Contribution from the ten major emission sectors in Europe and Denmark to the Health-Cost Externalities of Air Pollution using the EVA Model System – an integrated modelling approach, by Brandt et al.

Response to Anonymous Referee #1

General statement concerning the review:

We would like to thank the reviewer for a very thorough and extensive, however as we see it, also very critical review. The reviewer has questioned nearly everything in the paper; from the largest open questions within this interdisciplinary research field to the smallest detail in the paper. Making state-of-the-art science requires open discussions, and we welcome the discussions concerning all aspects of our work. We have therefore answered every scientific question and comment as they are indeed important. We consider our work as a step forward with respect to calculating health impacts and related external cost from air pollution using a regional chemistry transport model (including a tagging methods) for all individual scenarios, compared to the previously developed model systems, which are assuming linear atmospheric chemistry of pre-calculated source-receptor relationships. Of cause, we do not claim to have solved all problems in this very complex and highly interdisciplinary research field, which is impossible to do in one single paper. The work in this paper is the result of a 5 year interdisciplinary research center supported by the Danish Strategic Research Council, including leading research groups within the fields of air pollution modelling, health impacts (both epidemiologists and health statisticians) as well as environmental and health economics, working within this research area for more than two decades. The purpose of the paper is to define and demonstrate a new methodology for estimating the contributions from individual emission sectors to health related external costs with the main focus on developing the overall integrated model system, which is able to handle the non-linear behavior of atmospheric chemistry, which is new compared to previous integrated models systems. In recent years, these new and very interdisciplinary research fields have evolved, combining the research from e.g. natural sciences, health sciences and social sciences, and this work should be continued to pursue the many open questions which the reviewer has made. Not many scientific journals are presently aimed at publishing highly interdisciplinary research and therefore a journal must be chosen for which the readers find the results relevant and important. As this paper concerns air pollution modelling where an understanding of the atmospheric chemistry is essential for understanding the response in the atmospheres composition to emission reductions and since some of the new science in the paper concerns a new method for handling the highly non-linear atmospheric chemistry with respect to emission reduction scenarios and not the least, the response in the atmospheric chemistry to emission sector contributions, in an integrated impact assessment, the results are both relevant and important to the readers of ACP.

GENERAL COMMENTS

Reviewer 1: This paper is not within the regular scope of Atmospheric Chemistry and Physics, since it is not focused on atmospheric chemistry or physics but on health effects and economic valuation of different emission sectors. However, these issues are of interest to a large part of the ACP community. The paper presents results from calculations with the EVA model system, which uses results from DEHM model calculations of various air pollutants to assess health-related economic externalities of air pollution. The intention of this study is to estimate the health-related costs for different emission sectors in Europe and Denmark.

Answer 1: A large part of the work focuses on atmospheric chemistry and physics and the impacts from air pollution based on our present day understanding and knowledge. It puts the air pollution challenge into perspective and provides a new methodology for calculating source-receptor relationships taking into account non-linear atmospheric chemistry and proves that it is possible to carry out calculations for

decision support without using strongly simplified atmospheric models or assuming linear source receptor relationships. Therefore the paper is both relevant and important to the ACP community (see also our general statement above).

Reviewer 2: The DEHM model with three two-way nested model domains, with a horizontal resolution of 50km over Europe and 17km over a part of Northern Europe, could be useful to estimate ecosystem impacts (via nitrogen and sulphur deposition and ozone damage to plants) from different source sectors. It can also be used for estimating contributions to total PM2.5 (and constituents) to regional background concentrations in Europe. However, the background concentrations of PM2.5-constituents are typically not very useful, except as background/boundary fields for higher resolution models that can be used to model population exposure at more appropriate scales. A major problem in this study is that the grid resolution is too coarse for accurate estimates of population exposure, at least for all sources located near the population and with low emission heights (e.g., traffic and residential heating; population exposure from these sources are likely to be substantially underestimated in coarse grid models). For more diffuse sources (such as agriculture) the resolution is less problematic but it makes it very difficult to compare the impact of these sources to urban emissions. Some (much) higher resolution models (or methods to transform the rather coarse DEHM results to sub-grid scale concentrations with appropriate concentration gradients near roads and in residential wood burning areas) needs to be added to estimate population exposure from several important emission sources.

Answer 2: The DEHM model has been used for many years to estimate ecosystem impacts concerning eutrophication and acidification as well as impacts from ozone on crops. There is not a major difference from estimating these kinds of impacts to impacts on human health. Both the nitrogen species, the sulphur species, ozone and $PM_{2.5}$ are long-range transported and can be handled to calculate human exposure in an integrated assessment on relative course resolution, as has been conducted in similar systems for many year (e.g. the RAINS/GAINS system). It is clear that a regional CTM cannot handle the local effects from e.g. traffic or residential heating (and we do not claim that we are able to do so), and therefore the impacts calculated with such a system might be underestimated, since it cannot account for the near-source local effects. Nevertheless, it is only a small part of the exposure, which is missing, since the contribution from e.g. a highly trafficked highway decreases to background levels at distances more than 100 m away from the highway and the share of the population living within this distance from e.g. a highway is relatively small. However, as all the species in the model system are subject to long-range transport, and since the DEHM model performs well compared to measurements (see e.g. Brandt et al., 2012), the EVA system can be used to estimate the general exposure and health impacts in Europe (as has been done in Clean Air for Europe (CAFE)) on the regional scale. Furthermore, the concentration-response functions are only valid for the regional and urban background on an annual basis and not for short-term high exposure in e.g. street canyons in cities. There might be additional impacts on human health with respect to short-term exposures, however, we are presently not able to take this into account in impact assessments. Moreover, the so-called urban increment in the urban background in European cities is relatively small. A majority of the $PM_{2.5}$ concentrations in the urban background originates from long-range transport (80-90%) (Cuvelier et al., XXXX) and therefore the error of using a regional air pollution model for calculating human exposure to $PM_{2.5}$ in the urban background is minor. Nevertheless, we agree that a next step in developing the EVA model system for refining the results, is to take into account the urban increment using an urban background model, and this work is presently being conducted and will be documented in a future paper.

Reviewer 3: Brandt et al. seem to have missed a very similar article by S J Griffiths (Air Qual Atmos Health, 2011, 4: 189-197). Griffiths also looked at population exposure in Europe due to emissions in the different SNAP sectors (plus shipping) and discusses the importance of different sectors and the implications of different toxicity assumptions for designing effective emission reduction policies. In my opinion the present paper by Brandt et al., do not add much new useful information (at least not within atmospheric chemistry

or physics) compared to Griffiths (2011). Griffiths did not go into economic valuation but the usefulness of this part is questionable, considering the fact that knowledge about health impacts of different sources is not at a level where realistic comparisons of different sources is possible. At least Brandt et al., have to compare their results to Griffiths (2011) and, considering the results presented by Griffiths, revise the economic valuation part to properly handle the uncertainty range in impacts of different particle types.

Answer 3: We have now taken into account the results from Griffiths (2011) and compared with results in our work.

Reviewer 4: Brandt et al., present health impacts (and economic valuation results) as if they were very accurately known, and can be easily compared between different emission sectors. As Griffiths (2011) showed this is not the case. Emission sectors contributing primarily to secondary inorganic aerosol (e.g., agriculture) may have almost zero impact on health if ammonium and ammonium nitrate are harmless (or close to harmless compared to primary PM components, which is definitely possible). The model calculations by Griffiths showed that Agriculture contributed 38.2% to total European exposure of PM2.5 but only 2.4% to exposure of primary PM2.5. This means that the uncertainty range in health impacts from this sector is huge! This needs to be made clear also by Brandt et al. in order to not mislead readers. This is obviously not the only uncertainty component but possibly the largest one (at least for the agricultural sector).

