

This submission is made on an anonymous basis by a technical assurer who hopes it will make life easier for everyone involved in data management at future price reviews.

12. How can we simplify the price review while increasing value to customers, the environment and wider society?

The PR19 process included the publication of several iterations of the data tables. Multiple versions of the data tables is an almost unavoidable part of a complex planning and data gathering process such as the price review, but it can cause data management difficulties.

At PR19, changes to the data tables were logged in a separate file and could include changes to line definitions and line numbers, as well as additions and removals to the lines and of whole tables. I observed that to ensure the internal consistency of the submission and apply an appropriate level of governance, companies used complex software-based data management systems to allocate lines to individuals or teams, cross-reference them and track their progress through their internal and external assurance process.

The knock-on effects of line numbering changes were extensive because of the number of systems and processes which rely on the line numbers and the need to cross-check definitions with every update. The systems which rely on the numbering include commentary numbering, business plan references, internal and external assurance records, internal modelling processes and line ownership. The possibility that the items of data allocated to an individual might have a new line number almost meant that everyone providing data for the price review had to be informed and oriented every time there was a new table version. Every water company and their technical assurers had to undertake the complex task of mapping the new numbers to the old ones and applying version control for each change, which added complexity to the management process.

Three small changes to the update process can simplify the price review process.

1. **If a line is removed, keep the same numbers for the lines below it in the table.** This will result in there being 'gaps' in the numbering system but will avoid disrupting the reference system for the subsequent lines. As an alternative, to avoid creating numbering gaps perhaps it would be possible to grey lines out and mark them superseded, rather than removing them altogether.
2. **If lines are added, give them new numbers.** They can be added to the end of the table or, if they are so related to other lines that they have to be placed next to them, given new numbers that don't disrupt the numbering for the rest of the table. For example, if a need was identified for two more lines between lines 2 and 3, the numbers for the table could then run 1, 2, 2.1, 2.2, 3. This would avoid changing the reference system for the subsequent lines.
3. Give each line and definition a version number. If the definition is updated, increment the version number. The version number could be a separate column next to the line number. (On tables oriented around columns, each column would have a version number). That way it would be clear which definition the

information in the line follows and it will be inherently clear which definitions have been updated.

If a definition is updated without publishing a new set of tables, release the definition with the new definition version number and allow companies to update the definition number in their tables to show that they have checked and updated the data.

These changes would simplify data management, which is an important enabling activity for the price review. It would help companies to work more efficiently.

Finally, the completed data tables are very useful as a public data source. If Ofwat uses Python or other code to extract data from the tables as a long table, key-value pairs or similar, perhaps the code and/or data could be released to the public as open source in one place. This would add further value to the tables by making the data even more accessible.

(end)