Thanks for all the comments and suggestions. We have carefully revised the manuscript according to these suggestions. Our point-to-point responses are listed below:

General Comments: 1. The trend quantification in Section 5 is quite underwhelming. It is void of any statistical significance testing, which is necessary to diagnose meaningful changes. Moreover, trends are diagnosed for relatively short time periods, which provides little confidence in the result and leads to overfitting where there is substantial year-to-year variability. Confidence intervals (e.g., 95 and 99 percent) would be especially helpful here to demonstrate the degree to which diagnosed long-term changes (and reversals from the first to second half of the time period) are meaningful. Without statistical evaluation here, clear conclusions cannot be made for the diagnosed trends and the use of the term “significant” throughout the paper is inappropriate. Reply: We add statistical significance testing in Section 5. Most ozonesonde trends in lower troposphere and mid-troposphere before 2012 passed the 95% significance criterion. Since there are fewer samples after 2012, some trends only passed the 90% significance criterion.

2. There are at least two claims based on comparison of observations and CLaMS output that are not justified based on the analyses conducted. First, the authors claim at lines 124-125 that CLaMS overestimates transport from the stratosphere to the troposphere. This is based on comparing ozone concentrations in the model to that observed and assuming a certain missing control by tropospheric chemistry. Without additional analysis (or citations to other more thorough evaluation), I do not find this claim to be justified based on the analysis presented in this paper. Second, at lines 261-266 it is argued that a reduction in stratospheric ozone found near the ozonesonde location in CLaMS is a result of ENSO, but there are certainly several alternative explanations for this change that are not acknowledged. Notably, Beijing is near the climatological mean latitude of the tropopause break (the sharp discontinuity in tropopause altitude from tropics to extratropics). Latitudinal migrations of the tropopause break could result in Beijing being more on the tropical side in later years, thus less exposed to downwelling stratospheric air. The latter can certainly be evaluated using the CLaMS output. Reply: We claim that CLaMS overestimates transport from the stratosphere to the troposphere based on not only the comparison between ozonesonde and CLaMS, but also a study by Konopka et al. (2019). Although the current transport scheme in CLaMS shows a good ability to represent transport of tracers in the stably stratified stratosphere, there are deficiencies in the representation of the effects of convective uplift and mixing due to weak vertical stability in the troposphere. We give more explanation here to make it clearer to understand. For the second comment,

3. There is substantial repetition in the discussion of the time-evolving role of NOx. Namely, a succinct analysis and discussion is given and then followed shortly after by a less clear rehashing of essentially the same points while pointing to other work – lines 151-155. Perhaps the authors intend to make a slightly different point, but this is not clear. Reply: We modified this part to make it clearer. First, we cite some studies which revealed the reduction of NOx in many places of China in recent years. Then, we use OMI data to show the long-term variation of tropospheric NO2 over Beijing. After talk about NOx, we were thinking to cite studies to show the change of other precursors. But now we move them to the Discussion and conclusion Section. In Section 4, we revised the description as “The variations in the precursors of tropospheric ozone have dominant roles in the long-term variability of tropospheric ozone. In recent years, the Chinese government has started to invest time and resources in controlling air pollution. A review of 20 years of air pollution control in Beijing (UN Environment, 2019) reported reductions in NOx during the period 2013–2017. A clear decreasing trend in NOx emissions has been observed since 2012 (van der A et al., 2017). Zheng et al. (2018) also reported that emissions of NOx in China decreased by 21% during the time period 2013–2017. Wang et al. (2019) reported that NOx emissions in eastern China decreased by approximately 25% from 2012 to 2016. Tropospheric NO2, one of the precursors of tropospheric ozone, has gradually decreased over Beijing in recent years (Vu et al., 2019). We use the tropospheric column of NO2 from OMI to discuss the influence of precursors on the long-term variation of tropospheric ozone in Beijing. The

deseasonalized tropospheric columns of NO2 measured by OMI from 2004 to 2018 are shown in Figure 5. Tropospheric NO2 was increasing from 2004 to 2010, especially in 2009, leading to the increase of ozone in lower and upper troposphere. As Chinese government start to control air pollutions, tropospheric columns of NO2 were in a condition of relatively large fluctuation in the period of 2010-2013. Tropospheric NO2 over Beijing experienced two major fluctuations in this period, as shown by Gaussian-weighted means. Then tropospheric NO2 was gradually decrease since 2013, result in the hiatus of ozone increase in lower and upper troposphere.” In the discussion and conclusions section, the description is revised as “The exchange of ozone between the stratosphere and troposphere is also important to bring ozone into troposphere (Dufour et al., 2010, 2015; Neu et al., 2014)”.

Specific Comments: Lines 34-35: delete “which transports stratospheric ozone into the troposphere and tropospheric ozone into the lower stratosphere” as it repeats the previous words in this sentence. Also, this sentence is incomplete. What about the exchange of ozone between the stratosphere and troposphere? Reply: This sentence is changed as “The exchange of ozone between the stratosphere and the troposphere is also important to bring ozone into troposphere (Dufour et al., 2010, 2015; Neu et al., 2014)”.

Line 66: “relative” should be “previous” Reply: “relative” is replaced by “previous”.

Line 74: The accuracy and precision of the ozonesonde data should be listed here. Reply: The mean difference in the ozone partial pressure between the IAP and ECC ozonesondes was <0.5 mPa in the troposphere and <1 mPa in the lower strato-
sphere. The correlation coefficients for profiles by IAP ozonesondes and the ECC are greater than 0.99 (Xuan et al., 2004). The total ozone columns measured by the IAP ozonesonde and the Brewer spectrophotometer were in good agreement with a relative difference of 6%. For the total ozone column, the relative difference and correlation coefficient between IAP ozonesonde and Brewer instrument were 6% and 0.94.

Line 86: It is not clear where the “D” comes from in the ASAD acronym. A quick Google search shows that this should be defined here as “A Self-contained Atmospheric chemistry coDe (ASAD)”. Please revise. Reply: We revise it as “A Self-contained Atmospheric chemistry coDe (ASAD)”.

Lines 140-148 and Figure 5: Details on OMI data used belong in the data and methods section. Reply: We give more details on OMI data in Section 2.

Figure 6: It is not clear what exactly is being done to/with the data. Are the time series based on a three-month average of monthly means or something else? Reply: No, the data is still monthly. That means we have 3 samples each year for each season. To make it clearer, we revised it as “Figure 6. Trends of monthly mean column ozone (DU) from the ozonesonde observations (black) and CLaMS simulations (red) in four seasons. There are 3 monthly values in each year for each season.”