Answer 4: We do not think that we treat health impacts (and economic valuation results) as if they were very accurately know. The key figure in the calculations is the relative risk for all cause deaths, where we have used a relative risk for $PM_{2.5}$ of 1.06 which is commonly used in these kinds of health impacts assessments. This factor has even recently been modified to 1.062 (WHO 2013), so we consider the figure to be even more precise now. Concerning the economic valuation, this is, of cause, associated with uncertainty, however the economic valuation is based on a general accepted method on willingness to pay. In any case, the economic valuation is not sensitive to the emission sector, as the valuation is related to the health impacts and not the emission sectors. The central question is the health impacts associated with the particle composition, and this is answered below (see answer 5). It is clear that the model calculations by Griffiths showing that Agricultural sector contributed 38.2% to total European exposure of total $PM_{2.5}$, since most of the emissions result in secondary formed aerosols and only 2.4% to exposure of primary $PM_{2.5}$, since the primary emissions are much smaller for this sector. We get similar results using the EVA model system. We also agree that the uncertainty of the estimated health impacts from this sector very much depend on the toxicity of the secondary formed particles compared to the primary emitted particles. However, as we argue in Answer 5 below, there is no evidence that the secondary formed particles are not harmful, and since there is only indications that the primary particles are more harmful than the secondary on the short-term, but not on the long-term we have to treat the particles equally. We have in the paper clearly stated that this is the assumption and all calculations are based on this assumption, therefore the results are valid under this assumption (which is also assumed e.g. in RAINS/GAINS presently used to support the European Commission). It is not a problem to assign different toxicity to different particles in the EVA system, but we have not found evidence to do so. On the other hand, if we had assigned nearly zero toxicity to the secondary formed inorganic particles (SIA), other reviewers would have criticized our work to be bias towards e.g. supporting the agricultural sector. With the present knowledge, it is not possible to distinguish the toxicity from the different particles. Presently it is a matter of believe whether to assign greater toxicity to primary particles or e.g. to EC/BC. There are researcher within the community who believe that should be assigned with lower toxicity and other researchers who believe that this should not be the case. The discussions just show that we need much more research in this field. Since the state-of-the-art cohort studies are not able to clearly distinguish between the particles, we presently need to handle the effects from particles equally (see also answer 5). At the same time, we can carry out sensitivity

studies showing the effect of assigning different toxicity to different particles, which we also have conducted.

Reference:
WHO (2013) Summary of recommendations for Question D4 on 'Identification of Concentration-Response Functions for Cost-effectiveness Analysis. World Healh Organiziation, Regional Office for Europe, Bonn, Germany

Reviewer 5: There is no evidence indicating that all different PM components have identical (as assumed in the EVA model) or even similar health impacts. It is true that presently no PM components have been proven to be completely safe (which is extremely difficult to prove given the highly complex emission mixtures and complex atmospheric PM composition with many correlated species) any component as potentially causing some health impacts but it is very likely that some components are many orders of magnitude more dangerous than others (on a mass based scale). It seems likely that some components are mostly harmless (apart from indirect effects on lung deposition efficiency, due to changes in particle size, which could be both positive and negative from a health perspective, but this complex issue is not treated in this paper, and it is very unlikely that the impact is well represented by the simple linear increase of health effects due to particle mass increase).

Answer 5: It is not true that we have not treated the highly complex issue of health impacts from particles in the paper. As we acknowledge that the issue of health impacts from various particle compositions is very important, we have handled this shortly in section 2.6. However, it is out of scope to make a complete review of the health impacts related to atmospheric particles in this paper, since it is certainly not the main focus of the paper and we have only included references to the most important and substantial papers within the field. Furthermore, we have shortened this section considerably according to a suggestion from the reviewer in the initial review process of ACP. The EVA model system includes the possibility to assign different toxicity to the different kind of particles, however, after a very extensive review of this issue within the Danish research groups working with health impacts from particle, we conclude that presently, there is no evidence to conclude that any atmospheric particle is harmless and should be excluded. Futhermore, there is not much evidence that e.g. BC/EC are much more harmful compared to the total $PM_{2.5}$ mass. This is a working hypothesis in the research field, which is supported by studies of short-term health effects, but as long as clear evidence does not exists that some particle species are more harmful that others, we have to treat the particles equally.

According to the latest extensive metaanalysis (Hoek et al., 2013), which have assessed all cardiovascular and respiratory causes to death with respect to $PM_{2.5}$, PM_{10}, EC/BC and NO_2, they conclude that it is presently not possible to distinguish the effects from PM_{10}, EC/BC from $PM_{2.5}$ as they are strongly correlated. It is basically a matter of which species is used as a proxy for the health effects and the majority of studies concerns health effects using $PM_{2.5}$ as a proxy and therefore most evidence exists for the total $PM_{2.5}$. E.g. the effect from EC/BC has been found to be app. 10 times higher than the effect from $PM_{2.5}$ in the same studies, however, the concentrations are also found to be 10 times less and as we are talking about the same number of health impacts, one can chose to use BC/EC as a proxy or the total $PM_{2.5}$, however, evidence does not exists in these studies that BC/EC is more harmful on the long-term than the other chemical species included in the total $PM_{2.5}$ even though it is presently treated as a working hypothesis in the research field supported by the short-term evidence.

Elemental carbon (EC) or as black carbon (BC) as usually is about 10% of the fine particle mass and has shown to be associated with health effects so that an effect of 1 μg/m3 approximately corresponds to the effect of 10 ug/m3 $PM_{2.5}$, and that often is an independent effect in short term studies when adjustment for $PM_{2.5}$ is made (Hoek et al 2013; Janssen et al 2011, 2012). This means that EC/BC often can explain a

significant part of the short term effects from the total fine particle mass and that this provides independent information in addition to $PM_{2.5}$. However, this is not yet documented in cohort studies of long term effects, where it not to sufficient degree has possible to extract the effect of EC/BC in relation to the total $PM_{2.5}$. It is therefore really important to distinguish what we know about health impacts from short-term and long-term effects from the different kind of particles.

We agree totally that assigning different health impacts to different kind of particle composition is maybe the most important research issue within the field presently. However, we still have to conclude that as the evidence does not exists presently to distinguish between the particles. Since evidence exists that no particles can be treated as harmless, we have to treat the effects from the different particles equally. As this is already the common procedure in the present major decision support system (RAINS/GAINS) and as long as we clearly state that this is an assumption we have to make in our calculations, we do not consider this to be a major issue in the paper. The only thing, which presently can be done is to show the sensitivity of the calculations to assigning different toxicity to the different kind of particles, and this has been carried out in the paper.

References:
Hoek G, Krishnan RM, Beelen R, Peters A, Ostro B, Brunekreef B, et al. 2013. Long-term air pollution exposure and cardio-respiratory mortality: a review. Environ Health 12:43.

Janssen NAH, Hoek G, Simic-Lawson M, Fischer P, van Bree L, Brink Ten H, et al. 2011. Black Carbon as an Additional Indicator of the Adverse Health Effects of Airborne Particles Compared with PM10 and PM2.5. Environ Health Perspect 119:1691–1699

Janssen NAH, Gerlofs-Nijland ME, Lanki T, Salonen RO, Cassee F, Hoek G, Fischer P, Brunekreef B, Krzyzanowsk M. 2012. Health Effects of black Carbon. WHO Europe.

Reviewer 6: It is dangerous to present economic costs for individual emission sectors in the way done here without proper estimates of the uncertainty intervals, since it could lead to completely wrong conclusions regarding effects of different emission sectors (and consequently suboptimal, or even counter-productive, measures to reduce air pollution impacts). At the moment we can only make very rough guesses about the total health impacts of air pollution on the European scale. There are too many (large) uncertainties about emissions, model description, human exposure, health impacts of different substances for it to be possible to compare the "health costs" of the different SNAP sectors in a scientifically sound way.

Answer 6: We do not agree that it is dangerous to present economic costs for individual sectors. This has been conducted already in Clean Air for Europe (CAFÉ), and it provides the best foundation for decision support. The economic cost is of cause uncertain, depending on the criteria chosen for making the economic valuation. This is also the reason for making the results as relative contributions from the different sectors in percentage related to their contribution to the total external costs. However, even though economic valuation can be associated with uncertainties, we obtain similar results for the total health costs as obtained in CAFE using the RAINS/GAINS system (app. 800 billion Euros/year). This number is generally accepted as the health costs in Europe related to air pollution. Whether it is half or twice this amount is not really important, since the economic valuation is only a measure of the importance of air pollution problems compared to other challenges of our society. The general uncertainties of long range chemical transport models are presently +/- 30% on the annual basis within a model grid cell and mainly associated with the uncertainties in emissions. It is also clear that if one examines the uncertainty on the human exposure on the individual level, the uncertainty is large. However, the uncertainty in the population level is much smaller. We simply disagrees with the statement that uncertainties are too large

to be able to compare the health costs of the different SNAP sectors. On the contrary, we claim that this methodology id the only scientifically sound way to carry out integrated impacts assessments on the European and national levels. This has been done for many years using the RAINS/GAINS system, however, this system is not taking into account the highly non-linear atmospheric chemistry due to cost efficient reasons making source-receptor relationships for many sector and countries in Europe. Our goal with this work is partly to take a step further and show that this is feasible with present day technology using state-of-the-art methods in every part of the impact pathway chain.

Reviewer 7: Although the details are not yet known about which air pollutants are directly responsible for various health impacts it is definitely clear that different sources of air pollution have substantially different effects. The economic valuation of the health impacts of different sectors becomes meaningless when using unrealistic assumptions about the exposure response functions (ERFs). Economic valuation for individual emission sectors should only be attempted for effects that are at least reasonably well understood from a more fundamental perspective (e.g., crop losses due to ozone uptake). Ammonia emissions from agriculture (and other sectors) certainly need to be limited as much as possible due to the immense ecosystem effects but that is not an issue covered in the EVA model. Sulphur and NOx emissions also needs to be limited (e.g., for acidification, eutrophication and ozone production reasons). When it comes to health impacts, combustion generated PM, and possibly mechanically generated wear particles of different sizes (road dust, break wear PM, etc), ozone, various organic compounds, metals and other toxins are likely the most important to consider. Improving air quality is an important issue and the costs of air pollution are no doubt very large both from a monetary perspective and in terms of human suffering and ecological consequences. However, the EVA model (with the assumptions used in this study) does not produce useful results for guiding policy and it does not add any new insights into atmospheric chemistry or physics. For these reasons I have to recommend that the paper is not accepted for publication in ACP.

