Reviewers' comments:   
  
Reviewer #2 (Remarks to the Author):  
  
I thank the authors for their responses to my queries and suggestions and for the thorough revision of their paper. I found the revised version much clearer and easier to follow, partly because of the significant rewriting of large sections of the manuscript, and partly because of a more consistent use of terminology in relation to ecological time lags. I do have a few additional comments which I hope the authors will find useful to further improve the quality of their paper:

The authors thank the referee for noticing the effort that has gone into this revision.   
  
L111-119 – I fail to see how this grassland example of land cover change 'type' is different from the example of 'directionality' given for forest & grassland in L120-122 below. Please reword/clarify.

We have reframed the paragraph to highlight the three main factors that appear to modulate legacy effects: the type, amount, and directionality of land cover changes (L106-122).

L180 – ‘the legacy of the past landscapes on the current effective number are dependent…’ effective number OF SPECIES?

Corrected.

L198 – ‘avian communities are expected to re-equilibrate in the near future…’ – I would add ‘assuming no further changes in land cover’.

A valid point, now amended.

Also, unclear what is meant here by 'in the near future' – within 15 years as assessed in your study? This is only a short timeframe / snapshot of biodiversity change and it may take much longer for biological communities to pay extinction debts / colonisation credits and reach a new equilibrium.   
This is a good point - thank you. Indeed, with our current approach we cannot estimate how long these debts/credits will take to realise. Therefore, we have rephrased this paragraph and avoided discussion of the duration of any relaxation period.

L251 – Time of day included as covariate in the legacy model, but not in the equilibrium model? Why?

The written explanation was incorrect, and we apologise for the confusion caused. Time of day was included as a linear effect in the static covariates function, , which is included in both the legacy and the equilibrium models.

L408 – Is p<0.001 the significance value of the Moran's I, or is it the threshold used to assess significance? The former would indicate spatial autocorrelation (i.e. spatial distribution not random). Please clarify.

We have corrected this. Specifically, the p-value was higher than 0.05. We have also included additional information on the spatial scales tested for Moran’s I.

L493 – ‘four different buffer options’, but in L293 you mention six buffer 'options': a) a 4000m and 6000m circular buffer around the centroid of the polygon defined by the vertices of each segment, b) a 500m, 1000m, 2000m and 4000m buffer around the segment line. Please clarify.

The information at line 293 was from a previous iteration of the analysis presented in the first submitted version of the manuscript. The corrected buffers that we tested are described at line 493: a) a circular buffer around the centroid of the polygon defined by the vertices of each segment (4000m radius), b) a buffer around the segment line (500m, 2000m and 4000m radius). This is now also standardised between statements.

L484 – ‘in order TO identify…’  
Corrected.

Reviewer #4 (Remarks to the Author):  
  
This study uses data from the North American Breeding Bird Survey (BBS) to calculate extinction debts and colonization credits based on 15 years of land use change in the continental United States. The paper uses a creative modeling approach and vastly extends our knowledge of ecological time lags by using multiple land cover types at a continental scale. The existence of debts and credits after only 15 years of change, despite all the noise in the data, is compelling evidence for their conclusions.   
  
The authors have done an appropriate job addressing concerns about uncertainty in the results. I commend this effort, but I have some remaining suggestions that need to be addressed. These suggestions will help make the results more transparent while also helping readers correctly interpret the spatial data in Figure 1.   
  
(1) The authors use appropriate posterior simulation methods to assess uncertainty in predictions of debts and credits across the U.S (Figure S7). However, the way that these results are presented still lead to some concerns that likely stem from a lack of clarity. Therefore, my first suggestion is that the authors include more information in the text on how the coefficient of variation (CoV) was calculated for Figure S7. What was the exact equation used (e.g., sd / mean, or another equation such as the Geometric CoV)? Based on existing literature, how are you interpreting ‘low’ versus ‘high’ uncertainty based on these values?

We have added information, in text at line 469 and in figure S7 caption at lines 648-652, to explain the CoV used, and have provided the equation.

