consensus from the panel; however, as a result of this process, they will be in the receipt of a set of arguments in support and in contradiction that will enable them to more clearly articulating why they have opted to use that particular value.

Figure 4.2: Flow of the Delphi Method

Assemble expert group

Up to eight members are invited to participate

- We suggest a mix of academics, commercial practititioners and WOC personnel, with a maximum of two people in the following areas:
  - Representatives of SSC, one whose background relates to customer insight and one to senior particpation in policy planning
  - A representative from another WOC
  - Commercial practitioners in the field of customer engagement, ideally not directly involved in the collection of the main inputs to the triangulation
  - Academics with knowledge of the use of customer engagement measures in plan formulation for utility companies
  - A member of the Challenge Panel who would partiipate but also report to the wider Panel

Distribute draft framework and explain tasks

Preparation

- On completion of the first draft of the triangulation framework, a summary of the approach will be distributed, together with all relevant documents
  - Participants will not be expected to read all the material, but to have these documents to hand for reference as required
  - They will be given a fuller explanation of the Delphi process with timing milestones
  - This will include a process for assessing bias, agreed by SSC and developed along the lines of the Bristol Water example noted in the brief
  - We will empahsise the focus on RAG ratings and assessing the relative importance of the different data sources (wider views will be welcomed, but not included in the Delphi process)

Distribute Draft insight reviews, analysis and tabulation

First evaluation

- The draft values derived from the insight reviews, analysis and tabulation of any technical work required for WTP valuations will be sent to each participant
  - This material will be accompanied by a questionnaire asking participants to assess the WTP values in the light of the evidence share with them
  - Participants will be encouraged to express dissention / interpretation, with open-ended statements justifying their view

Distribute synthesis of responses to first questionnaire

Disseminaton

- The final version of the agreed triangulation framework will be distributed, incorporating draft values for WTP components that reflect the feedback from the first evaluation above and the arrival of new values based on recent survey work
  - It will be accompanied by a summary of this feedback, together with the updated WTP values and ranges.

Distribute summary of findings for final comments

Final Comments

- The final values derived from the insight reviews, analysis and tabulation of any technical work required for WTP valuations will be sent to each participant
  - Each participant will be invited to submit a final statement summarising their views on the process and the validity of the final values used, together with suggestions for future improvements to the process
  - It will be essential to accept that consensus of opinion may not be possible for some plan areas, but the process will deliver explanations for the different opinions and assist SSC in determining which opinion to side with, for the purposes of proceeding with a clear guide to preparing the plans.
  - The final output will be a summary report of the process with appendices containing all anonymised options.

External Validation

Representation to External Bodies

Working with the SSC insight team, regular updates and formal reports will be delivered to the following:

- SSC's WRMP24 and PR24 independent assurance consultancy (Jacobs), who will require assurance that the triangulation approach is high-quality and 'line of sight' on how the outputs have been used to shape decisions in the PR24/WRMP24.

- SSC's newly formed Independent Challenge Group (ICG), who will require regular updates and a full opportunity to critique all aspects of the process. A member of the panel will be included in the Delphi expert team.

- The SSC Board, as per Ofwat's policy paper from Feb 2022, have certain obligations around engagement with the research programme. This includes triangulation, the effectiveness of any challenge and line of sight of outputs and how they are used to shape the plans.

Independent Review

Role of the Peer Reviewer

The role of the peer reviewer is to assess the proposed triangulation procedures (SMARTS, RAG weightings, all calculations).

The reviewer has been invited to review at each of the following four milestones:

- Draft framework
- Draft values
- Draft triangulation report and
- Final triangulation report with considerations for future proofing the framework to support delivery of the PR24 plan.