Answer 7: We agree that the details are not yet known about which air pollutants are directly responsible for the various health impacts. On the contrary, the main cohort studies are not able to distinguish between the different kinds of particles. There is a clear and significant correlation between the total $PM_{2.5}$ concentrations with a relative risk of 1.06, confirmed in many studies and adopted by both the WHO and EU. Therefore, this concentration response function related to this relative risk can be used for health impact assessment studies on the long term and on the regional and urban background for which the relative risk is valid for. The economic valuation of the impacts from air pollution is also an established research field, where results e.g. have been used in Clean Air for Europe (CAFE). The effects from crop losses due to ozone exposure are not better understood than the air pollution related health impacts – on the contrary – much more evidence exists that health impacts associated with PM2.5 is a fact and that the relative risk is relatively uncertain and has especially been verified in recent adaptation by the WHO to be even more precise (1.062).

The reviewer expresses a very personal view on which effects related to which chemical compounds should be regulated (nitrogen and ozone with respect to ecosystems and crops, respectively) and which emissions needs to be limited. The reviewer suggest a list of species “combustion generated PM, and possibly mechanically generated wear particles of different sizes (road dust, break wear PM, etc), ozone, various organic compounds, metals and other toxins” that are likely the most important to consider, however, these are based on working hypotheses and not on the present confirmed knowledge. We totally agree that health impacts related to these species are very important to study in the future, but for present day health impacts assessments, one has to base the results on confirmed knowledge, and therefore we have used the generally accepted relative risk for the total $PM_{2.5}$ which has been confirmed and adopted by the WHO, which is the only thing which can be done at the present stage of knowledge.

SPECIFIC COMMENTS

Reviewer 8: Model uncertainties need to be taken into account in the economic valuation. You need to specify how uncertain the model results are. How well do you model the different components (especially in densely populated areas)?

Answer 8: The DEHM model has been thoroughly tested against measurements in many publication, see e.g. Brandt et al. (2012). The model performs well on the regional level for the whole of Europe, which it has been used for. It is not possible to assess the model uncertainties concerning the emission reduction scenarios, since no measurement data exists to validate scenarios. The only thing modelers can do is to develop models based on best knowledge on atmospheric processes concerning chemistry and physics and then compare model results with available measurements to ensure that chemical and physical processes are described in best possible way. When we believe that process descriptions in the models are adequate for reproducing measurements, we have to believe that the models can also describe realistic and precise responses to e.g. emission reductions. Any uncertainty that we assign to the results would be wrong, since the whole system is very complex. However, one can do the experiment of assigning uncertainties to every link in the impact pathway chain, with 30% uncertainty to each overall link of emissions. atmospheric modelling, eksposure, concentration-response function and economic valuation, then the total uncertainty of the whole system would be SQRT($(1/3)^2$ +$(1/3)^2$ +$(1/3)^2$ + $(1/3)^2$ + $(1/3)^2$) = 31%, which an acceptable uncertainty for such a system. However, the biggest uncertainty in the system is probably the lack of secondary organic aerosols (SOA), which presently is being implemented in the DEHM model, but we already clearly stated that in the paper and the impacts assessment for SOA is future work.

Reviewer 9: The uncertainty in model results (per sector) should be included in the impact assessment. It is quite likely that some components are better modelled than others and that some are consistently over- or underestimated. Since the work tries to compare the relative importance of different emission sectors this information is very important. At least information regarding model bias for “health-relevant” parameters should be given (SOMO35, SO2, CO, Ammonium, Sulphate, Nitrate, EC/BC and OC concentrations). Tables with this information are most suited as Supplementary material but it is important to know if, e.g., model calculated agricultural related components (ammonium and nitrate) have a very different bias compared to primary components (EC/BC and OC). Coarse (insoluble) particles (e.g., from road and car brake wear and desert dust) are also potentially harmful. Some motivation for excluding the coarse fraction from the EVA model is needed.

Answer 9: As stated in the paper, the DEHM model has been thoroughly tested against nearly all components in previous papers, see e.g. Brandt et al. (2012) with satisfactory results. There are no clear bias’s for individual species, which all performs well, except a clear underestimation for the total $PM_{2.5}$. due to lack of mass closure for the total PM. This is a normal condition for present long range chemical transport model, where some processes are missing. We consider one of the main missing species missing to be SOA and this is why we are presently implementing this group of species in the model. However, it is clearly stated in the paper that SOA is missing, so this contribution to the total $PM_{2.5}$ is not assessed. However, the major part of the SOA formation is due to natural sources, and therefore it will not change the absolute or relative contribution from the anthropogenic emission sectors much in the assessment part of the paper.

Reviewer 10: Is there really any reason to believe ammonium nitrate is more dangerous to health than sea salt? Is there any reason to believe that sulphuric acid particles get more dangerous for human health by

being (partially) neutralised by ammonia? If sulphate particles have any direct health impacts (which is doubtful at ambient air concentration levels) I would assume that they are coupled to the particle acidity, which decrease when ammonia is taken up by the sulphuric acid particles. Ammonia emissions need to be minimized for various severe ecosystem effects, but not because of direct health impacts as air pollutants (there is no evidence that ammonium nitrate or ammonium sulphate are impacting human health and it is extremely unlikely that they would be as dangerous per mass concentration unit as directly emitted combustion particles).

Answer 10: As stated before, we are using the WHO accepted standards associating health impacts with the total $PM_{2.5}$ mass as this association has been most significantly verified. See answer 5 for more details.

Reviewer 11: Even ExternE 1999 (which is the major reference of the present work) indicates that nitrate may be harmless and that this could explain a higher toxicity of PM in the USA compared to Europe. It is not proven that nitrate is totally harmless but compared to primary particles from combustion (soot etc.) it seems likely to be very much less toxic. Thus it does not make sense to apply equal toxicity to all different PM components. It seems clear that the ExternE methodology was not meant to be used in the way done in the EVA model.

Answer 11: The main difference between ExternE and the Eva system is the application of state-of-the-art air pollution models, where ExternE used extremely simplified models in the beginning. ExternE is now being updated with state-of-the-art models, which are comparable to the models used in the EVA system. An indication in a reference from 1999 concerning toxicity of the different species cannot be used today, where much more knowledge is available. The present day cohort studies in the US and Europe does not indicate a major difference in toxicity – on the contrary – the results confirms each other. As we stated before, we agree that primary particles might be more harmful than secondary formed particles, but with the present knowledge we can still not distinguish the toxicity between the different particles, and therefore this difference remains as scientific hypotheses to be tested further. As we acknowledge, that primary particles can be more harmful to human health, we have performed a sensitivity test to examine the impacts of assigning different toxicity to the different particles. This is the best which can be done with the present knowledge.

Reviewer 12: In general, I find it rather unsatisfactory that Brandt et al. refer so much to ExternE and other "technical reports" from various sources in a scientific paper. In my opinion peer-reviewed scientific papers should primarily be cited.

Answer 12: We agree that peer-reviewed scientific papers should primarily be cited. Most of the references in the paper are to peer-reviewed papers. However, Since the EVA model system is based on the impact pathway method used in ExternE and in RAINS/GAINS, which are mainly documented in technical reports, we have to include these references too.

Reviewer 13: The abstract is not very easy to read and feels a bit repetitive (for an abstract). I would prefer it somewhat more compact. Some suggestions for shortening are given here: Page 5873, line 10: Change "the most to human health impacts using this tagging method." to "the most to human health impacts." line 12: "from the ten major emission sources" -> "from ten major emission sources" line 16: remove "of the ten major emission sectors in Europe and Denmark" line 22-25: change the sentence: "The results in this work emphasize the importance of defining the right questions in the decision making process, since most of the atmospheric chemical compounds are linked via non-linear chemical reactions, which are important to take into account." to "The results in this work emphasize the importance of defining the right questions in the decision making process." [the rest has already been pointed out]

Answer 13: All the suggestions from the reviewer have been changed.

Reviewer 14: Also in the rest of the paper some things could be written somewhat less repetitively. I will give some suggestions for changes in the following but mostly leave this aspect of improvement for the authors. The paper is very long so cutting it down by avoiding unnecessary repetitions (and possibly moving some details to a Supplement) could improve the readability of the paper.

Page 5874, line 5: I think one general reference about "health-related external costs" would be good to include since economy is outside the regular scope of ACP some readers will likely be unfamiliar with this.

