The manuscript, “Colonisation debt: when invasion history impacts current range expansion”, aims to test whether the environmental conditions of patches behind the invasion front alter invasion speed, creating a “colonisation debt”. Using a simulation model together with experiments, the authors find that the colonisation debt occurs when there is positive density dependence in either dispersal or growth, and that the effect of this on invasion speed depends on the direction and strength of the environmental gradient. The manuscript addresses interesting questions about the role of density dependence during spread with environmental heterogeneity.

While overall I found that the manuscript makes a novel contribution to our understanding of population dynamics during range expansion, I’m less convinced that the authors’ interpretations and framing of their results is appropriate given some of the choices made in their model. At minimum, a reframing of the results is warranted, but some additional simulations that decouple the gradient length and steepness may further strengthen the results. I would also have liked to see the experimental results presented more thoroughly, and worked more clearly into the discussion. I have attached some major and minor comments that I believe should be addressed prior to publication.

Major comments:

1. The authors have modelled a periodic landscape as a series of successive gradients, where both the steepness and length vary. In their results, the short / steep gradients are strongly influenced by the opposite gradient behind the minimum or maximum carrying capacity patch. In my view, it is not possible then to attribute any changes in invasion speed to the gradient steepness or direction in these cases. It seems like the authors get around this in discussing the role of the gradient direction only in the shallower gradient scenarios, but this is not clearly specified. Even in this case, the influence of positive density dependence is much stronger on the upward patches just after the minimum carrying capacity patch. Some additional simulations that explore this effect without the descending gradient behind might help tease apart the role of the gradient, a low carrying capacity core and the periodicity of the environment. At minimum, the possibility of each of these could be further explored. As is, I’m not sure calling these “gradients” is quite right. This language makes me expect you are modelling an environmental gradient, as in a range shift for example.

2. The experiment got lost a bit in the results and discussion. As is, I’m not sure the inclusion of the experiment is fully justified. Some stronger connections between the experimental results and simulations are warranted, and perhaps including some further analysis (if possible) using similar methods as the simulations. For example, looking specifically at the role of gradient steepness. It is unfortunate that the authors did not also include a treatment with a shallow, upward gradient as the difference between the up- and down-ward gradients seems like an important result, and with both gradients the difference between periodicity and a straight gradient could have been studied further.

Minor Comments

Abstract
3 - I'm not sure I agree with this assertion, at least this is not the only way that density dependence can affect invasions.

15 - The switch in terminology here to "extinction debt" is confusing.

Introduction
- The first two paragraphs here lack focus, and may be too general for the audience that is likely interested in these results. Consider revising down to one paragraph.

35 - this sentence is unclear, consider revising. Do you mean that density dependence causes emigration from the core and therefore expansion?

46 - consider removing one of "however" and "yet"

51-52 -I don't agree with this assertion. In at least some of these papers, carrying capacity is not explicitly modelled but is an emergent property of the population dynamics. So reducing growth also changes the carrying capacity. The point on 52 about binary patches is more meaningful and accurate.

54 - Again, this is not necessarily true. You can have variable carrying capacity emerge from population dynamics alone. Consider hedging here, at least.

62 - I'd like more here about what the significance of this relationship is. The authors do a nice job of this in the conclusion, it could come a bit here too.

65-68 - Wouldn't this depend strongly on the mode of dispersal and/or dispersal distances? With limited dispersal distances or local dispersal and a downward gradient, populations might saturate habitat close to the core quickly and then get stuck by the low carrying capacity patches? I know this is not what the results show, but I wonder if there is a threshold carrying capacity below which there can be pinning without long distance dispersal?

69 - Would the ability for populations to grow large at the front lead to accelerating speed with positive density dependence? There is previous research on acceleration with evolution, but here it's shown that acceleration can be just from positive density dependence in heterogeneous landscapes. This is an interesting result that I think is worth elaborating on here and in the results / discussion.

73-75 Does this mean that population dynamics at the edge might reflect habitat quality in the core, vs. just at the edge, with positive density dependence? Or that the previous patches matter more than the current patch? Consider clarifying here.

83 Some justification for this choice of landscape would be helpful. Why not just a single gradient in either direction, which would mimic environmental gradients in nature?

86 Consider adding a reference here.

Methods
Some additional info about growth and dispersal might be worth adding to the main text, in particular the authors should make it clear that dispersal is local and the positive density dependence comes in the form of an Allee effect for growth and changes in dispersal probability for dispersal.

124-36 - I'm curious about the different measures of speed. Are they correlated? I would expect the landscape level speed to average out across simulations over time, but early on it should be correlated with the gradient speed? The local measure of speed taken at the midpoint was interesting, but I didn't see this come through in the results.

143 - I'm curious why the authors chose the downward landscape and not to include an upward gradient as well?

Results
It would be nice if the experimental results came through more. I'd like to see a figure from the experiment and perhaps some further analysis linking experimental and simulation results. I also wondered about including a figure showing speed as a function of carrying capacity in the current patch vs. the patch directly behind.

164 - Is this just because of the periodicity in the landscape? Do changes in speed due to q average out as carrying capacity goes up and down?

Figure 3 – Consider making the lines / circles / crosses more prominent

Discussion
In other research, for example Garnier & Lewis 2016 (in the Bulletin of Mathematical Biology), it is shown that environmental gradients can mimic positive density dependence in that they can create pushed waves. I thought it was interesting that here the biological mechanisms induced something beyond what we see with just the gradient. I wondered whether the density independent simulations still spread as pushed waves in the gradients? And would be interested to hear the authors interpretation of their results in this context in the discussion. With longer gradients, more akin to environmental gradients studied with range shifts, would we see something different?

218-21 - I'm not sure I follow this, consider rewording and expanding on this explanation. I would expect that with density dependent dispersal populations would establish in the smallest patches, but then stall there due to low dispersal. Whereas, with density dependent growth populations might lag in establishing at all in the smallest patches. I don't think this is what was found though?

224 - I found this confusing, the shallower gradients are also the longer ones, is that correct?

245 - There's an extra word here.

286-91 - Really interesting connections made in this section.