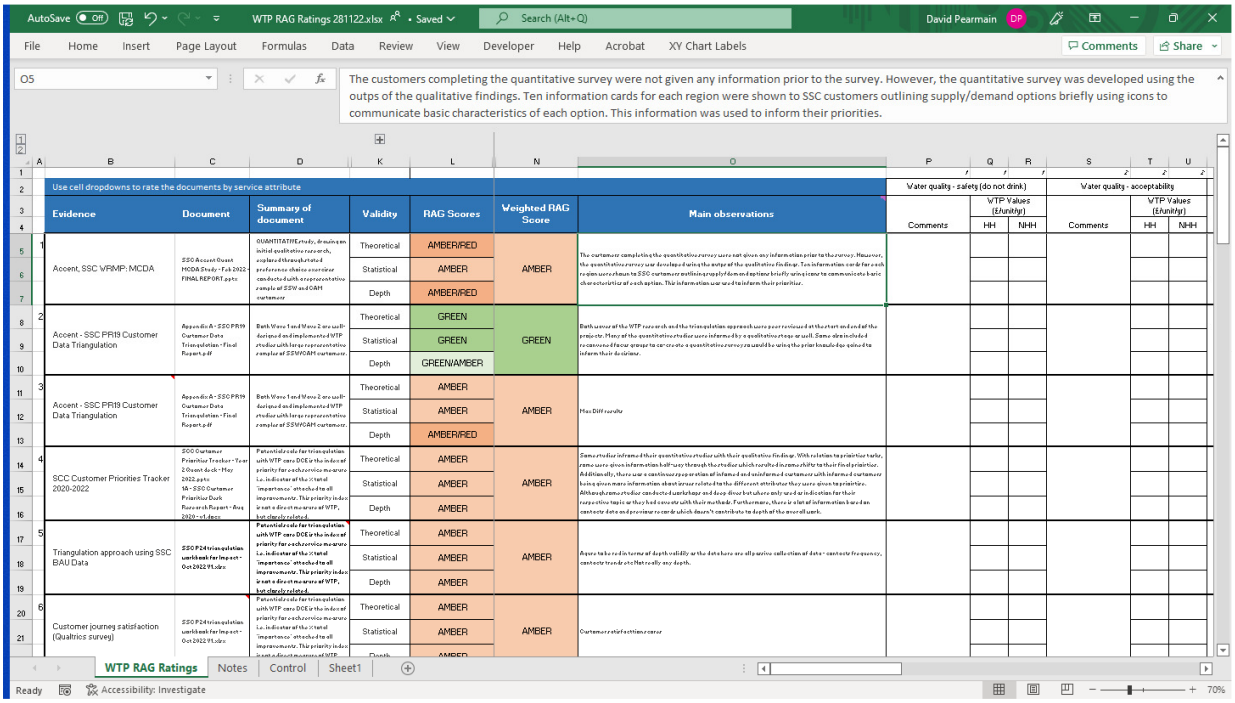
While we expect a good working relationship with the peer reviewer, we will also be mindful of the need to ensure the engagement does not compromise their independence. We will maintain a suitable level of formal working throughout, embodied in the careful construction of technical reports and records of all communication.

We will also liaise with other agencies working for SSC through the company's PR24 supplier framework and wider club projects who are likely to be involved in supplying new customer engagement or using their expertise to review and input to the triangulation approach.

5. APPENDICES

Appendix 1: RAG Rating Tool

A short description followed by key features of each source will be provided:



For source, ratings will be determined for the three types of validity, together with notes that relate to the service attributes covered. When all sources are rated, the scores are combined into a single rating for use in the WTP analysis.

Click on the embedded workbook to call up a working example.

WTP%20RAG%20Ratings%20281122.xlsx

The criteria for determining the RAG ratings may develop over time, especially when the process is used by the Delphi panel, but initial guidelines are offered, based on the PR19 experience:

Theoretical validity	How appropriate are the findings from this source for their intended real-world application?
Statistical validity	How robust are these findings in statistical terms (sample size, quality of design, representativeness, randomness)
Depth validity	The quality and detail of information given to survey participants and the level of discussion and education that contributed to participants' views. The foundation of explorative and developmental work behind the design of the study.