Answer 14: We have inserted the following to explain external costs "(i.e. the indirect cost associated with the activities resulting in emissions)".

Reviewer 15: line 16: remove "than the most obvious and visible sources"

Answer 15: Done.

Reviewer 16: line 17-19: Remove the sentence: "When quantifying emissions, more than ten major emission sectors are defined of which the major power plants and road traffic constitutes two." It is strangely formulated (there are hundreds of different emission sectors) and contains no useful information.

Answer 16: Reviewer #2 had another comment to this sentence, which we have followed.

Reviewer 17: line 27: "from the ten major emission sectors" I suggest you change this to "from ten major emission sectors (SNAP categories)"

Answer 17: Changed accordingly.

Reviewer 18: Page 5877, line 13-15: The purpose of the RAINS/GAINS system is not to extrapolate the results to 100% reductions in individual emission sectors. The typical aim of the RAINS/GAINS model runs is to study impacts of relatively modest emission changes to estimate impacts per ton emission change to optimise emission measures to reach certain air quality targets as cost efficiently as possible. This is very different from trying to estimate the total impact of a single emission sector. The total impact of a single sector is usually not very well defined since the emissions may interact strongly with emissions from other sectors. E.g., the agricultural sector emits mostly NH3, which only forms particles if there is sulphate or HNO3 available to react with. In a simplified case: assume a region where Agri only emits NH3 and, which is influenced by the road traffic sector emissions of NOx (neglecting the small sulphur emissions from this sector) and not influenced by significant sulphur sources. For this case the methodology used by the authors will count the formed NH4NO3 (by the reaction NH3(g) + HNO3(g) <-> NH4NO3(p)) BOTH as a contribution from the agricultural sector (because of the NH3-emissions) AND from the road traffic sector (because of the HNO3 formed from the traffic NOx emissions). Similarly the NH4 part of ammonium sulphate will be allocated to two different emission sectors. This means that the summing of the emission sectors (1-10) will necessarily overestimate the total impact of all emissions. And this means that the "health related costs" associated with the agricultural sector (and the NOx and SOx-sectors) will be unfairly high (the cost for NH4NO3, if you really can prove that it is dangerous, should be split between the different sources contributing to it). You could argue that the removal of all agricultural NH3 emissions (while keeping all other emissions at their present level) would lead to the "saving" of the same amount as presented in the paper but this is not the way it is formulated (and it is a very strange way of counting costs that should be split between the contributing sectors). Also, extrapolation of the EVA/DEHM model results from the present study to a 10 or 20% decrease in emissions would lead to similar (or perhaps even larger)

"linearity problems" as the conventional RAINS/GAINS model approach. The CTM used in RAINS/GAINS treats the chemistry at a similar level of detail (and non-linearity) as the DEHM model.

Answer 18: This is more a statement than a comment to the paper. It is, of cause, true that the CTM (the EMEP model) used in RAINS/GAINS treats the chemistry at a similar level of detail (and non-linearity) as the DEHM model. The main difference is in the way scenarios are calculated; in EVA using the CTM every time, and in RAINS/GAINS extrapolating from linearized source-receptor relationships. Of cause, extrapolation of the EVA/DEHM model results from the present study to a 10 or 20% decrease in emissions would lead to similar linearity problems as the conventional RAINS/GAINS model approach, however, this is actually the main point in the paper – that the system is not used to make extrapolations, but new source-receptor relationships are calculated every time and for every scenario. We are also not stating in the paper, that the purpose of the RAINS/GAINS system is to extrapolate the results to 100% reductions in individual emission sectors. However, we indicate that serious errors are being made even when source-receptor relationships for one country, for one species, for one reduction of e.g. 15% for one year cannot be used to assess the impacts on air pollution levels from multi-country, multi-species, multi-year reductions or even for different reduction rates for different countries/species/years etc.

The reduction scenarios in our paper are included to demonstrate the EVA systems capability, not to be extrapolated to other scenarios. As we write in the beginning, the results in this work emphasize the importance of defining the right questions in the decision making process. The scenarios should be designed for every single question. It is totally clear, as the reviewer states, that "the total impact of a single sector is usually not very well defined since the emissions may interact strongly with emissions from other sectors". This is exactly the reason for the scenarios we have chosen to demonstrate the system. The main question we ask is: What happens to the atmospheric chemistry and subsequently the air pollution levels when the emissions from one sector is removed and this we define as the contribution of this sector to the total air pollution levels. Of cause, we agree that the formation of e.g. ammonium-sulphate inorganic particles in the atmosphere surely depends of emission of both ammonia and sulphur from different emission sectors and if any of these are missing the particles are not formed. The contribution from any emission sector depends clearly on the chemical composition of the atmosphere from all sources and therefore, the contributions are changing over time and space in a highly non-linear dynamic system. This is exactly the reason for the necessity of running emission reduction scenarios every time a new question should be answered and not running predefined source-receptor relationships, which are extrapolated.

At last, we would like to emphasize that the discussion given by the reviewer has already been addressed in the paper as follows in section 5:

"The results in this work emphasize the importance of defining the right questions in the decision making process, since most of the atmospheric chemical compounds are linked via non-linear chemical reactions. The results represent the case where all the emissions (of all the different chemical compounds) from the sector of interest are reduced simultaneously. In the case where one would like to examine the external cost for the impacts of sulphur alone (e.g., by reducing the sulphur content in the fuel), the model system should be run with the individual emitted compounds one at the time, for example, the sulphur emissions from traffic alone. This would result in different external costs, especially for the countries surrounded by large sulphur emissions from other sources and where the NO_x emissions are relatively high compared to the SO_2 emissions from the traffic sector, as is the case for the Scandinavian countries.

The atmospheric system is highly non-linear, complex and interdependent. Therefore the results from assessing the impacts from each emission sector depend clearly on the assumption that the other emission sectors are not changed. As an example, consider whether the health impacts of emitting ammonia from the agricultural sector would be much lower if all the other emissions from burning fossil fuels were

removed. If there is no sulphur- or nitrogen-oxide available, the ammonia cannot be transformed into ammonium particles. Another example in this work is the relatively high cost of sulphur emissions from e.g. the road traffic in Denmark. This is related to the life time of SO_2 from the international ship traffic in the Baltic Sea and the North Sea, which is influenced by the emissions of NO_x from the same sector (e.g. road traffic). The NO_x emissions increase the concentration of OH in the area, which enhances the transformation rate of the emitted SO_2 from ships to sulphate. In this way, the NO_x emissions increase the sulphate concentrations, which are accounted for in the "Total S" external costs. The emissions are all linked in the chemical composition of the atmosphere via non-linear chemical processes".

Reviewer 19: Section 2.2. Important model information is missing: A critical issue when using CTM results for estimating population exposure is the vertical resolution of the model near ground. How thick is your lowest model level? And how do you estimate concentrations at "human height" (ca 1.5m, unless the lowest model level is extremely thin some adjustment may be necessary)?

Answer 19: A critical issue when using CTM results for estimating population exposure is not necessary the resolution of the lowest model layer, but whether the model is able to reproduce the air pollution levels compared to the monitoring networks, usually located at 2 m heights. We do not anticipate that there will be any major difference in 1.5 m height or 2 meter. The model has been thoroughly tested against measurements in Europe in numerous publications - see e.g. Brandt et al. (2012). The DEHM model in the present setup has a height of the lowest model layer of 15-20 m, and even included functions for describing the levels at 2 m height, however, these functions made no significant difference, and we have chosen not to use them.

Reviewer 20: Related to this is the issue of emission heights. How do you distribute emissions in the vertical?

Answer 20: The distribution of emission in the vertical is described in Brandt et al. (2012). In a paper describing an integrated modelling approach, we cannot give a detailed description of the CTM used in the study. The model is described extensively in the references.

Reviewer 21: The model horizontal resolution is very coarse for estimating population exposure. 50km resolution cannot capture the very strong concentration gradients near large emission sources located close to cities (which are not resolved well even at the 16km resolution). This means that comparing "area-type" sources such as agriculture to more local sources (e.g., traffic and residential heating) is not trivial. As far as I understand this problem is not handled in the EVA model. You need to discuss this issue and estimate the uncertainties in population exposure due to (too) coarse model resolution for (at least some of the) important sources.

Answer 21: We simply disagrees that the resolution in the model is too coarse for estimating population exposure at the regional and urban background level for these long-range transported species. The resolution of 50 km x 50 km has been used for many years in the RAINS/GAINS system to calculate similar exposures. Furthermore, we do not claim that we are able to calculate the local exposure from traffic (which is only important up to a 100 meters from the source) and residential heating. Of cause, no regional model can calculate local exposures and no integrated model systems are presently taking into account the local exposure from local sources like these. However, we are presently developing a new version of the system using an urban background model to take into account the exposure at much higher resolution. The results from this work will be published in a future paper.

