Summary

The authors investigated the presence of a potential bet-hedging strategy in the parsley frog *Pelodytes punctatus*, in which species a bimodal breeding phenology is observed. They use both empirical observations and theoretical modelling to test whether or not this phenomenon could indeed be a bet-hedging strategy, which dual approach makes this study quite robust and valuable. Overall, the manuscript is well written and concise, with only a very few errors. The introduction lays out the theoretical background neatly, and presents the study system well. Empirical knowledge on bet-hedging is still limited, and the authors did an excellent job highlighting the relevant knowledge gaps, and the manuscript made a substantial contribution to the scientific literature on the subject. The sampling methods appear to be reasonable and justified. The results are communicated mostly clearly, and the discussion was focused, avoiding over-arching conclusions or speculations. Alongside with the main findings, I think a great strength of the presented results is that it provides strong empirical support for the notion that shorter lifespans increases the selection pressure for mixed strategies. Also, the evidence of inter-cohort competition is a unique and very valuable finding, which puts emphasis on a yet scarcely documented and understood phenomenon with potential relevance to the understanding of the evolutionary ecology of a wide range of species where cohort splitting occurs.

I only noted a few major comments, which should be addressed in my opinion, in order to improve on some segments. Specifically, one of the tested measures (offspring survival from egg to metamorph) should be revised, or at least given a firm rationale and explanation (see Major comments: Methods). In addition, while most of the results are clearly discussed, I think a bit more detail on the likely (autecological) advantages and disadvantages of each (autumn vs. spring) strategy would greatly improve the concluding paragraphs of the discussion, helping readers to appreciate more the nuances of the presented system.

Major comments

Methods

193-198: I might be just misunderstanding something, but why is it necessary to take the product of hatching rate and metamorph rate? The latter seems to need the former to take place, i.e. they are not independent events. It seems that the former is a time interval like t1-t2, and the second is t1-t3, so taking their product may not be necessary, as the second already takes into account hatching success (i.e. will always be less than (or, at best, equal to) hatching rate). Hence, overall metamorph success [metamorph/eggs] seems sufficient as a proxy for larval survival. Or, alternatively, juvenile survival (using the above time nomenclature: t2-t3) is simply [metamorphs/tadpoles]. This should be clarified, and if needed then corrected, as it would affect some parts of the conclusions (e.g. Discussion l:386).

264-266: “Overall the mean number of individuals produced per female is c s1 when the autumn cohort doesn’t fail and $(1 - c)$ s2 when it does” – wouldn’t that be the expected proportion of offspring reaching sexual maturity, rather than the “mean number of individuals”?

371: “The breeding effort in our population was higher in autumn than in spring” – in the “Statistical analyses“ season was not mentioned as a predictor for the fitness-related dependent variables, nor was it mentioned in the “Explanatory variables”. In their current forms, these sections indicate that no models were fitted on breeding probability, breeding effort, etc., with season as predictor. I think it would be important to include such models (e.g. “breeding effort ~ season”) and present their results prior to describing results from the “pond-characteristics” models, as they would likely provide general insight into the dependent variables in question. For instance: how was breeding probability affected by season? This approach may help us to understand why such a bimodal phenology is apparent. Based on the reported results, it might seem that autumn tadpoles gain substantial advantages, but if there is higher variation in autumn whether or not a given pond survives (or, indeed, in larvae survival), this could indicate that throughout this risk spreading strategy, the relatively “low risk, low benefit” spring cohort represents a safer route, but the “higher risk, higher benefit” autumn cohort is a necessity to decrease variance in the long term (lineage, i.e. geometric mean) fitness, hence neither the strategies can dominate the other on the long run.

Results

In the results, I think it is important to report all estimates and statistics, even when non-significant, e.g. at lines 328-330. Alternatively, these could be also reported in table(s): continuous trends (if not in interaction with other variables) can be simply reported as appearing in the model summaries, whereas between-group differences (from categorical predictors) can be reported as estimated marginal contrasts (the R-package “emmeans” may be useful for this). Just in case, such marginal contrasts can be acquired as:

```
# model specification, where “x” is a categorical factor
m1 = lm(y ~ x)
# getting group-level comparisons and corresponding marginal estimates
emmeans(m1, pairwise ~ x)$contrasts
```

Discussion

455-465: while the **frequency**-dependence indeed appears to be supported by the findings, the cited reference of Gremer and Venable (2014) highlights **density**-dependence, which was not (explicitly or implicitly) modelled by the authors, therefore I’m not quite certain that this reference is adequate here. Alternatively, the authors could elaborate on how density dependence could also play a role in the observed patterns: for this, it would be useful to fit models on the fitness-associated variables (breeding effort, etc.) of the spring cohort in relation to the estimated conspecific density (i.e. pond-size-corrected number of parsley frogs [originating both from spring and autumn] of ponds). If I’m not mistaken, this was not modelled, only the presence/absence of conspecifics.

Minor comments

57: “genotypic” → “genotype”
90: “insect” → “insects”
113: “sensible” → “sensitive”
114: delete “than”
162: delete “)”
207: “undistinguishable” → “indistinguishable”
233: “apply” → “applied”
236: was the binomial model quasi-binomial here as well?
264: delete whitespace after “below”
307: does “spawning probability” correspond to “breeding probability”? If so, please unify throughout the manuscript.
308 and others: when writing mean and SE estimates, I think the “x ± y” format may be better (e.g. 0.18 ± 0.02)
332: no space between “%” and “(“, and there is a whitespace between “(“ and “0.61”
337-338: “From the point of view of spring breeders, in 28/57 cases, they found autumn tadpoles in the pond” – not quite clear: does that mean that in 28/57 cases, those females laying eggs in spring shared the ponds with autumn tadpoles? Please clarify.
338-339: “success of spring breeding event” → survival of spring tadpoles?

343: all stage-dependent survival rates for spring tadpoles, or for both spring and autumn tadpoles? In other words: only spring tadpoles were affected by such mixed ponds?
422: “predates” → “predate”
491: “and sensitive to environment , a phenomenon proposed” → delete whitespace after “environment”