We thank the reviewer for his/her detailed comments. We have revised the paper to address these comments. Our point-by-point response is given below in bold font.

Reviewer:
In this submitted work, the authors classify the KBOs in the MPC and use the results of this classification to constrain the midplane of the Kuiper belt. This is an interesting question, which can constrain the amount of additional masses in the solar system and may be better constrained with the current catalog of known KBOs. The midplane calculation method utilized here is a reasonable approach for using the MPC catalog, which has unknown discovery biases.

Unfortunately, the reviewer is not able to recommend this work for publication at this time. The selection of an appropriate sample of KBOs for a midplane analysis is critical and can significantly affect the results of the analysis.

The utilization of all KBOs in the MPC including single opposition objects is of particular concern. This paper notes that previous work (Gladman et. al 2008) recommends 2 or 3 oppositions, but claims that followup observations do not significantly improve estimates of the orbit's orientation. It is true that orientation is more quickly constrained than semi-major axis, however, this work appears to make no cuts in arc length at all. Short arcs and especially extremely short arcs (<7 days) will be biased based on the assumptions of the orbital fitting software used to generate the orbits. For example, Bernstein & Khushulani orbital fits for very short arcs favor near-circular solutions, and will often give the smallest inclination compatible with the discovery location on-sky as well as an eccentricity near zero and a semi-major axis uncertainty of +-20AU. Objects with a semimajor axis uncertainty of 20AU cannot reasonably be assigned to any semi-major axis bins of 2-3AU, and best case scenario these result in random noise added to any real plane measurement. It is likely, however, that there are additional concerns based on the typical way short arcs are fit and the likelihood of fits being increasingly poor as distance/semi-major axis increases (short arcs on slower moving objects are even more concerning). For context, of the current TNO and Centaur lists in the MPC, 729 are single opposition, 522 have <30 days arc, and 376 have <7 days arc. A spot check comparing a few short arc KBOs to the MPC list to this paper's included table identified multiple extremely short arcs, for example 2015 SP21 and 2015 SO21, which have 2 and 3 day arcs and 32 AU and 60 AU semi-major axes, respectively. These orbital fits should not be considered to be sufficiently constrained to include them in this analysis.

A related issue to using single opposition KBOs is the impossibility of identifying resonant objects. The two examples above would have ~20AU semi-major axis uncertainties, which would make both compatible with 3+ major Neptune mean motion resonances. As resonant KBOs are extremely common (for example, 40% of the Outer Solar System Origins Survey discoveries were resonant), we should expect that a large fraction of the single opposition KBOs are resonant. Resonant KBOs should be excluded from the midplane analysis, but cannot be properly identified for objects with large semi-major axis uncertainty. Previous work such as Volk & Malhotra 2017 used a semimajor axis uncertainty of 5% as a limit for inclusion in the dataset. (In this work, this uncertainty is miss-stated as an evolutionary change of 5% or scattering, but Volk & Malhotra 2017 refers to the uncertainty in the orbital fit.) The reviewer recommends that the authors implement a similar uncertainty limit on the orbital fit for KBOs in their analysis. The reviewer would also recommend that the authors consider a higher perihelion cut than 7.35 AU, such as Neptune's semi-major axis.

The authors also state that the Smullen & Volk 2020 classifier does not have the functionality to include uncertainty, which is correct in that it doesn't accept uncertainty as an input. Instead, as describe in that paper, they explore the uncertainty region by inputting multiple clones which explore the parameter space, which is computationally efficient using that classification method and should not require re-training. Utilizing multiple clones would also help to identify which KBOs are too near the resonance boundaries for classification of resonant/nonresonant. If the authors refine their selection criteria to repeat the midplane calculation, the reviewer recommends classifying multiple clones per object to ensure that resonant KBOs are properly identified.

The reviewer's concerns about sample selections and clones are well founded. We have revised our sample selection pipeline as follows, after the point where barycentric elements at 2021-11-15 have been retrieved from Horizons for each object:

1. **Select objects with 34.79 au < a < 150 au and q > a_Neptune.**
2. **Download 1-sigma semimajor axis uncertainties from Horizons Small Body Database Query. Keep only objects with sigma_a/a < 5%.**
3. **Cross-reference with Minor Planet Center database to eliminate objects observed for fewer than 3 oppositions.**
4. **Download 6x6 covariance matrix (e, q, tp, node, peri, i) from Horizons Small Body Database.**
5. **Generate 300 clones of each object from the 6x6 covariance matrix and classify each. Discard objects where >50% of clones are Resonant.**
6. **Assign remaining objects to semimajor axis bins based on nominal orbital elements and compute midplane uncertainties as in the original manuscript submission based on nominal orbital elements.**

Data Editor's review:
One of our data editors has reviewed your initial manuscript submission and has the following suggestion(s) to help improve the data, software citation and/or overall content. Please treat this as you would a reviewer's comments and respond accordingly in your report to the science editor. Questions can be sent directly to the data editors at data-editors@aas.org.

