Thank you for your comments and suggestions. We note that many of the comments relate to issues with making sure the visualisation technique is as engaging, accessible, and understandable as possible. This includes suggestions about the number of lines to use, the comparisons and range of countries being used to illustrate the method, and the associated amount of detail about each of these populations which is discussed, and the colour scheme. There are also suggestions about the colour of the contour lines and other aesthetic features.

As aesthetics is a somewhat subjective matter, and different users are likely to have different preferences in regard to these matters, we have developed an online app to accompany the paper which allows a great deal of customisation of such options. This includes options to change the colour scheme with which the levelplot part of the visualisation is displayed, the number of shades, the position and number of contour lines, and so on.

The app also allows any combinations of the 45 countries within the combined HFD and HFC collection to be visualised. This means that users are free to customise the visualisations and apply them to any particular use cases of interest to them, and to adjust aesthetic options such as to be most preferable to them. Once the composite level plots have been produced, they can be downloaded in a wide range of image formats, including both bitmap and vector graphics formats, and at a wide range of sizes.

Alongside the CCLP itself we recognise the importance of adding as much guidance and supporting information to these plots so as to allow viewers to fully understand what they convey. The visualisations are best understood as comprising elements from two distinct (but linked) Lexis surfaces, and so an important challenge is therefore to make the viewer fully cognizant of the surfaces being represented. A more conventional way of visualising such data are either in the forms of summary statistics, like TFRs, which sweep one of the dimensions into a margin, or as sweeps through of the primary axes of the surface - producing either fertility age a given age for different birth years, or fertility for a given birth year given age. To support understanding both of how the CCLP can be understood as comprising multiple sweeps through these axes, an additional option provided is for users to select a birth cohort and visualise the corresponding age specific fertility schedule and cumulative fertility schedules. These sweeps can be selected at will with a slider, and display simultaneously for as many countries as the user selects. The use of plotly, an interactive visualisation tool which is web-based, allows users to hover over elements of these schedules and see corresponding information.

By linking the app with the article, we hope both that the method of visualisation which the article introduces can be more easily understood, and also that the reader is then able to start using the visualisation method very quickly after reading the article in order to address those demographic issues of most interest to them.

Alongside the CCLP, the app also supports development of a 'Lexis surface' way of thinking about the data and qualities of fertility variation between ages, cohorts and geographically distinct populations through interactive 3D visualisations of the same surface, which can themselves be represented in a number of different ways, and alongside orthogonal sweeps through the data surfaces which appear interactively when the user hovers over positions on the Lexis surfaces produced.

In terms of changes to the manuscript, we have aimed to simplify the visualisations shown here by reducing the number of contour lines from four to two as requested. These contour lines are at 1.5 babies/woman, a level considered a lower bound of a sustainable population [?], and 2.05, i.e. replacement fertility in low infant mortality populations. We have also included a number of additional annotated versions of these visualisations which separately discuss important features relating to both the age-specific fertility surfaces (represented by the colour/shade), and cumulative fertility schedules (represented by the contour lines).

The specific countries and comparisons between countries we have focused in is now more constrained than in the original submission, and we have gone into more detail in order to explain the features shown within and between populations which are compared. The focus is now on two comparisons: between East and West Germany, which both now have low fertility rates, but where East Germany saw a rapid fall in fertility rates in the late 1980s; and between Norway and the USA, which are both comparatively high fertility populations, though have reached high fertility rates, slightly above replacement levels, through different means: through higher fertility at peak age of fertility in the case of Norway, and through a wider range of ages at which fertility rates are moderately high in the USA. We have discussed features within and between each of these populations, as they relate to both age-specific fertility, and cumulative fertility contours/milestones, in more detail than before.

In addition to the static annotated figures within the manuscript itself, and the app which allows a great deal of customisation and comparison of countries, we also include a number of 'animated' (through non-interactive) PDF 'flip-books' for the annotated figures as well. These introduce the annotations and features of the visualisations one step at a time, meaning that the figures should be more effective in guiding the viewer through the process of understanding and orientating themselves through the visualisation.

Combined with careful editing of the manuscript itself to make it more concise, we hope these supplementary materials will both help support the reader in understanding the visualisations so they are able to use them effectively, and to start using such visualisations for their own research activities soon after, citing the paper as they do so.

