Response to Referee #1

We thank all referees for their constructive comments and suggestions. We respond to the specific comments below in blue color, repeating the reviewer's comments in italics for reference.

1. General comments
The study uses a comprehensive AGCM-chemistry model to investigate how drastic measures to reduce US carbon dioxide emission by 50% by 2050 would affect both air quality and climate in the coming years. Furthermore, it incorporates air quality regulations together or separately from carbon emission reductions. While world leaders seek solutions to curb carbon emissions and prevent climate change during the Paris COP21 meeting, this type of articles comes well timed, in my opinion. The future of US climate and air-quality regulations is relevant, because the country is the largest historical climate polluter and US policies exerts global influence. In my opinion a key message of the article is that US climate policy would generate air quality co-benefits, even if no direct air-quality measures are taken, this would be already a clear win-win situation. Air quality regulations alone would have important health benefits, but PM reductions would lead to a global positive forcing, having negative climate consequences. If both air-quality and climate regulations are implemented, the health benefits are large and the negative climate effects from PM reduction is smaller than the positive climate benefits from carbon reductions on a global scale. I agree that US air-quality and climate regulations should be implemented together to address these two closely related environmental problems as the authors suggest. In my opinion the study is interesting and novel, but I am still missing three important things in the it:

I. There doesn't seem to be any attempt in the manuscript to evaluate the model's ability to correctly simulate short-lived climate pollutants. The comparison between a one-moment aerosol scheme and a two-moment aerosol scheme is definitely very valuable as it provides a measure of the model precision, but by no means it yields the uncertainty in aerosol modelling as the authors suggest in the conclusions. To evaluate uncertainty, models should be compared to observations. Can the model simulate present-day PM2.5 and ozone concentrations? Has someone compared GISS ModelE2 with any of its two aerosol schemes with observations? I think a section discussing these key issues is necessary.
Response) We have previously published a paper with the detailed evaluation and comparison between the both models (*Lee, Y. H., Adams, P. J. and Shindell, D. T.: Evaluation of the global aerosol microphysical ModelE2-TOMAS model against satellite and ground-based observations, Geosci Model Dev, 8(3), 631–667, doi:10.5194/gmd-8-631-2015, 2015.*). Since the details of the model comparison are available in Lee et al., (2014), we have added the following statements in the Section 3 to provide a short summary of model evaluation and differences.
"Ozone in the ModelE2 was previously evaluated in Shindell et al (2013), which found that around 900 hPa ozone tended to be overpredicted in the model by around 5-8 ppbv. Though ozone in this version of the model was improved at higher altitudes, values near the surface were similar to the prior ModelE, which displayed little mean bias relative to a network of 40 surface ozone measurements although the correlation was only R=0.7 (Shindell et al., 2006). The atmospheric residence time of methane in modelE2 is in

excellent agreement with the value inferred from observations, indicating that OH levels are also well simulated. Additional analysis of seasonal maximum 8-hourly surface ozone showed that the model captures the summertime observed levels in the western US very well, but substantially overestimates values in eastern North America (Schnell et al., 2015)

The detailed description and evaluation of ModelE2-TOMAS and the difference between OMA and TOMAS is available in Lee et al. (2014). In brief, the ModelE2-TOMAS and ModelE2-OMA models capture the observed sulfur species and other aerosol species as well as aerosol optical depth mostly within a factor of two. However, anthropogenic aerosols in both models differ from each other by a few percent to a factor of 2 regionally due to differences in aerosol processes such as deposition, cloud processing, and emission parameterizations."

Shindell, D. T., G. Faluvegi, N. Unger, E. Aguilar, G. A. Schmidt, D. M. Koch, S. E. Bauer, and R. L. Miller, Simulations of preindustrial, present-day, and 2100 conditions in the NASA GISS composition and climate model G-PUCCINI, ***Atm. Chem. Phys.****, 6, 4427-4459, 2006.*

Shindell, D. T., Pechony, O., Voulgarakis, A., Faluvegi, G., Nazarenko, L., Lamarque, J.-F., Bowman, K., Milly, G., Kovari, B., Ruedy, R. and Schmidt, G. A., Interactive ozone and methane chemistry in GISS-E2 historical and future climate simulations, ***Atmos. Chem. Phys.****, 13, 2653–2689, 2013.*

Schnell, J., M. J. Prather, B. Josse, V. Naik, L. W. Horowitz, P. Cameron-Smith, D. Bergmann, G. Zeng, D. A. Plummer, K. Sudo, T. Nagashima, D. T. Shindell, G. Faluvegi, and S. A. Strode, Use of North American and European air quality networks to evaluate global chemistry–climate modeling of surface ozone, Atmos. Chem. Phys., 15, 10581–10596, 2015

II. Carbon dioxide concentrations from future emission scenarios are derived by using a simple carbon cycle approach as in the IPCC reported in 2007. The manuscript shows no estimate of the uncertainty in future carbon dioxide concentrations estimated with that simple carbon cycle method approach. It would be interesting to learn how variable this estimates are and by how much could they affect the associated carbon dioxide forcing by 2050. I don't mean to run more simulations, but to discuss it in openly the manuscript, since CO2 is the main climate forcer after all.