Reviewer 22: For this particular study the emissions used need to be described much better than what is done on p5879, lines 6-10. There is no way other researchers could even try to do comparable model runs

without better information. I suggest that the authors add a Supplement to the paper where detailed information about the emissions is given. At least the following info is needed: Emission inventories used for the different parts of the domain. Emission amounts of the different emitted species for each emission sector (both total and Danish emissions). Make sure that the information is detailed enough that someone who wants to try to reproduce the study could get a system with similar emissions for all sectors. Reproducibility is an important criterion for scientific papers. It is also important to know the emission amounts for different emission sectors to get some idea about the relative importance of the different sectors per ton of emission.

Answer 22: We think that the present description of the emissions included in the paper should be sufficient for anybody to reproduce the results. We state that: "Emissions are based on several inventories, including EDGAR (Olivier & Berdowski, 2001), GEIA (Graedel et al., 1993), retrospective wildfires (Schultz et al., 2008), ship emissions both around Denmark (Olesen et al., 2009) and globally (Corbett and Fischbeck, 1997), and emissions from the EMEP database (Mareckova et al., 2008). For a more detailed description of the DEHM model, see Brandt et al. (2012).". It should also be obvious that EMEP emissions are used for the European domains and that global emissions are used in the hemispheric domain. Of caouse, it should in principle be possible to reproduce any results that are published in the literature, however, it is not at all a normal condition that all emission databases or emission amounts should be described in detail in modelling papers or even published in an appendix. To our opinion, anyone can reproduce the results, based on the information given in the paper using a similar model.

Reviewer 23: Page 5789-5881: Section 2.3 (The tagging method) is quite long but unfortunately it does not contain any detailed information about improvements by using this approach. One thing that could make this paper more interesting for the ACP would be if the improvements due to using the tagging method (compared to the more "conventional" difference between two runs) could be illustrated with numerical examples. It would be easy to check this (I guess the authors must have done it) and it would be valuable information for the modelling community if the magnitude of the problem could be illustrated. How different are the results without the tagging method? What is the magnitude of the numerical noise in the DEHM model? This information is important in judging one (of many) uncertainties in the EVA model system. The authors just state "These disadvantages must be weighed against the increased accuracy" but do not give any details about the outcome of their weighing of the alternatives. Some details about this should be given in the main text (more extensive results/figures can be given as Supplementary material). My guess (and experience) is that most modern advection schemes, used with proper time steps, would have relatively small noise problems when using the difference method. Since implementing a tagging scheme is costly it is important to know if it is worth the cost.

Answer 23: We do not think that the description of the tagging method over 1½ page is quite long. We also think that we have explained the tagging method to an appropriate level of detail, where everyone wanting to reproduce the method is able to do so. Furthermore, we have argued and explained why we have developed the tagging method. We think that a full testing of the tagging methodology is not within the scope of this paper, since it would be twice as long and we cannot handle all issues in one paper. We plan to submit a paper about the tagging method alone in another paper. Numerical noise is a challenge in all modern higher order advection schemes. Using a difference method, as proposed by the reviewer, is not within the group of modern advection schemes and the difference method itself introduces a lot of numerical noise. However, this paper is not about advection schemes since advection schemes is a scientific area in itself. We would, however, like to refer to the reference Frohn et al. (2002), where the advection scheme used in DEHM is thoroughly described and tested.

Reviewer 24: Page 5879, lines 24-25: a better reference about the Gibbs phenomenon is needed (Brandt et al., 1996, is not the best reference for this, it does not discuss the details of the phenomenon at all)

Answer 24: The Brandt et al. (1996) is provided in order to illustrate the Gibbs phenomenon in its worst case, when handling a single strong point source. We could have referenced Gibbs (1899), however, as scientific papers does not make a reference to Newton every time the second law is applied, we found that is more appropriate to reference to an example of Gibbs phenomenon, where one can see the phenomenon in action using a computer model.

Reference:

Gibbs, J. Willard (1899), "Fourier's Series", *Nature* **59** (1539): 606, doi:10.1038/059606a0, ISSN0028-0836

Reviewer 25: Page 5881, lines 12-14: "On the European scale, a gridded data set was obtained from the EUROSTAT 2000 database (http://epp.eurostat.ec.europa.eu/), covering Europe." This is not precise enough to be able to find which data were actually used in the study. This must be specified better (the EUROSTAT database is huge). Also, the definition of Europe is somewhat imprecise. Eurostat sometimes only include EU-countries, but I guess you have at least included Norway and Switzerland as well? How about other (non-EU) European countries? And did you include all of Russia (or just part of the country)? What about Turkey and the parts of North Africa included in DEHM domain 2?

Answer 25: This is actually an error in the manuscript. Population data was not obtained from the EUROSTAT 2000 database, but from EMEP (www.emep.int) already gridded to the EMEP grid and Including all countries in the DEHM model domain 2 including Turkey, Russia, etc.

Reviewer 26: Page 5881 line 25 – page 5882 line 4: Newer studies by e.g., Pope et al. (2009, 2011) show non-linear ER-relationships for cardiovascular disease. Newer references regarding this are needed.

Answer 26: Non-linear ER relationships are expected for chronic YOLL in general, but Pope et al. (2009, 2011) also shows that the ER relationships can be considered linear in the present range of air pollution concentrations. Furthermore, the non-linear relationship is difficult to quantify based on the present studies, as is the challenge of distinguish the toxicity of the different particle compositions. Therefore, we are using linear ER relationships.

Reviewer 27: Page 5882, line 10: NH4+ is missing in the list of included compounds (as far as I understand from the rest of the paper particulate ammonium is included in the health impact estimates).

Answer 27: $NH4^+$ is now included in the list "NH_4^+ is furthermore included when it is contributing to the particle mass by reactions with sulphate or nitrate".

Reviewer 28: line 18: Krewski et al., 2009 have a very extensive analysis of the data and provides many different alternatives. They also show significantly different ERFs for different time periods (and locations). This certainly could be an indication that the health impacts are different for different types of particles. If the health impacts really were due to the PM2.5 mass it is difficult to understand the results in Krewski et al., 2009.

Answer 28: This is a very open interpretation of the results in Krewski et al. (2009). The extended follow-up and spatial analysis of the American Cancer Society Study linking particulate air pollution and mortality, clearly concludes that "the epidemiological results reported here are consistent with those from other population-based studies, which collectively strongly support the hypothesis that long-term exposure to

PM2.5 increases mortality in the general population" (Krewski et al., 2009). Of cause, there is a statistical variation in the data, but it cannot presently be concluded whether this is due to statistical variability in the data or whether the health impacts are different for different types of particles. See also answer 4 & 5.

Reviewer 29: lines 21-22: "we assume a 10-yr time-lag between the exposure pulse and subsequent changes in mortality risks for the relevant age-groups above 30". Please motivate the choice of a 10-yr time-lag. And what is the reason for assuming that the "relevant" age-groups are above 30? And is it above 30 at time of exposure or 10 years later? In general, I think this section is not written with enough detail to be easily understood by many ACP readers (and there are no references regarding how chronic mortality risks are converted to mortality cases and there is no unit for the 1138 per 100 000 individuals; I guess it must be YOLL?). It would be good to explain this a little bit more detailed (possibly in a Supplement).

Answer 29: We have shortened this section as the chronic mortality is explained also in section 2.6. The unit for the 1138 per 100000 individual is YOLL.

Reviewer 30: line 24: You use the RR=1.06 from Pope et al., 2002. What do newer studies suggest for RR? A discussion about this is needed and some discussion about uncertainties.

Answer 30: See answers 4 and 5.

Reviewer 31: Page 5883, line 1-2: "Several studies have established a link between sudden infant death and exposure to SO2." The references to these "several studies" are missing. Also, Table 1 indicates that PM is used rather than SO2.

Answer 31: It is correct that PM is used rather than SO2 for sudden infant death. The sentence has been removed in the paper.

Reviewer 32: lines 2-4: "It has also been established that O3 concentrations above the level of 35 ppb involve an acute mortality increase, presumably for weaker and elderly individuals". References are missing to studies that have established this. As far as I understand there is no scientific basis for a safe level at 35ppb for ozone?

Answer 32: See answer 35

Reviewer 33: line 5: The reference to an AEA technology report (Hurley et al., 2005) for two of the ERF's is not very satisfactory for a peer-reviewed research paper. I suggest that you give the original references to the studies that provide the ERF's you have applied. Otherwise it is very difficult to critically review the assumptions and to check the newer literature that may be important to evaluate this work.

Answer 33: We think that the reference to Hurley et al., 2005 is appropriate, since this paper is not a review of ERF's but a documentation of a new integrated model system. We also refer to answer 35.

