We thank the reviewers for the detailed feedback. We will revise the final manuscript to address all the minor remarks, and answer to the main remarks and direct questions below.

Reviewer 1

a) It would be great to have an explicitly related work section. b) The linearization in 3.2 should be described to make the paper self-contained. c) The estimation of M_1 using VI should be expaned as well.

For the final version we expand the related work by citing recent variational approximation techniques deviating from direct ELBO optimization [1, 2], make the paper more self-contained by re-iterating the linearization process (following [14]), and expand the description of the procedure for estimating M_q (which we presume you referred to with M_1). Briefly, we first run standard VI for sufficiently many iterations (until the losses stabilize), and then compute the loss for every training instance. We then sort the resulting losses and set M_q to match the desired quantile of this empirical distribution of losses.

e) Also the recommended one also need to motivate why a continuous utility is needed. For this task, I believe that discrete unity can be used as well.

Discrete utilities are appropriate for recommender systems, but our demonstration is about predicting consumption volume, which is better addressed with continuous utility due to very large set of possible values, and because overall magnitudes matter much more than the exact counts.

Reviewer 2

Why can we assume smoothness / differentiability of the expected utility ?

We would like to note that smoothness is not a necessary requirement, simply something that helps the optimizer to converge faster. Based on our experiments and supported by the empirical findings of [23] the expectations of utilities with point-wise non-differentiability tend to be smooth, even though we do not have a rigorous proof for this.

Reviewer 3

The paper should address if there are ways to use unbounded losses (e.g., by switching from utilities directly to inference in loss-based posteriors). If there is a fundamental reason to use utilities instead, it would be good to have a thorough explanation of this reason.

The assumptions on utilities (and hence on losses) arise from the derivation of the optimization objective (Eq.1) that requires the log of the expected utility to be defined. In Section 3.2, we relax these assumptions and provide two practical ways of handling unbounded losses either by linearization of the utility-dependent term (Eq.7) or by a non-linear transformation (Eq. 3) that effectively compresses extreme losses into very small (but still non-zero) utitilies. In practice, the procedure seems to work well for unbounded losses as well.

One potential way of extending/adapting the method that comes to mind is using recent advances in loss-based/PAC/generalized Bayesian posteriors, see e.g. Bissiri, Holmes & Walker, '16 or the work on Generalized VI. I believe this could be possible by designing a kind of compound loss...

These works provide solid basis for modifying the posterior inference procedure directly and hence offer very interesting future directions for loss calibration as well. However, it is non-trivial to design such a compound loss that would directly link the variational objective to the exact definition of gain/risk of Bayesian decision theory. Our framework achieves this, and we also show (Fig 3c; linearized case) that constructing objectives ad hoc may lead to suboptimal results. We do not rule out the possibility of alternative loss-calibration strategies building on generalized posteriors/approximations, and look forward to papers proposing such techniques.

Why exactly the utility infimum should be 0?

Optimal decisions $\{h\}$ in Bayesian decision theory remain invariant under linear transformations of $u(.)$. This does not hold for the loss-calibration bound. Instead, the relative (vs. ELBO) importance of $\mathbb{U}$ is maximal for utilities (linearly) transformed such that $\inf u = 0$. Whenever $\inf u = \beta$; the bias $\beta > 0$ can be moved outside the integral and also the logarithm. The remaining term inside the logarithm is $1 + \frac{\textit{integral}}{\beta}$ that for $\beta \to \infty$ converges to a constant (e.g., $\nabla \mathbb{U} \to 0$) removing calibration completely. For $\beta \to 0$ the magnitude of the $\mathbb{U}$ term – and hence the calibration effect – is maximal. We typically want this, but $\beta > 0$ can be used for reducing the calibration effect if so desired.

References

[1] T. Rainforth et al. Tighter Variational Bounds are Not Necessarily Better. In *ICML*, 2018.

[2] Tao et al. Variational Inference and Model Selection with Generalized Evidence Bounds. In *ICML*, 2018.