Appendix 2: Peer Reviewer and Other External Challenge

Over the course of the triangulation process, the work will have been regularly reviewed by an independent peer reviewer, Professor Iain Fraser of Kent University. This appendix summarises key questions raised at various stages:

17 July 2022 (Peer Review of initial draft document)

Peer Reviewer feedback	Response
Abbreviations need to be defined as not all readers will be familiar with all of them.	Glossary added
Under the title "The main challenges", you mention there are 5 stages then only report 4.	Titles updated, as stages 3 and 4 are combined in a single section
I'm not sure what to make of the weights attached to RAG and what the figure below the table is showing - it isn't clear. I'm unsure what mean when you refer to "effectively logarithmic"?	Section expanded and revised
Will you be providing an overview of the Delphi method? There are plenty of references on strengths and weaknesses. **Also,** • **how will the Delphi survey instrument be constructed/designed?** • **Who and how have been involved in this process?** • **How will you ensure that participants in the Delphi believe that all participants are "experts"?** • **Why only two rounds of interview? The literature generally advocates 3?** • **How are you going to evaluate consensus?** • **How will you evaluate the text responses?**	The section on Delphi has been expanded on the initial draft version and the detail of proposed panel composition and survey design will contribute to develop. In this latest version of the framework report, we have taken the questions posed and addressed them directly in the main text.
In terms of sensitivity analysis - I'm unsure what you mean by a statistical (orthogonal) design in terms of the data. You could use PCA maybe?	The intention is to use a select combination of WTP values to reflect the impact on overall CBA results of each individual service attribute, without having to test every combination of values. Fractional factorial / orthogonal designs are a long-established method to achieve this. This is where a fraction of all possible combinations of values are used, ensuring the impacts of individual items ('main effect') can be tested independently of any potential interactions.
I don't fully understand what the paragraph that starts "We will also establish..." is telling me? This needs more explanation.	This is simply to encourage more differentiation in the way sources are evaluated, in case RAG ratings are very similar – text now revised to reflect this.
More generally, once the Delphi survey is generated and an extended discussion of how it has been designed is provided then it will probably become easier to understand how all of the information fits together.	-

7 August 2022 (Peer Review of draft final document)

Peer Reviewer feedback	Response
Background (references to ROAMEF and Magenta): **This amounts to a literature review and a set of subjective judgements about what matters in these reports.**	An expanded section on this has been added
Stage 5: Link to business plan: **Not sure I understand what is stated here actually makes a link to a business plan...**	This section expanded to explain reporting stages and the relevance of the parallel thematic reporting
3. TECHNICAL TRIANGULATION DETAILS, SMARTS, Screen: **What is being proposed is a mix of new stated preference research and benefit transfer - there are some standard references on how benefit transfer can be undertaken - I suspect it is explained the Green Book (March 2022)** **https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent/the-**	Given the hierarchy of valuation methods listed in the Green book (Market prices, generic prices, revealed preferences, stated preferences (SP), subjective well-being), SP is the most relevant data source for addressing specific service levels, but we will endeavour to draw in results from wider sources where available.