Response) Thank you for the good suggestion. We have included the followings in Section 3.

"The impulse response function characterizes the complex behavior of the climate response to CO_2 emission changes as a first-order approximation. Due to the linear system assumption in the function, it has a limitation on representing non-linear and path dependent processes (e.g., Joos et al., 2013; Olivie and Peters, 2013). However, since CO_2 emission changes in our scenarios are much smaller than 1 Gt C per year whereas an impulse response function is likely in a linear regime when the CO_2 impulse size is below 100 Gt C (Olivie and Peters, 2013; Joos et al., 2013). Nevertheless, in order to estimate the variation in CO_2 RF associated with the choice of an impulse response function, we have estimated CO_2 RF using additional impulse response functions derived from multi-model intercomparison projects such as C^{4}MIP and CMIP5, which are obtained from Olivie and Peters (2013). We found that our CO2 RF differs only by 3-4% when using the impulse

response functions fitted to the multi-model mean of CMIP5 and by 10-17% when using the impulse response functions fitted to the multi-model mean of C⁴MIP."

Olivié, D. J. L., and G. P. Peters. "Variation in Emission Metrics due to Variation in CO2 and Temperature Impulse Response Functions." Earth System Dynamics 4, no. 2 (August 8, 2013): 267–86. doi:10.5194/esd-4-267-2013.

Joos, F., R. Roth, J. S. Fuglestvedt, G. P. Peters, I. G. Enting, W. von Bloh, V. Brovkin, et al. "Carbon Dioxide and Climate Impulse Response Functions for the Computation of Greenhouse Gas Metrics: A Multi-Model Analysis." Atmos. Chem. Phys. 13, no. 5 (March 8, 2013): 2793–2825. doi:10.5194/acp-13-2793-2013.

III: I think the article should clearly acknowledge that the study is based in numerical model experiments and openly discuss the disadvantage of running the same AGCM in all experiments. This relates to point I. How well does GISS ModelE2 represent aerosol properties and the direct and indirect aerosol effects? Due to the policy relevance of the article the authors could motivate similar studies with different models.
Response) We believe the original manuscript is clearly stated this is a modeling study throughout the manuscript, esp. in the Abstract and Introduction, but we have made a small change in the Abstract to be clearer.
"Using NASA GISS ModelE2," ➔ "Using the NASA GISS ModelE2 general circulation model,"

We agree with the review that using multiple AGCMs would provide more robust results than a single GCM-based study. New discussions (shown in bold) are added in the conclusion section to address the limitation of using single GCM in the study.
"Utilizing two independent aerosol models in the same host GCM, we have found that overall conclusions agree well between the two aerosol models, but missing species such as nitrate can influence the air quality and climate impact moderately. Our climate estimates reinforce that aerosol RF is a dominant forcing agent for regional climate change, and AIF is as important as ADF. A climate impact only based on aerosol direct forcing can be misleading, and we strongly suggest including AIF for more complete assessment of the climate impact of emission scenarios. **Since our study utilized a single host GCM, and we recognize that there are large model-to-model differences among GCMs (e.g., Shindell et al., 2013), we encourage other modeling groups to perform similar work using other host GCMs, to obtain more robust results.** "

2. Specific comments
Line 15 page 31386: Add "to" in "leading a strong positive forcing ..."
Response) Done.

Line 20 page 31386 & Line 1 Page 31404: I don't agree that having a regional positive forcing of 0.22 W/m2 by 2050 over the US is necessarily a climate dis-benefit, while the global radiative forcing is still negative. The number is relatively small and no climate effects (beyond the forcing) have been estimated in the article or even cited from other studies. Why is it a dis-benefit? Please clarify.
Response) First of all, we have toned down the statement.

" it leads to climate dis-benefit~" ➔ "it could lead to potential climate dis-benefit~".

We have included new discussion (shown in bold) in the Section 4.3.
"There is overall negative RF globally (-0.015 W m^{-2} in 2030 and -0.056 W m^{-2} in 2055) but positive over the US regions (0.14 W m^{-2} in 2030 and 0.22 W m^{-2} in 2055) because of positive aerosol RF. The localized aerosol RFs is due to its short lifetime, while the well-distributed negative CO_2 RF over the globe is due to its long lifetime. **The strong positive RF from aerosols are mostly localized over the U.S. especially over the eastern US (in Figure 10 for the 2030 case). Previous studies show a large influence of regional RF on the regional climate response (i.e., surface air temperature) over the US (Leibensperger et al., 2012) or the NH mid-latitude regions (Shindell et al., 2009). Our regional RF over the US is only 0.22 W m^{-2} in 2055 and therefore the resulting climate response would be small. Nevertheless it is likely to contribute to warming rather than cooling at least in the near-term and thus the CO_2 reduction policy used in our study could potential lead to mild regional climate dis-benefits over the US, especially during the summer (Shindell et al., 2016).**"

Leibensperger, E M, L J Mickley, D J Jacob, W T Chen, J H Seinfeld, A Nenes, P J Adams, D G Streets, N Kumar, and D Rind. "Climatic Effects of 1950-2050 Changes in US Anthropogenic Aerosols - Part 2: Climate Response." Atmospheric Chemistry and Physics 12, no. 7 (2012): 3349–62. doi:10.5194/acp-12-3349-2012.

