Author's response

Referee's comment #1

The paper presents an interesting case study for Lenton et al.'s positive social tipping point framework.

The manuscript is well structured in the beginning but in total, too lengthy. Especially part 4 and 5 could be shortened and lack clarity in several aspects. It is for example not entirely clear why reuse systems are better than recycling systems. What are the measures of a 'good' system? Economic efficiency, the carbon footprint or the return/recycling rate? A clearer definition would help the reader to understand the framework.

In more detail:

The problem of a decreasing share of reusable glass bottles it not clear to me when the alternative is a plastic bottle recycling system. According to this study, e.g., recycled PET is more sustainable across a range of indicators compared to reusable glass: Stefanini et al., 2021. Plastic or glass: a new environmental assessment with a marine litter indicator for the comparison of pasteurized milk bottles.

Given that the advantage of reuse glass bottles over recycling PET is not clear, arguments in the discussion section might also not hold true. 'Policymakers need to regulate to address the market failure' regarding reuse system could become a misguided policy intervention if the assumption that reuse is better than recycling is not certain anymore.

Figure 3: Does ESD accept previously published figures?

Figure 5: What's shown on the x-axis? And how is stability defined?

Figure 6 is confusing in several aspects:

- The axes are missing, especially the y-axis. It is also not clear what reinforcing and balancing refers to here.
- In the example, R2 (innovation and technology intervention) happens before the social contagion effects, which is contrary to what you describe in Fig 4.
- How can a reinforcing and balancing feedback loop have the same direction on the y-axis?

Figure 8: similar to Fig. 6, the axes are missing. While it seems that the x-axis (time) is consistent between the two figures, the y axis is changing which is confusing to the reader. In addition, in Fig 6 the downward trend seems to be something desirable in terms of efficiency and sustainability (or whatever the authors intend to show here), whereas in Fig. 8, the downward trend seems to be something 'negative'.

R6: How can network effects (R3 in Fig 6) and a lack of network effects both be a reinforcing feedback loop?

R7: Similarly, it is counterintuitive that the lack of something is a reinforcing feedback mechanism.

I recommend using one single variable for the y-axis in Fig 6 and 8, which corresponds to the x-axis in Fig 5.

As statistics of reuse glass bottle and recycling PET bottle use seem to be available, the figures would gain significant power if the shares were shown quantitatively in the figures to demonstrate tipping effects over time.

It does not make sense to describe B2 before R10 if R10 happens before B2 according to fig 8.

In the discussion part, the difference between 5.1 (enabling/destabilizing conditions) and 5.2 (feedback loops) is not clear. In general, the discussion section could be shortened, given the great length of the manuscript. Policy options in 5.3 are partly covered in 5.1 already and it is not clear why an additional section on business and policy options is needed.

Overall, the empirical example of glass bottle reuse systems in Germany is an interesting illustration of Lenton et al.' s PTP framework but it requires considerable improvements, especially regarding the clarity of the figures.

Authors' response:

Thanks for your interest in this manuscript and your detailed review, which helped to improve it significantly. Below, we respond to your comments and suggestions one by one, where the paragraphs in cursive are your original comments:

The paper presents an interesting case study for Lenton et al.'s positive social tipping point framework. The manuscript is well structured in the beginning but in total, too lengthy. Especially part 4 and 5 could be shortened and lack clarity in several aspects.

RE: Thank you for the invested time and effort in providing the valuable and much appreciated feedback. We can see the point the reviewer is making. As with historical cases, this case is very rich and there is much to say about it. However, we will shorten our discussion of the case narrative (section 4 - results) by means of providing a short introduction and condensing the remainder into a table accompanying the images of both tipping phases. This way, we can move some of our observations about the insights into tippings this generates from section 5 (discussion) to section 4. At the same time, in section 5, we can focus more on the relationship of tipping points with the traditional timelines reserved for such transitions and choosing time appropriate timescales to study change (i.e. although tipping itself in this case happened relatively fast, it seems that other slower changes also played a role) and the role of interventions that initiate tipping (i.e. how specific interventions managed to bring about a sudden acceleration). Particularly the latter raises an interesting question for further research: if and how 'tipping' actions can be recognised or designed given a certain set of circumstances. We are exploring if this can be linked to other concepts such as 'the adjacent possible theory' by Kaufmann to provide concrete directions for further work (e.g. how to recognise the constraints of the situation and use this as a basis for interventions). All in all, we think that these changes will both shorten sections 4 and 5 as well as deepen the contributions from this study.

It is for example not entirely clear why reuse systems are better than recycling systems. What are the measures of a 'good' system? Economic efficiency, the carbon footprint or the return/recycling rate? A clearer definition would help the reader to understand the framework.