Peer Reviewer feedback	Response
green-book-2020#a1-non-market-valuation-and-unmonetisable-values **See section 9.A1 - it is called Value Transfer in this document**	Annex 1 lists low-central-high values for the following: Air pollution, noise, nature-based recreation / benefits Local and visual amenity, water availability, water reliability, flooding, carbon reduction, soil erosion, land value uplift and Carbon emission values. We have included a selection of potentially relevant values from this annex in Appendix 3a
3. TECHNICAL TRIANGULATION DETAILS, SMARTS, Rate, Statistical validity: **Odd - no mention of model robustness - everything relates to sample size. There is also no discussion of experimental design (why in theory as it is an applied feature?) and how robust, etc. So, might suggest that existing statistical validity "considerations" are at best weak and would not stand up to typical academic review for published work.**	Comment added in this section: "A further aspect of statistical validity is the implicit assumption that 'best practice' applications of stated preference methods have been used (for example in the efficiency of experimental designs and robustness of the statistical models), though this is not cited specifically".
3. TECHNICAL TRIANGULATION DETAILS, SMARTS, Rate, RAG: **If you have Red for the Theoretical Validity for any aspect then this in principle means that the study should carry a Red flag, full stop - seems that that the RAG needs to be applied in some sort of hierarchical way for it to make sense.**	This is consistent with the suggestion that rankings should play a part in capturing the views of Delphi participants (see comment re best-worst scaling below). This is addressed with this addition to these section: 'A hierarchy should be assumed by which a minimum level of theoretical validity is sufficiently established (eg rated amber or better) for a study to be included. We shall therefore ask panel participants to give overall *rankings* of the data sources as well as RAG ratings.'
3. TECHNICAL TRIANGULATION DETAILS, SMARTS, Sensitivity: **This would seem to imply that the values are drawn from a symmetrical distribution…?**	This is not an explicit assumption – the distribution could be asymmetrical. PR19 used a pragmatic +/- 20% range: '*for sensitivity testing we define the low and high values such that the Low value is calculated as the minimum WTP value plus 20% of the difference between the minimum value and the central case value, and the High value is calculated as the maximum value minus 20% of the difference between the central case value and the maximum value. The justification for redefining the confidence intervals in this manner is to avoid having extreme range of values for the Combined WTP.*'
4. VALIDATION, Our Approach: **I can't help but wonder if asking about the relative importance of the attributes, in a best worst scaling type of approach would be an easier task than simply "is the WTP high or low?" And will it be the household WTP values or will it be aggregate WTP values that will be considered? In that sense, is it the WTP as it enters the CBA that will be considered? I suppose you have information (previous estimates) so as to establish a reference point but "asking for the appropriate WTP" is a difficult task...**	We suggest focussing on the WTP values, as we really want to look at what best represents customers' priorities, rather than the technical issues related to how to incorporate this in CBA. We would also suggest that ranking is sufficient (Best-Worst is more a quantitative technique for large samples). For clarity, and ease of comparison, we will split WTP separately by HH/NHH, given that the values and priorities of these groups are different. Text reflecting the above has been added to this section.
4. VALIDATION, Our Approach, Reaching a Consensus: **A "consensus" can be thought of as a point estimate. What you seem to be seeking is more a range of values - so in this way if you end up with a wide range of values such that the consensus explicitly reflects the uncertainty inherent in the WTP values. This is credible given that WTP estimates are typically subject to large standard deviation and that focussing on a point estimate can be very mis-leading - I think you can highlight this aspect of the uncertainty more to explain the proposed approach. In fact, as part of the Delphi you could ask participants to express their degree of uncertainty about the WTP values - this would then relate to how some CVM WTP studies ask respondents the WTP question and then ask then how certain they are of the response they have just given.**	Text added to this section to reflect these points

September/October 2022 (Peer Review of draft Delphi approach)

Peer Reviewer feedback on Draft approach	Response
The more I read through what is being proposed and what is required the more I think that the task in hand is very complicated. **To take issue with WTP estimates and to propose alternative values implies a view about what participants think to be the truth. The does require them to understand what the WTP estimates actually represent. And this is hard...**	A discussion on how and why the approach was re-designed is given in '**SSC10 PR24 Technical Triangulation – Application of the Delphi Method**' Appendices
These [*WTP*] values would make more sense if you included error bars for all estimates. I think unless an estimate is expressed in this way it should not be presented as meaningful data	Agreed that low and high values should be included for all central WTP values shared with participants
On the text – *'The first part of the questionnaire therefore invites participants to prioritise the topic areas in order of their perceived importance to customers (based on their reading of the relevant literature).'* **This is clearly subjective - who says what is relevant?**	This relates to the need for a re-think in terms of how to keep the exercise manageable for participants. We cannot ask them to do their own literature reviews as there simply is not enough time. The exercise needs to be more along the lines of: 'this is the evidence we have compiled - based on this, plus any other knowledge you may have, please make your assessments
The time it will take to read and evaluate the material is significant. Is the task too demanding? How can you be sure that the participants won't become tired or bored? How to ensure data quality is maintained seems a real challenge. **Will you randomise the order of topics by participants as a way to check if any variations in evaluations are topic and not time related.**	These are certainly issues that could be addressed with randomisation of items, but the key point is the complexity. This led to a major re-draft of the approach, in which extensive summary material was prepared, drawing on a wide range of customer engagement sources, both quantitative and qualitative in nature. The questionnaire was also extensively revised, focussing more on how participants responded to the collated WTP values, in the light of what they had learned from the summary materials and their own familiarity with the issues. The contrast in the questionnaire can be seen by comparing the versions in '**SSC10 PR24 Technical Triangulation – Application of the Delphi Method**' Appendices. This includes discussion on how and why the approach was re-designed.