Shindell, Drew, and Greg Faluvegi. "Climate Response to Regional Radiative Forcing during the Twentieth Century." Nature Geoscience 2, no. 4 (March 2009): 294–300. doi:10.1038/ngeo473.

Shindell, D., Y. Lee and G. Faluvegi, Climate and Health Impacts of US Emissions Reductions Consistent with 2 °C, Nature Climate Change, doi:10.1038/nclimate2935, 2016.

Line 23 page 31388: Remove on of the "the"s.
Response) Done.

Line 15 page 31391: Describe a bit more what is not included in micropyshics in OMA model, I understand that deposition, both wet and dry is taken into account, but what about nucleation, coagulation and condensation/evaporation? I understand that they are not represented in GISS-OMA, having no effects on the aerosol size distribution. Is this correct?
Response) Yes. ModelE2-OMA does not compute aerosol nucleation, coagulation, and condensation and thus aerosol size distributions. We have modified the following sentences to address the reviewer's question.

OLD: "ModelE2-OMA uses a default aerosol module, which has no microphysics."
NEW: "ModelE2-OMA uses a default aerosol module, which has no aerosol microphysics such as coagulation, condensation and nucleation and thus does not calculate aerosol size distribution."

Lines 15 to 20 page 31395: Why is valid using annual PM2.5 and hourly O3 concentrations to estimate mortality rates. Why not using hourly or daily PM2.5? Aren't daily exceedences of

PM2.5 also very important for mortality estimates?
Response) The choice of air quality metrics is dictated by published epidemiological studies linking those metrics to mortality changes and hence not chosen by us. The health analyses in this study draw directly upon studies relating annual PM2.5 and hourly O_3 exposures to health outcomes. Therefore these are the most appropriate quantities in this case.

Lines 8 and 9 page 31400: Are You sure that the main difference between models is due to the missing nitrate aerosol. What about aerosol microphysics in TOMAS vs. no aerosol microphysics in OMA. How important is it to have a microphysical treatment of aerosol particles in a AGCM? Since You have done similar simulations with both aerosol schemes, it would be very useful to have more discussion on the differences and maybe even a table or figure.
Response) For the PM2.5 evaluations done in Lee et al. (2014), both models performed quite similarly in general, but showed some difference when non-negligible amount of nitrate aerosols is present in ModelE2-OMA. PM2.5 is less or little impacted by aerosol microphysics, as PM2.5 over the US is mainly contributed by anthropogenic aerosols, which is mostly present in fine mode. ModelE2-OMA has a good skill to simulate aerosol mass concentrations. Aerosol number and size information is the major improvement by having aerosol microphysics.

Figure 4: Why is the difference in PM2.5 not shown?
Response) We excluded the PM2.5 in Figure 4, as the difference in PM2.5 is roughly the same as nitrate shown in Figure 2. As we explained in the manuscript, ModelE2-TOMAS predicts more sulfate reduction and less OM reduction, compared to ModelE2-OMA. This cancels each other and it makes the overall PM difference driven by nitrate, which only ModelE2-OMA simulates. We have modified the sentence below to clarify this point.

OLD: "These effects cancel each other and overall PM is little influenced by the choice of model."
NEW: "These effects cancel each other and overall PM2.5 difference between the models is almost equivalent to the amount of nitrate shown in Fig. 2."

Figure 5 caption: What do the error bars mean?
Response) We used the error bars to represent the higher mortality rate using CRF ($CRF_{high, PM}$) and the lower mortality rate using low CRF ($CRF_{low, PM}$). We have mentioned in the Figure 5 caption. However, in case it wasn't clear, we have modified the caption to clarify.
OLD: "The upper and lower bars are for mortality rates using $CRF_{high, PM}$ and $CRF_{low,PM}$, respectively."
NEW: "The higher ($CRF_{high, PM}$) and lower ($CRF_{low,PM}$) bars indicate the spread in mortality change predicted using the range of concentration-response functions used in this study (see Table 2)."

Figure 6 caption: Do You really mean "upper error bars" or is it "lower error bars" ?
Response) See our response above. For Figure 6 with ozone, we used the horizontal upper bar to present the higher mortality rate using $CRF_{base,O3}$. Since the upper bar is used to

present higher value, it should be named “upper bar”. We’ve changed “the upper error bars “ to “the horizontal upper bars”.

Figure 12 caption: Is it really “Same as Fig. 8” or is it “ Same as Fig. 9” ?
Response) Thanks for catching the error. It should be “Same as Fig. 9”.