RE: The reviewer has pointed out something that we see can cause confusion. Please let us clarify: our study focuses on understanding the dynamics of tipping from one paradigm (or one system state) to another. The quantitative assessment of which of these states is the "most sustainable" we consider out of scope for this work, as it is highly context dependent, the domain of active scientific inquiry, and our paper does not focus on Life Cycle Assessment or the quantification of environmental impact. Although, as we mention in the paper, there is evidence that points to the system under study comparing favourably to single use systems (based on circulation rate and distance travelled), and it is also the assumption of many current policy interventions in the circular economy domain that aim to increase reuse rates. Our objective instead is to focus on understanding tipping dynamics in a qualitative manner and we intend to revise the manuscript to highlight this better.

In more detail: The problem of a decreasing share of reusable glass bottles is not clear to me when the alternative is a plastic bottle recycling system. According to this study, e.g., recycled PET is more sustainable across a range of indicators compared to reusable glass: Stefanini et al., 2021. Plastic or glass: a new environmental assessment with a marine litter indicator for the comparison of pasteurized milk bottles. Given that the advantage of reuse glass bottles over recycling PET is not clear, arguments in the discussion section might also not hold true. 'Policymakers need to regulate to address the market failure' regarding reuse system could become a misguided policy intervention if the assumption that reuse is better than recycling is not certain anymore.

RE: The rapid establishment and almost equally fast undoing of the initially successful change described in our case study is relevant for current circular economy change efforts that are the focus of many current environmental policies, the EU among them. Many rapid transitions are pushed for in this area, including packaging. Here, the objective is often to move away from recycling and towards increased reuse. However, the reception is not always positive and these efforts so far do not always show the desired results: there is critique on the usability of solutions , on how this – perversely – leads to increased plastic usage, and how return rates fall behind expectations. This undermines reuse systems and it raises the question of whether they can be expected to be permanent or that they will – like in our case – be replaced by single-use systems once more. Given the current investment of effort and resources in reuse systems it therefore seems prudent to learn from similar efforts in the past to see if lessons can be learned that can inform these current change efforts; hence the choice of our case study. However, as

mentioned in our reply to the previous comment, we regard the environmental impact assessment to fall outside of the scope of our study. In this work we are interested in the tipping dynamics themselves, and by unpacking and better understanding this to contribute both to a better understanding of tipping dynamics as well as circular economic change efforts. As previously stated in the above, we intend to revise the document to make this focus explicit. In addition, it seems prudent to, also in the discussion and conclusion section, once more point out that the quantification of environmental impact is not included in this work - and that other or further work is needed to establish under which conditions reuse systems are also environmentally favourable. We will add that it is here, at the intersection of our work and these quantification efforts, that a more comprehensive picture can be built about reuse systems, how to bring them about and under what conditions they are environmentally favourable.

Figure 3: Does ESD accept previously published figures?

RE: Yes, as long the reproduction rights are secured, which is the case here. However, we intend to align the images so they are presented in the same style in the final document, which entails redrawing the existing figure.

Figure 5: What's shown on the x-axis? And how is stability defined?

RE: We will change the figures in question so that the x-axis is labelled as the "level of reuse". From lowest to highest level of reuse: Single-use recycling, individual company reuse, pool reuse. (Dynamic) stability in this context refers to the ability to resist perturbation and recover from it, which is visualised as steepness of the valley sides and the height of the hill that has to be overcome. We will include this more clearly in the text, already where we explain the positive tipping points framework, so it is clear to the reader when subsequent images are introduced.

Figure 6 is confusing in several aspects:

- *The axes are missing, especially the y-axis. It is also not clear what reinforcing and balancing refers to here.*
- *In the example, R2 (innovation and technology intervention) happens before the social contagion effects, which is contrary to what you describe in Fig 4.*
- *How can a reinforcing and balancing feedback loop have the same direction on the y-axis?*

RE: Many thanks for your careful scrutiny of the images and your attention to detail. We agree that this figure needs to be revised. Firstly, the green arrows in this version are inadvertently mislabelled as 'destabilising' dynamics - they should be 'enabling' dynamics. Apologies for any confusion this caused.

The reinforcing feedback loops refer to the dynamics that reinforce the tipping towards a new system. The balancing feedback loops refer to the dynamics that stabilise the new system. We will make this clearer in the text in section 3 where we will say more about how we analysed the data and the key of the figure.