Reviewer 34: lines 6-8: "Finally, there are studies, which have shown that SO2 is associated with acute mortality and for this response we apply the ERF identified in the APHEA study – Air Pollution and Health: A European Approach (Katsouyanni et al., 1997)." Only one reference is given here (Katsouyanni) and it is more than 15 years old; I guess there are newer studies that have either confirmed or modified the ERFs for SO2? Much of the association of SO2 with acute mortality may be that it is working as a surrogate for other substances. Is there not a risk of double counting impacts if this is the case? E.g., Chen et al., 2012 (Env. Res. 118, 101-106) found that the mortality impact of SO2 in Chinese cities did not persist after adjustment for NO2.

Answer 34: SO_2 is a gas most commonly originating from combustion of sulphur-containing fuels and thus in many locations closely correlated with the other major combustion-derived pollutants: fine particles, NO_x and CO. SO_2 can cause bronchoconstriction in both normal and asthmatic subjects but at concentrations well above those found in ambient air. Evidence for direct effects of SO_2 as a gas (other than by contributing to sulphate secondary particles) remain limited and somewhat controversial but in contrast to NO_x there are several studies suggesting some independent effect of this gas. The plausible mechanism is that by adsorbing to particles SO_2 renders these more acidic (by oxidation into sulphuric acid) and adds to the stress in the peripheral airways caused by air pollution. In addition a unique study in Hong Kong by Hedley et al. (2002) where mortality was significantly lowered during an intervention on SO_2 alone (not affecting the other pollutants of consideration) suggest a separate role of SO_2. Results from the analyses of the American Cancer Society (ACS) cohort also suggest an independent adverse effect of SO_2 (Pope et al. 2002). It is plausible although not clear that the effects of SO_2 only become evident in locations with a significant contribution of SO_2 to the air pollution.

In contrast to other recent reviews, but from the point of view that independent effects of SO_2 do seem to exist and that they are plausible from a mechanistic point of view and despite the risk of double counting effects in EVA we have used a short-term CRF of SO_2. The CRF for all-cause mortality for a 24h previous day 10 µg/m3 increase in SO_2 used in EVA is 1.012 (95% CI 1.007-1.016) and is taken from the study in Athens by Touloumi et al. (1996).

References:
Hedley et al. Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: an intervention study. Lancet (2002) vol. 360 (9346) pp. 1646-52 Page 50 of 76

Pope et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA (2002) vol. 287 (9) pp. 1132-41.

Touloumi, G., Samoli, E., & Katsuyanni, K. 1996: Daily mortalty and "winter type" air pollution in Athens, Greece - a time series analysis within the APHEA projecy. J Epidem Comm Health 50(suppl 1); S47 - S51

Reviewer 35: lines 10-13: Regarding chronic exposure and lung cancer, what time-lag do you assume from exposure to the time of disease onset (or diagnosis)? I also have to say that I find it hard to believe that pure secondary inorganic aerosol would be as carcinogenic as combustion generated primary particles.

lines 14-20: Chronic bronchitis (CB). The ERF used for chronic bronchitis and PM2.5 exposure is based on an old paper (Abbey et al., 1999) based on a single study from California and a distinct subgroup of the population (seventh-day Adventists); furthermore the results from this study seem to be not statistically significant. There are many new studies available and it seems like the issue of chronic bronchitis is more related to traffic emissions (and distance to major roads) than to total PM2.5 concentrations (e.g., one European-wide study is Sunyer et al., 2006, Occup Environ Med 63:836–843; that showed no association between PM2.5 (or S-concentration) and chronic bronchitis; Smoking, rhinitis, poor education, and low social class were associated with CB for both genders, occupational exposure for males, and traffic intensity and NO2 for women). It is likely that the methodology used in the EVA model gives erroneous estimates of the impact of different emission sectors on CB. Unless the authors can show some evidence of a relation between ammonium sulphate/nitrate and CB this morbidity cost should not be included for secondary inorganic particles. In any case the paper needs an update on the literature about chronic bronchitis and air pollution. It is also not clear how the used RR=1.007 (unit?) is transformed into the value 8.2E-5 cases/(ug/m3) in Table 1.

lines 19-20: Why did you choose a Norwegian study rather than a pan-European estimate? Some motivation for this is needed. lines 21-26: Restricted Activity Days (RAD). This section gives very little details and the only reference given is the old ExternE Methodology 1998 Update. The ExternE methodology seems to be based on very old morbidity data (from 1976-1981) from cities in the USA. The ERFs are based on rather crude estimates of fine particle concentrations. Furthermore there were large year-to-year differences in the results. In ExternE there is an unexplained downscaling of the E-R functions by a factor of two for European implementation. There is also a scale factor (0.6) to transform from PM2.5 to PM10 effects. Regarding RAD I think that the presentation needs to be clearer (it is ok to put details in supplementary material but it has to be possible to find out what assumptions the EVA model is based on, and what data/relationships have been used). Did you use the rescaling of E-R for RAD from US to the European region? And if so, please motivate this. Why would Europeans be less sensitive to PM than Americans? Also, if there are no newer studies than the ones cited in ExternE 1999 I think it is highly questionable to include the RAD at all in the model system. There are at least some newer works that could be used instead, e.g., a study of air pollution and Disability Days (Stieb et al., 2002, Environ. Research 89, 210-219), which seems more relevant for the generally lower levels of air pollution in Europe. And there is a study from Norway (Hansen and Selte 2000). There are probably a lot of more recent papers to consider.

lines 26-28: Hospital admissions. Here again very little information is given. The ExternE methodology from the 1990s should be updated to take into account new information. And many details are unclear in the ExternE 1999 reference. How did you handle conversions from American to European conditions. Transformation from PM10 (or PM13) to PM2.5? How do you motivate the assumptions you have used? Some of the studies, that ExternE is based on, only looked at people aged 65+. Did you restrict your study to this age group? Many of the ERFs used in ExternE in the 1990s were considered very uncertain already then and it is not reasonable to continue using these old assumptions when the field of air pollution and health effects have expanded enormously during the last decades. For example, the case of Cerebrovascular (CV) hospital admissions due to PM exposure: This is based on a single study from Birmingham (UK) covering a 2-year period (1992-1994). The results showed that CV admissions were (just barely) significantly associated with PM10-concentrations. Since these results are only for a single city it is difficult to generalise to a whole continent. If the effect was mostly due to fine particles from traffic (or other combustion sources) there may be no effect (or at least a very different effect) from secondary inorganic aerosol (ammonium nitrate and sulphate) produced on larger scales. There are many new studies that should be taken into account when trying to estimate CV (and other) hospital admissions due to air pollution.

The authors need to do a literature review of the field to get an updated view of recent developments regarding health impacts of air pollution. Using ExternE methodology from the 1990s is not good enough anymore. ERFs needs to be updated and very importantly if you want to compare impact of different sources scientifically motivated uncertainty ranges have to be used. Some examples of newer studies include: Bedada et al., 2012 (Env. Health 11:77) that looked at transient ischaemic attacks and minor strokes in the Greater Manchester region between 2003 and 2007 and found a modest association between NO and these health impacts but not for PM10; this points towards traffic emissions being the most important in this case (NO may be a surrogate for some other unmeasured pollutant(s) such as ultrafine particles (UFP) or there may be some direct effect of NO). Other studies for various locations have shown other results, e.g., Mechtouff et al. (2012, Int J of Stroke) found no association between air pollutant exposure and ischaemic stroke in Lyon (France), Villeneuve et al (2006, Eur J Epidemiol) found no association between PM2.5 or PM10 and stroke in Edmonton (Canada), but some association for CO and NO2 (possibly indicating vehicular traffic as the important factor) with ischemic stroke during the summer half-years. O'Donell et al. (2011, Epidemiology) found a negative (but non-significant) association of PM2.5 and ischemic stroke in a study covering 8 cities in Ontario (Canada). These are just a few, more or less random, examples from recent literature (results regarding CV hospital admissions are inconsistent

between different studies); obviously the authors need to make a more complete literature survey if they intend to include, and try to put monetary value on, various health impacts from very different pollutant sources. As far as I can judge most studies point towards relatively small effects from air pollution on CV disease and the studies that show effects tend to single out vehicular traffic as the most likely source of the impacts. The literature is rather consistent that traffic is the responsible air pollution source for stroke (see, e.g., the references in Andersen et al., 2012, Stroke 43, 320-325). The literature on respiratory effects hospital admissions also needs to be taken into account (I have not had the time to go through it in detail but ExternE 1999 is almost certainly outdated): Sunyer et al., 2003 found no effects of SO2, except for asthma in children but this effect disappeared after controlling for PM10 or CO. In ExternE 1999 effects of ozone on hospital admissions were also included but they are not included in your Table 1. Why were they excluded in this work? Health effects for asthmatics: How does the ExternE 1999 ERFs compare to more recent scientific literature? ExternE1999 points out that bronchodilator use and wheeze E-R functions may have important problems of representativeness since they are based on small numbers of subjects, in only one European country (Netherlands).