**Peer review comments and responses**

Dear Nikola, Tim and Sebastian,

I would like to thank both you and Professor Bijak for providing very helpful comments and hope that our paper has much improved as a result. Following your suggestions, we made extended revisions. For example, we have re-written the introduction to strengthen the novelty of the visualization approach introduced in the paper. Crucially, we have provided a much clearer description of this approach by using the cases of West and East Germany as illustrations. These also highlight substantive insights that can be derived from the examination of the visualizations concerning both West and East Germany. We have changed the colour scheme in all the visualizations to be more colour blindness/CVD friendly.

Please find our responses to each of your comments (CO#) as follows:

**C01: Major revisions required. This paper needs reworking: the novelty aspect beyond overlaying the Lexis surfaces with CFS isolines is not obvious, so this aspect would need strengthening.**

**(Nikola adds:) I'll chime in to try to clarify:**

**The novelty aspect could be addressed by adding a section in the introduction that explains in more details how your visualisation is superior to other existing ways of visualising that kind of data.**

We have extensively re-written both the introduction and concluding sections of the paper to address CO1. In the introduction, we have added a paragraph which describes the novelty of the visualization approach used in the paper, as it enables more efficient and at-a-glance comparisons of multiple and interrelated aspects concerning fertility trends and variation across countries and geographic regions. We highlight the value and benefits of the visualization approach proposed in the paper, by using West and East Germany as examples and pointing out some substantive features which arise from their examination. In the concluding section, we provide a summary and a brief discussion of our results, along with directions for future research.

**C02: The legibility of Figures 1-4 could be improved by adding a new figure 1 that shows just one or two Lattice Plots and uses direct labelling and perhaps a concrete example to guide the reader. East and West Germany would make a nice case study to explain what the plots show and why they are superior to other existing visualisations of the same data type.**

Response: See main response below.

**C03: I admit that I currently struggle to read the figures; especially the contour lines are hard to read, also due to the fact that the labels (for solid, dashed,...) are hidden in the notes below the plots. Perhaps you could adjust the colour scheme of the surface to use fewer colours and then use colour to distinguish the contour lines.**

Response: See main response below

Main Response: We have attempted to address C01, C02 and C03, by providing a much clearer introduction to the approach using West and East Germany as examples. In addition to illustrating the construction of the method, aiding interpretability, they also highlight, through inline annotations, a number of substantive features identified within Germany demographic data when both age-specific fertility rates, and cumulative cohort fertility milestones, are presented in the same visualisation. Given the word limit we have not attempted to provide a summary comparison with existing visualisation approaches, but hope through this detailed introduction to the approach, illustrating its substantive value through the German examples, the overall novelty and benefits of the approach are now much clearer.

The specific sequence of new illustrations is as follows:

The first subfigure (1a) demonstrates the intuition behind the contour lines by presenting cumulative cohort fertility schedules for two birth cohorts in West Germany: 1935 and 1938, alongside both cohorts’ age-specific fertility schedules. This subfigure shows how the contour lines are produced at whichever ages the cumulative fertility schedules of a given cohort intersect specific fertility thresholds or milestones. In the example given three of these thresholds are reached at similar ages for both cohorts, but the replacement fertility contour is reached at a substantially older age for the 1938 than the 1935 cohort. The complete cumulative cohort fertility for both cohorts is evident through the cumulative schedules becoming vertical from the early 40s onwards, illustrating why the associated contours tend to move near-vertically from these ages in many figures (i.e. if a birth cohort has not reached replacement fertility by around age 45, they are unlikely to reach replacement fertility at any age). This first subfigure also includes an inline sketch showing the monotonic ordering of four contour lines added to all figures. As the ordering of contour lines is consistent by definition, and each line uniquely identified through its combination of thickness and dash type, we hope to remain using hue and intensity to indicate age-specific fertility rates exclusively as before.

The second subfigure (1c) shows the two key elements of the composite plot – colour and contour lines – separately for West Germany, as well as combined. Each of the separated plots are heavily annotated to highlight both methodological features, which apply to all other visualisations, and substantive features specific to West German patterns of fertility change. The methodological features include: the fact population label backgrounds are coloured by region, where the colour key is located; and the reason the contour lines do not extend as far to the left as the first coloured tile. The substantive findings include the fall in fertility at all ages in the late 1960s; the upwards drift in age of peak fertility and corresponding fall in magnitude of peak fertility; the last birth cohorts to have experienced replacement or near-replacement fertility levels; and speculations about the likely trajectories of the 1.5 and 1.3 contour lines, and thus the likely the total cohort fertility levels at which West Germany appears likely to stabilise at. In total 20 features within the figures are highlighted using within-figure annotations, in order to make much clearer the density of information revealed about population fertility within a single plot.