R2 refers to "economies of scale and reuse", which is in the correct order in the figure. The innovation and technology intervention you refer to is not shown in the figure, as interventions have been excluded from the figure to reduce complexity and visual clutter. However, we see the value in adding those details and will include all identified enabling/destabilising conditions/dynamics, feedback loops and interventions and will add these concrete interventions to the image so that they have a clear place in the timeline, yet do not confuse the tipping dynamics overview.

Figure 8: similar to Fig. 6, the axes are missing. While it seems that the x-axis (time) is consistent between the two figures, the y axis is changing which is confusing to the reader. In addition, in Fig 6 the downward trend seems to be something desirable in terms of efficiency and sustainability (or whatever the authors intend to show here), whereas in Fig. 8, the downward trend seems to be something 'negative'.

RE: Y-axis is still "stability" as in Fig. 5, both axes will be labelled accordingly.

'Down' in the tipping points framework merely represents a new system state. For tipping phase 1, we have chosen to represent the change from the individual company reuse system to the pool system as a change from left to right and from up to down to indicate the dynamic nature of the 'valleys' and 'walls' of system states. For tipping phase 2, to emphasise a reversal or undoing of the pool system we have chosen to represent it as a reverse movement compared to the tipping phase 2. We will emphasise this better in both the textual explanation of the tipping diagrams as well as in the visuals.

R6: How can network effects (R3 in Fig 6) and a lack of network effects both be a reinforcing feedback loop? R7: Similarly, it is counterintuitive that the lack of something is a reinforcing feedback mechanism.

RE: A reinforcing feedback loop can operate to amplify change in either direction. Network effects and Social contagion reinforced the tipping towards pool reuse in phase 1 (the more companies cooperated, the more attractive it became for others

to join), while the lack of network effects reinforced the tipping towards single-use recycling in phase 2 (the fewer companies participated in pool reuse, the less attractive it was to join the pool system). Reinforcing (positive) feedback can always operate to amplify change in either direction. In one direction increasing network effects leads to further increasing network effects. In the other direction breakdown of network effects causes further breakdown of network effects.

I recommend using one single variable for the y-axis in Fig 6 and 8, which corresponds to the x-axis in Fig 5.

RE: We will keep the y-axis consistent for "stability" and the x-axis for "levels of reuse" in the figures.

As statistics of reuse glass bottle and recycling PET bottle use seem to be available, the figures would gain significant power if the shares were shown quantitatively in the figures to demonstrate tipping effects over time.

RE: Thank you for this input, we will do so where available.

It does not make sense to describe B2 before R10 if R10 happens before B2 according to fig 8.

RE: Right, we will rework this so that the order of the dynamics in the text fits accordingly to the figures.

In the discussion part, the difference between 5.1 (enabling/destabilising conditions) and 5.2 (feedback loops) is not clear. In general, the discussion section could be shortened, given the great length of the manuscript. Policy options in 5.3 are partly covered in 5.1 already and it is not clear why an additional section on business and policy options is needed.

RE: We agree, and will revise as indicated in the above. This means that we will rework and restructure the discussion section to be less repetitive, clearer and shorter.

Referee's comment #2

Overall, based on a sustainable transition perspective, the paper provises an interesting case study on historical account of rise and fall of a bottle reuse system in Germany. The paper in general is well-written and clear in its description. At the same time the reviewer felt that the chapters 4 to 6 are too lengthy. Especially, the discussion in the chapter 4 tends to be a desriptive account of historical development of the bottle reuse system rather than an analysis on how sustainable transition was

established and failed (or taken over by another system). Framework presented in figure 4 is reasonable. However, the storyline of connecting chapter 4 and 5 via the framework presented in Figure 4 is rather weak.

Authors' response

Thanks a lot for the interest in our manuscript and your feedback. As indicated in the other comment reply above, with historical cases, this case is very rich and there is much to say about it. However, we will shorten our discussion of the case narrative (section 4 - results) by means of providing a short introduction and condensing the remainder into a table accompanying the images of both tipping phases. This way, we can move some of our observations about the insights into tippings this generates from section 5 (discussion) to section 4. At the same time, in section 5, we can focus more on the relationship of tipping points with the traditional timelines reserved for such transitions and choosing time appropriate timescales to study change (i.e. although tipping itself in this case happened relatively fast, it seems that other slower changes also played a role) and the role of interventions that initiate tipping (i.e. how specific interventions managed to bring about a sudden acceleration). Particularly the latter raises an interesting question for further research: if and how 'tipping' actions can be recognised or designed given a certain set of circumstances. We are exploring if this can be linked to other concepts such as 'the adjacent possible theory' by Kaufmann to provide concrete directions for further work (e.g. how to recognise the constraints of the situation and use this as a basis for interventions). All in all, we think that these changes will both shorten sections 4 and 5 as well as deepen the contributions from this study.