(2) Figure S7 nicely addresses the concern that there could be large spatial biases in the results. Yet, to the reader, uncertainty in estimates appears to be incredibly high. If the conventional sd / mean equation was used to calculate CoV, then the uncertainty values on the map are unacceptably high. For example, a CoV of 10 means that the standard deviation is 10 times larger than the mean, and classifying a hexagon with that much uncertainty as a “credit” or “debit” is not appropriate. The scale in Figure S7 reaches values well above 100, which leads me to believe that some other equation must have been used, but readers are currently in the dark and unable to interpret these numbers.

The CoV used in the previous version was the standard CoV of sd / mean, but we agree that this should have been made explicit. We have now done this (469 and 648-652). We also agree that our plot was not completely clear in its presentation of uncertainty. The presence of large CoV numbers was caused by a number of predictions resulting in debts and credits of magnitudes of the order of 10-7. This arose in areas of minimal landscape change, leading to mean values that were very close to zero. The propagation of uncertainty over these values resulted in large coefficients of variation for those areas. It also came to our attention that in the code we had rounded down certain intermediate outputs, causing small values of debt/credit to be zero. We have now removed this rounding. Furthermore, to clarify this and ease interpretation we have recalculated our uncertainty map, now using the geometric coefficient of variation and included explanatory text in the figure S7 caption, as well as upper and lower credible intervals maps of the same figure.

(3) While I appreciate the use of CoV to represent uncertainty, I do think that maps showing upper and lower quantiles of the expected value would be incredibly useful to readers. Given that you have already written a script to propagate uncertainty via posterior simulation and calculate a per-hexagon CoV, it should be fairly straightforward to save the median and quartiles from each hexagon as well. Figure S7 could then become a 3-panel map showing CoV along with the lower and upper bounds of estimates.

Thank you for this additional suggestion. We have now included panels to Figure S7 to show the lower and upper credible intervals of our predictions and facilitate interpretation. The intervals were calculated as the lower (0.025) and upper (0.975) quantiles from the thousand values predicted from posterior draws.

In addition, I have several minor comments:  
Currently, the wording “2880 bird communities”, which occurs throughout the manuscript, seems ambiguous to readers. I can infer that 960 BBS routes x 3 sections along each route = 2880 different survey units. But referring to these sampling units as “communities” without an explicit definition in the main text is misleading. Some people might interpret a “community” as a single assemblage of birds associated with a particular habitat (e.g., wetland bird community), rather than a specific sampling unit. Please provide a definition and added clarity to this phrase (preferably in the results section instead of buried in the methods).

We have added a clarification on this point at lines 60-62: “We identified a community as the assemblage of birds associated with the landscape surrounding each survey unit (i.e. not a prespecified habitat type).”.

In several places (e.g., lines 80, 143, 147, etc.) regions of the U.S. are referred to as the “North East” or “South West”. To remain consistent with previous literature, I suggest changing these to a single word, e.g., “Northeast” or “Southwest”.

We have made these changes as suggested.

37: I suggest rewording the sentence starting with “Notably…”. Many would argue that species responses to habitat change at a local scale are in fact instantaneous, as when birds respond to deforestation or wildfire. Here, non-instantons changes refer to gradual changes in species composition at the broader landscape scale.

Rephrased accordingly.

63: Typo: “as a biodiversity metric”

Corrected.

125-126: There are typos in the title for Figure 2. Consider rewording to something like: “The contribution of past landscape to the current effective number of species depends on the type, amount, and …”

Corrected.

139: Typo: Please change “associated to” to “associated with”.  
Corrected.

141: I suggest rewording this sentence to provide more clarity. Ex: “…much of the area in the central United States has experienced…

Adjusted.

Figure 4. I commend the authors for propagating uncertainty into these estimates.

We thank the reviewer for this comment and do agree that this, together with your suggested upper and lower boundaries figure, was an important missing element.

209: Incorrect package name: “exactextractr”

Corrected.

532: This line references JAGS model code, but the main text says that models were run using STAN.   
Corrected to clarify that models were run using STAN.