The third subfigure (1c) provides a side-by-side comparison of East with West Germany on a common scale. A large number of substantive features have already been pointed out in the second subfigure for West Germany, and some additional corresponding features are highlighted through in-figure annotations for East Germany.

We have also reduced the number of specific examples we cover in the main manuscript, and moved some of the current material to an appendix. Whereas previously with included examples of more than two countries, with no within-figure annotation, we now focus on the comparison between the USA and Norway, which are unusual amongst rich countries in both re-establishing replacement fertility levels after first losing them. As with the East/West Germany comparison, this comparison is now greatly facilitated by extensive within-figure annotation, pointing out important similarities and differences between the two populations as they relate to both cumulative cohort and age-specific fertility values. As before, then then conclude with the full visualisation comprising all 45 countries, ordered by fertility in the last observed period.

**C04: In Figures 1-3, the grid lines seem too dense and distracting from the message; likewise, in Figure 4 there seem to be too many tick points.**

Response: The gridlines have now been removed.

Response: The size of figure 4 has been increased slightly, from 25x25cm to 26x26cm, to reduce the density of tickpoints. The manuscript will also include clearer links to a larger version of this figure for printing on an A3 printer or similar, as this print size seems more appropriate for a visualisation of this complexity. The larger version of this figure also includes full labelling of each countries, rather than use of HMD codes, to improve readability further.

**C04: The equation objects within the text need improving**

Response: All equation objects have been rewritten so that less information is presented in the form of subscripts, because these are harder to read. For example  would now be written as .

**C05: all the visualisations need to be double-checked from the point of view of being colour blindness/CVD-friendly."**

**C06: We (the guest editors) agree that optimising the figures for colour blind readers is a valid point, but we also realise that this may not always be achievable. The most common type of colour blindness is the red/green weakness, meaning people cannot distinguish red from green. Hence, it's a good starting point not to use red and green together in the same figure. Please check if you could adjust (without too much extra work) the colour scheme accordingly.**

**More details on "How to Optimize Charts For Color Blind Readers Using Color Blind Friendly Palettes" can be found here:** [**https://venngage.com/blog/color-blind-friendly-palette/**](https://venngage.com/blog/color-blind-friendly-palette/)

We thank the reviewer and guest editors for these additional considerations regarding use of colour. We have changed the main colour scheme used for representing age-specific fertility rates from RColorBrewer’s ‘Spectral’ to ‘Paired’ colour scheme, and then further modified the Paired colour scheme so as to be more CVD friendly.

Whereas the Spectral colour scheme involves using changes in hue to indicate changes in values, and uses red and blue hues throughout the majority of the palette, the Paired colour scheme makes use of both hue and lightness to indicate such changes.

The standard Paired colour scheme is a palette with the following sequence of colours:

Blue, green, red, orange, purple, brown.

We have modified this colour scheme such that green and red are no longer adjacent in the sequence:

Blue, green, purple, orange, red, brown.

This is because, as you note, the most common forms of colour deficiency are red-green. Using the modified version of the Paired colour scheme these colours are no longer adjacent to each other, and instead will be separated by colours that are easier to distinguish.

The Paired colour scheme involves alternating between desaturated (‘lighter’) and saturated (‘darker’) versions of colours with the same hue:

Light blue, dark blue, light green, dark green, light purple, dark purple, light orange, dark orange, light red, dark red, light brown, dark brown.

This means that, even in the extreme case of complete colour blindness, the colour scheme will still be somewhat interpretable as a monochrome series of alternating light-dark-light-dark bands. Given that age-specific fertility schedules are usually unimodal - first increasing with age, then decreasing with age – it should be apparent even in monochrome whether these alternating light-dark bands represent increases (‘ascents’) or decreases (‘descents’) along the value surfaces.

Although the Paired colour scheme is intended for representing categorical data, we have found it useful for supporting relatively accurate colour-value lookups using the corresponding colour scheme. This is in contrast to, for instance, using brightness alone to represent a value, where a given shade of brightness will tend to be perceived as brighter when surrounded by duller cells, and duller when surrounded by brighter cells (as illustrated by the Checkershadow Illusion).

Whereas gradations in value using the ‘Spectral’ scheme are indicated primarily through shifts in hue from red to blue, which would be difficult from someone with the most common forms of CVD to distinguish between effectively, the ‘Paired’ scheme employs both hue and brightness in order to distinguish between values.