[73 children in two small Dutch towns during one winter!] This needs to be updated to take into account other studies as well. For children the ExternE1999 cites Roemer et al. 1993, which has been cited ca 170 times (according to Web of Science) so there is very much new information available to update the E-R relationships. Since there are so many new studies published after 1993 I have not managed to go through all the literature but the authors need to do this in order to update the EVA model. A relatively brief look at published results show that in many studies respiratory conditions are associated with NO2, PM10 and CO, which suggest that traffic-related pollutants may be most important. Some studies also show that ozone may have an impact. A review of 22 panel studies on particulate air pollution and children were published in 2004 (Ward & Ayres, Occup Environ Med 2004, 61:e13). An important implication of the review is that "the degree of heterogeneity between panel study results questions the transferability of estimated effect sizes between locations or populations, and limits the use of summary measures in quantitative risk assessment" Another thing Ward and Ayres discuss is the important issue of publication bias, which I think should be considered when adding different health effects in integrated assessment models. Another paper of great relevance for children and asthma is the PEACE study by Roemer et al. (1999, Eur Respir J 1998; 12: 1354–1361), which included 14 different European centres (in 10 countries) including one urban and rural panel per centre (totally 2100 children in 28 locations). In conclusion, the PEACE project did not show clear effects of PM10, black smoke, sulphur dioxide or nitrogen dioxide on morning or evening peak expiratory flow or the prevalence of respiratory symptoms and bronchodilator use. Results for health impacts on asthmatic children in the literature are not fully consistent and it is not reasonable to continue to use the ExternE1999 (Roemer et al., 1993) for E-R relationships. In fact Roemer et al., 1999 indicates that it is better to assume no impact than the Roemer 1993-E-R.

Weinmayr et al. have relatively recently written a review and meta-analysis of shortterm effects of PM10, NO2 on respiratory health among children with Asthma (2010, Environ Health Perspect 118:449-457). This could be a good starting point for developing new ERFs for the EVA model; but the relation between NO2 and health effects may well be due to NO2 being a marker for the urban air pollution mix (UFP, EC etc) rather than direct effects of NO2; this also means that different components of PM10 may have different impacts for asthmatics.

Table 1: How do you motivate the setting of the E-R relationship for PM2.5 by scaling PM10 values by a factor of 1.67? Some studies have even shown higher impact of PMcoarse than PM2.5 on asthma. The E-R relationship in ExternE1999 for cough is also based on an old and very limited study (Pope and Dockery, 1992, 39 children in Utah Valley, USA, during one winter).

This also needs to be updated to take into account work during the last 20 years. Weinmayr et al., 2010 included cough as well as other symptoms in their review/metaanalysis. Table 1: You seem to have used a two times higher E-R for cough (4.46E-1 days/[ug/m3]) than ExternE1999 (0.223). What is the motivation for this? Also please motivate the scaling factor used to transform the original PM10 E-R to PM2.5 E-R.

Regarding asthma and ozone, ExternE 1999 included an ERF for asthma attacks due to ozone. This seems to be excluded in the EVA model, probably for good reasons, since the ExternE reference is from 1980, but this should be mentioned/motivated in the text. Adult Asthma ERFs are also based on a single study (Dusseldorp et al., 1995) of 32 persons near a large steel industry in the Netherlands during a limited time (67 days, Oct-Dec). This also needs to be updated. Two studies by Penttinen et al. (2001, Eur Respir J 17: 428–435; 2006, Inhalation Toxicology, Vol. 18, No. 3, Pages 191-198) could be useful; the 2001 study showed that number concentration of particles (PNC), but not particle mass (PM) was negatively associated with daily PEF deviations. Particle number concentrations in the size range smaller than 0.1 um had the strongest effects; the results in the 2006 study suggest that the negative effects of PM2.5 on PEF in adult asthmatics are mainly mediated by particles related to local combustion sources. There is also a study from Erfurt, Germany, by Peters et al. (1997, Am J Respir Crit Care Med 155: 1376 – 1383) showing larger health effects of the number of ultrafine particles than those of the mass of the fine particles. In another Erfurt study von Klot et al. (2002, Eur Respir J 20: 691–702) found evidence for independent effects of fine and ultrafine particles.

Answer 35: We have to say, that it is impossible to answer all these carpet bombing of comments in detail, as it would take more than several paper to handle all these questions. Of cause, it is important to have updated ERF's in an integrated model system like the EVA system presented in the paper. However, the mail goal of the paper is not to review all ERF's with respect to all kind of impacts as this is far out of scope for the paper and the paper would be tremendously long. The paper is focused on presenting the integrated model system EVA, using a state-of-the-art regional scale air pollution models including a tagging method for handling the non-linear chemistry when calculating contributions from emission sources. The concentration-response functions used in the system are documented to the extent, which can interest readers of ACP, and this section should not be longer – rather shorter, since the paper is already relatively long. All concentration-response functions can be discussed and questioned, however, the important issue is, that we clearly present what functions we have used, so that results in the paper can be reproduced by others if they wish, and of cause are able to update the concentration-response functions in future works. To the end of this, I would also like to emphasize that all the concentration-response functions in the EVA system have been reviewed and documented in a report "CEEH Scientific Report No. 7a - Description of the CEEH health effect model" by Bønløkke (2011) available at

http://www.ceeh.dk/CEEH_Reports/Report_7a/CEEH_Report_7a.pdf

Reviewer 36: Reviewer Page 5884, lines21-22: VOLY=52 000 Euro. This seems to be a mistake? You quote Alberini et al. (2006) for this value but their study gives two values: 54 000 Euro and 163 000 Euro. If you used a lower value please provide an explanation for this choice.

Answer 36: Yes – this is a mistake in the table, which has now been corrected.

Reviewer 37: Page 5885, lines 1-3: Reference is missing for the panel advising US EPA.

Answer 37: The sentence has been removed as it was not important for the results in the paper.

Reviewer 38: lines10-12: Regarding purchasing power parities: This is not explained/referenced well enough (considering that this is a journal devoted to atmospheric chemistry and physics and not economy).

Did you use the PPP for Denmark for all of Europe or just for Denmark? Why did you choose 2006-prices? How does the Danish PPP compare to European average levels?

Answer 38: The purchasing power parities is just a technique used to determine the relative value of currencies – this explanation is now inserted into the text for clarification. As explained we used the PPP for Denmark, but this can be scaled to average values for Europe by decreasing the valuation by 30% on average (see Brandt et al. (2010)) – and is a easily applied scaling factor. We have used 2006 prices for historical reasons, and this can be easily transformed into different years using inflation rates, with simple scaling factors.

Reference:
Brandt, J., Silver, J. D., Gross, A. and Christensen, J. H., 2010: "Marginal damage cost per unit of air pollution emissions", Roskilde: National Environmental Research Institute. 23 p. Specific agreement 3555/B2010/EEA.54131 implementing framework contract ref. no. EEA/IEA/09/002. (The results in the report was used in "Road user charges for heavy goods vehicles (HGV) Tables with external costs of air pollution", EEA Technical report No 1/2013, ISSN 1725-2237, pp. 88.)

Reviewer 39: Lines 12-13: Why is infant mortality valued higher (no reference given).

Answer 39: The valuation of mortality is found from willingness to pay methodology. The reason for the higher value for infants is that people are more willing to pay for additional life years for infants than for adults.

Reviewer 40: Line 13: "there is no cancer premium for adults": What does this mean? This is not a health/economy journal. At least provide a reference to explain this.

Answer 40: The sentence has been removed as it is not important for the results in the paper.

Reviewer 41: Lines14-18: There are no references given here about the morbidity costs. Please add this.

Answer 41: There is a reference to The Diagnosis Related Groups (DRGs), a patient classification scheme which provides a means of relating the type of patients a hospital treats (i.e., its case mix) to the costs incurred by the hospital.

Reviewer 42: Line 18: For work loss days, a 20% productivity loss has been added. What does this mean? It is totally unclear to me.

Answer 42: The sentence has been removed.

Reviewer 43: Line 19-20: The reference about chronic bronchitis is erroneous. Pizzol et al. (2010) deals with Pb and IQ-impacts as far as I can see.

Answer 43: Yes – but as stated: The complex calculations for chronic bronchitis is explained in Pizzol et. al. (2010).

Reviewer 44: Line28: "3000-4000 people die . . . due to present levels of atmospheric pollution". These "present levels" the reference is from 2002 so I guess it is not exactly present levels of air pollution but rather levels about 10-15 years ago. Since the reference is a Danish journal I could only check the English summary and according to that the figure is rather 5000 deaths. I also note that Raaschou-Nielsen et al.

(2002) state that “The ultrafine particle fraction may cause a much greater impact on health than indicated by the mass.” I do agree with this comment.

Answer 44: The number is published in a Danish peer-reviewed journal, which of cause is not optimal. Howewer as these are the latest numbers calculated, we have to include this Danish reference.