The .csv file for the full Table 1 data set is fine. It will be converted to machine readable format for you at acceptance (although if you want to try this yourself we do have an online tool to help at https://authortools.aas.org/MRT/upload.html).

I would suggest that you review the number of significant digits in this table. For example, is the semi-major axis for the first source known to 8 digits (43.93181826 AU)? You only report 2 significant digits in the example table. A new version can be uploaded with the revised manuscript.

Because we do not fit any orbits or generate any orbital uncertainties ourselves, there is no

need to include a table of orbital elements. Instead, we now present in the body of the paper only our categorical classifications for a sample of five objects. The revised Table 1 will include only the name and classification of each object in the main sample set, together with the number of clones that are classified as Resonant, Classical, Scattering, and Detached. For the sake of reproducibility, we will archive on Github and/or Arxiv the database query results used as inputs for our computations, as well as the source code used to produce the results. We have more carefully examined the accuracy of our results for the midplane inclination and longitude of node in each semimajor axis bin, as well as their upper and lower bounds, and choose to report inclinations to the nearest tenth of a degree and longitudes of node to the nearest whole degree.

Statistics Editor report :
Sec 4 describes a 'parametric bootstrap' procedure for estimating the uncertainty of the data-based Kuiper Belt mid plane estimate. But that is not the right term. The parametric bootstrap draws samples from a statistical model, often obtained from the data itself by maximum likelihood estimation (Wikipedia Bootstrapping_(statistics)#Parametric_bootstrap). But here the authors take some variables from the data, and choose constrained-random values for other variables ... there is no statistical model. Then they reject many candidate object, accepting one synthetic object for each real object based on an arbitrary "acceptable match". The authors' procedure really has no name; they are creating randomized datasets in an idiosyncratic fashion. Therefore, the theorems underlying the bootstrap do not apply (reviewed in Hall 1990, DOI 10.1016/S1573-4412(05)80008-X). This is a heuristic procedure and the term 'bootstrap' should not be used.

Our use of the term "parametric bootstrap" was a misunderstanding. We will no longer use that term. However, we don't believe it was entirely misapplied. Our method uses the following partial statistical model of KBO populations: uniform distributions in argument of pericenter and mean anomaly, q and p as independent uncorrelated Gaussian distributions with their mean and standard deviations taken from the data, and empirical distributions for semimajor axis and eccentricity, approximated by fuzzing randomly selected real objects.

The main concern of the Statistics Editor is that because the theorems underlying the bootstrap do not apply to an incompletely constrained statistical model, our heuristic procedure may not converge to reproducible results as the repetition count increases. Our revised manuscript includes a plot to show convergence of the inclination lower bound for one semimajor axis bin; convergence plots for other bins and other bounds are similar.

Sec 5 states that potential KBOs were identified with a "gradient boosting classifier from the Python package scikit-learn on 55 features from the first 100 kyr of the integrated trajectories". This explanation is inadequate for reader understanding and possible reproducibility of the calculation. First, the features used for a classifier should be presented, usually in a table. Second, the methodology should be described (at least in outline form) mathematically with details on the many options of the code (sec 1.11.4 at https://scikit-learn.org/stable/modules/ensemble.html#gradient-boosting). Graphical outputs should be shown and described to inform and convince the readers. References to gradient boosting methodology should be added such as Friedman 2001 (10.1214/aos/1013203451, 15K citations).

We emphasize that we did not invent the classifier. The 55 features used by the classifier and the methodology used to define it, along with ample graphical demonstrations of that methodology, are fully explained in Smullen & Volk 2020. We think it unnecessary to repeat that information in full for context. In the original manuscript, we examined four slightly different variations in the training set that could be used for the classifier and picked a training set slightly different from the final fiduciary training set used by Smullen & Volk 2020. To allay any concerns that this may render use of the classifier invalid without more extensive testing, our revised manuscript uses exactly the same training set as the final fiduciary training set in Smullen & Volk 2020. This means that we use their classifier as made available on Github without any changes whatsoever, and there is no need to present further information for reproducibility. We include clear and explicit language in the revised manuscript to direct readers with questions about the classifier and its reproducible use to Smullen & Volk 2020 and their Github page.