November 2022 (Peer Review of Final Delphi approach)

Peer Reviewer feedback on Final approach	Response
Various comments on clarity of text and style	Incorporated in final questionnaire
Future Customers: **These groups are clearly defined? Future customers are only households?**	Added (HH) and a more detailed description for future customers.
Supporting nature and wildlife: **Non-market – will be a non-use value – issues around valuation will become important** **I think a statement about the scope of the benefits matters. What you see here is the fact that a benefit such as protecting a wildlife habitat, which yields a small WTP, will be a non-use value and as you multiply by all SSC consumers it becomes very large very quickly. I think the habitat measure distorts things here. Also SSC will likely be paying farmers to manage land in terms of water quality and it is the cost to land users and not the public that matters – even if the benefit estimate is very large, only a fraction of that amount will be required to induce land use change by farmers. This is probably an over-estimate in aggregate.**	A helpful summary of the use/non-use issue, but the question for the study was: do we share this with participants or leave as it is? It was agreed with SSC to leave it as is, and see what feedback we receive, as to how much participants perceived this as an issue.
Resource Material prepared by Impact **- there is a lot of info, but I suppose that is inevitable given the number of themes being examined. The overview document is very nicely put together - I like the mix of media and research information. A lot of care and attention has clearly gone into preparing the materials. Just one thing - will all participants know the geographic coverage of the two firms? Do you need a map showing the areas covered?**	Suitable maps and summary stats provided by SSC were added to the resource materials

Alongside the peer review, we also received initial feedback from Simon Sperryn, outgoing Chairman of SSC's Customer Challenge Group (CCG):

CCG Challenge raised	Response
I have Impact's triangulation proposal. I think this is a well-judged commission. The need for, and benefits offered by, triangulation are easy to agree. But it is hard to get any clarity on how to do it. Furthermore, in PR24 you will have not only to triangulate the various research data collected by you for your business plan, but, for the first time, you will also have data related to WRMP to triangulate with the business plan research, and regional research to triangulate with SSC research, *and* national research managed by the national steering group to triangulate with local research. Having third party support through this 'knot-garden' makes good sense and will strengthen your ability to evidence the logic of your conclusions and the use of best practice in methodology. My main concern is whether the brief reflects this wholly new scenario adequately. In particular, the national versus local issue, on which COG members are pressing Ofwat for clarity, might be worth including specifically. Also, I wonder whether you could use Impact's advice on how to make best use of your organs of customer engagement and the revised Stakeholder Group in challenging the methods or the conclusions of triangulation. Impact's experience in the energy sector chimes well with Matt's experience which could be a useful bridge.	We propose to include national WTP figures in the triangulation (where available) to complement those derived from SSC regional / local studies
I am delighted to see in the business objectives that you propose to assign upper, middle and lower values for WTP results. That seems to me a very practical way of embracing the fact that WTP produces random values that vary with the wording the timing and the weather. A range will reduce the vulnerability inherent in assuming any accuracy in WTP values and give you more scope to pick values that triangulate with more meaningful research results. The flexible approach outlined in the proposal makes sense given the uncertainties surrounding triangulation, the fact that Ofwat will gradually reveal its requirements for PR24 as the work progresses, and the iterative nature of the triangulation task as different research projects produce results. The idea of regularly reviewing Terms of Reference and of retaining the agency to work alongside the development of the business plan seem practical to me.	-
There is a danger that developing a tool for managing multiple research inputs could descend into a formulaic process based on weightings and scores that takes the place of interpretation and judgement. Impact seem aware of this danger and I liked their reassuring "we think it likely that the final ratings will still rely largely on 'expert judgement'" (page 11).	The Delphi method is critical in this respect, as it aims to gather opinions as well as ranking data from participants. Please see Chapter 4 for an outline of the Delphi method, where it is described as '*a flexible tool for drawing together the opinions of several independent assessors with the aim of moving towards a consensus, or at the very least a well-documented and clear basis for disagreement*.'
This looks like a really difficult project to manage, given its duration, flexibility, and organic development. I expect it will depend on a particularly strong working relationship between you and the agency.	Please see the section: 'Stage 5: Link to business plan' which outlines the close working approach between SSC and Impact

Appendix 3a: Relevant environmental values referenced in the Green Book, 2022, Annex 1 (20/21 prices)

	Description	Low	Central	High	Unit
Nature based recreation	Welfare value of outdoor recreation sites	£48	-	£120,067	per hectare (various land covers)
Physical health benefits from nature	Indicative health savings/benefits from every physically active visit to green space	£3.36	-	£14.34	per marginal physically active visit to greenspace
Local amenity	Average additional value per property within 100m - 500m of accessible green or blue space	£1,538	£3,076	£9,471	per property (capital value)
Visual amenity	Average price premium for a property with a view over green or blue space	-	£6,164	-	per property (capital value)
Loss of amenity	Welfare cost from significant litter accumulation in residential areas	£20	-	£76	per household
Water availability	Industry average present value lifetime social cost of providing water supply	-	£5.7m	-	mega litre per day
Water quality	Improvement in water quality status	£22,000	£25,400	£29,500	per km
Flood damage	Typical damage per property from a flood event	£8,000	-	£45,000	per property (flooding at different water depths)
Flood regulation (woodland)	Avoided water storage costs from woodland water storage in flood catchments	£97	-	£242	per hectare (woodland)
Nature based carbon reduction (peatland)	Carbon reduction value of restoring eroded peatland	£497	-	£5,297	per hectare (peatland)
Soil erosion	Average indicative cost of soil erosion (production, water quality, flood risk)	£130	-	£211	per hectare of average erosion
Greenhouse Gas values	Target consistent value	£121	£241	£362	per tonne CO2

Appendix 3b: Use and Non-use values

In cost-benefit analysis, and the WTP values that are used, a 'non-use' value is the value that customers assign to economic and public goods, even if they never used it and do not plan to use it in the future. This is different from a 'use value', where customers derive value from direct use of the good. An example of a 'non-use' value covered in the PR24 research might be natural wildlife habitats, where the value derived by customers is based simply on the knowledge that they exist/are protected, rather than direct experience/consumption. Most of the other investment areas may be regarded more as 'use' values, in that their value (or the value of avoiding their occurrence) is based more on how customers perceive themselves to be directly affected (eg risk of flooding, likelihood of a temporary use ban, etc).

However, it could be argued that the distinction becomes blurred when service improvements are expressed in terms of 'number of households affected', in that it implies risk of personal experience (use) but is expressed in terms of customers in general (non-use). For instance, a householder living in an area they know to have lead pipes is more likely to think of improvements to lead piping as a use value than a householder who is not in such an area. Another example, leakage, may be considered a 'use' value in that all customers are indirectly affected by leakage (ie translating to higher bills due to inefficiencies and lower resilience of supply), or it could be considered a 'non-use' value because it does not directly relate to their personal consumption of water.

The main concern for this study is whether the WTP values obtained for some items that have 'non-use' characteristics, notably environmental attributes but arguably a number of others, are over-stated. That is, we might assume that the WTP value obtained per household can then be applied to all households equally. However, it may be that customers are thinking of their current use/adjacency/knowledge of, say, local wildlife habitats, and if actual improvements to habitats are to locations far from their home, they would have zero WTP. The values for environment from the Green Book, listed in Appendix 3a, imply this, as they relate to proximity and use of green spaces.

When considering the range of values to use in the Copperleaf investment analysis, we therefore consider it advisable to include options that limit the scope of non-use attributes to the number of customers who would be considered as local, if not direct beneficiaries.