Reviewer 45: Page 5886, lines 3-5: “Support of adverse health effects of PM is also found in a long range of laboratory, animal and human experimental studies.” References are missing for this very important part. More details are also necessary to show what type of PM these laboratory experiments have shown adverse effects for. Since epidemiology can very seldomly prove that observed associations are due to a specific air pollutant (and not some other co-emitted/correlated species) laboratory (exposure) are crucial to find out what PM components are likely to be most toxic.

Answer 45: This is a very general statement, and we disagree that a reference is needed for this.

Reviewer 46: Lines 15-17: the sited study by Yap et al. (2012) and two other studies by the same group (Beverland et al., Environ Health Perspect 120:1280–1285 (2012) and Beverland et al., Atmospheric Environment 62 (2012) 530-539) are very interesting and show associations between Black Smoke (BS) and cardiovascular mortality (and thus not of PM2.5 or PM10). They also highlight the critical importance of reliable estimation of exposures on intraurban spatial scales. This means that in order to estimate the mortality impact of air pollution in a realistic way the relatively coarse resolution of the DEHM model is not sufficient for the purpose of health impact assessments for a majority of the European (or Danish) population. Some method for estimating the strong concentration gradients from major roads (and other large sources) are needed for this.

Answer 46: Se answers above – especially answers 4 and 5.

Reviewer 47: Lines 17-18: I also note that the study referred to about morbidity effects (Schwartz et al., 2012) shows an association between black carbon (BC) and blood pressure. Again, this points towards the importance of traffic emissions (PPM and possibly UFP) for this health impact (rather than total PM2.5 mass). Another study that has shown the importance of local traffic and other combustion (EC) for cardiovascular disease is the Helsinki panel study of Exercise-Induced Ischemia on elderly subjects with coronary heart disease by Lanki et al. (2006, Environ Health Perspect 114: 655-660). The study indicates that PM2.5 originating from local traffic is the most toxic and also that the effect seen of long-range PM2.5 was probably more related with carbon products than with secondary sulphate.

Answer 47: Se answers above – especially answers 4 and 5.

Reviewer 48: Lines 26-27: “No simple pattern, however, has emerged on which sources or which PM constituents matter the most.” It is true that this issue is not simple but there are plenty of indications in recent literature that local combustion sources (e.g., traffic) and EC/BC and other primarily emitted particulate components (OC, metals) are more consistently associated with health effects of different types than secondary particle components and total particle mass. The EC/BC etc. can of course be markers for unmeasured co-emitted components or UFP from traffic.

Answer 48: Se answers before – especially answers 4 and 5. However newest studies indicate that UFP is not a good marker for health effects in contrast to previous hypotheses.

Reviewer 49: Lines27-29: “the evidence of links between health effects and the sulphate fraction of particles is good” I do not agree completely with this. There are certainly studies showing an association

between observed (and/or modelled) sulphate and health impacts but many of these are rather old and did not include EC/BC or other primary PM2.5 species. In the 1970s and 1980s sulphur was to a large extent emitted from High-sulfur diesel fuels (with co-emissions of BC, PAHs etc), Residual oil (with co-emission of metals, e.g., Ni), Steel mills (various metals co-emitted), Coke ovens (organics and PAHs co-emitted). So a lot of the co-emissions may have been highly toxic and could have influenced the epidemiological studies. As far as I know there is no indication that sulphate causes cancer and it seems generally accepted that sulphate is not harmful per se (e.g., Schlesinger and Cassee, 2003 Inhalation Toxicology 15, 197-235; Schlesinger et al., 2006 Inhalation Toxicology 18, 95–125).

Answer 49: We see this as a comment or an opinion and has no further comments.

Reviewer 50: Page 5887, line 8: As far as I can see Andersen et al., 2007 do not show any specific results for nitrates (there are some results for secondary PM10 but this seems to be the sum of ammonium, sulphate and nitrate).

Answer 50: As nitrate is a part of the secondary formed inorganic aerosols and since these are treated equally, we keep the argumentation.

Reviewer 51: lines 9-10, Regarding the association of ammonium ions and cardiovascular disease: Peng et al. (2009) is an interesting study but the conclusions are that "Ambient levels of EC and OCM, which are generated primarily from vehicle emissions, diesel, and wood burning, were associated with the largest risks of emergency hospitalization across the major chemical constituents of PM2.5." As pointed out by Brandt et al., the effect seen in the single-pollutant model for ammonium is not statistically significant in the multipollutant models. If this is the only "evidence" for an impact of ammonium on human health I would be very reluctant to include ammonium in any health related impact assessment model! And I think that Brandt et al. should change the text in this section to rather indicate that epidemiological studies show very mixed results for secondary inorganic aerosol; especially for ammonium and nitrate there are very few studies that connect them with health effects (it is a bit unfair to use the Peng et al. (2009) reference to (indirectly) motivate that agriculture has a huge impact on human health when the conclusions of Peng et al. points out completely different sources). It is also important to point out that there are many studies that included secondary inorganic components but did not show any significant association with health effects (see e.g., Rohr and Wyzga, 2012).

Answer 51: The main issue as explained previously is that it is impossible at this stage to assign different toxicity to different particle composition as Rohr and Wyzga (2012) conclude in their paper.

Reviewer 51: Page 5888, lines 7-8 (and the section above): "Thus the choice in this study: to assign equal health effect to all components of particles is in line with other recent major reports." This may be true, but in my opinion it is not a good enough argument, for use in a scientific paper in ACP, to refer to a number of "technical reports" that for various (maybe not always totally scientific) reasons have chosen a certain approach. Methods should be based on scientific arguments that can be checked/verified/discussed (the original scientific literature should be cited rather than CAFE/DEFRA/ExternE/NEEDS reports, which often are too cumbersome to check for the reader and may at least partly lack scientific peer-reviewing; it is good to include the reports in the reference list for the interested reader but crucial arguments should be taken from peer-reviewed journals when at all possible).

Answer 51: We would again like to refer to answers 4 and 5. We agree that it is a major scientific question which components of $PM_{2.5}$ are more toxic than others (if any). As long as we only have clear evidence for the health impacts associated with the total $PM_{2.5}$, we have to apply this in the assessments.

Reviewer 52: lines 8-12: The review by Rohr and Wyzga (2012) does not support the choice to assign equal health effect to all components! On the contrary the review gives a very different picture than the text here suggests. Rohr and Wyzga show that there is growing epidemiological evidence (supported by controlled human exposure experiments) that suggests that EC and OC components are most strongly associated with adverse health outcomes. Toxicological studies suggest that various metals are important and carbon-containing components have been implicated as well. Please read the paper carefully and revise the usage of this reference.

Answer 52: See answer 51. It is presently a hypothesis that EC and OC are more toxic and the results in Rohr and Wyzga points towards future research. However, the main problem in Rohr and Wyzga is that they include both short-term and long-term effects in a single evaluation, which might be questioned.

Reviewer 53: Another recent review by Kelly and Fussell (2012, Atmos Environ 60, 504-526) about size, source and chemical composition as determinants of PM toxicity complements Rohr and Wyzga and is also useful to refer to in this paper. The review covers both epidemiology and toxicology and gives a lot of references that show the importance of carbonaceous aerosol, metals, particle size, traffic emissions (e.g., health risks associated with distance gradients from major roads and heavy traffic). Primary and secondary aerosol components are covered and among the conclusions drawn in the review, based on epidemiological and toxicological studies are: EC is more consistently associated with health-effects than OC (which could partly be due to masking of primary OC-effect by less toxic secondary OC); Some metals are associated with adverse health effects at or near ambient concentration levels (especially Ni and V); UFPs and species that are rich in this fraction have toxic properties and strong oxidizing potential; Evidence is consistently growing for an association between traffic emissions and detrimental effects on human health; There is generally less compelling evidence to connect secondary inorganic particles with adverse health effects. In yet another recent review, Janssen et al., (2011, Env Health Persp) showed that there is a much stronger association of black carbon (BC) than total PM2.5 to mortality; For cohort studies, pooled estimates for all-cause mortality per 1 μg/m3 were 5–14 times higher for BCP than for PM2.5! Also for morbidity, estimated effects of a 1 μg/m3 increase in BCP were greater than estimated effects of a 1 μg/m3 increase in PM2.5.

Answer 53: We would like to refer to answer 51.

Reviewer 54: lines 12-14: The sensitivity study included by Brandt et al., uses a much too small range of uncertainty in the toxicities for primary and secondary PM2.5. Just changing toxicities by 30% is not meaningful and there is no scientific argument for the choice of the factors 1.3 and 0.7 for primary and secondary particles respectively. The approach used by Griffiths (2011) (2.0 and 0.0) is more useful since it better represents the uncertainty about toxicities. Even this just gives a rough range of possible impacts and nothing better seems possible with the present state of knowledge about health impacts of particulate matter.

Answer 54: The choice of factors in a sensitivity studies is only meant to demonstrate the sensitivity of the hypothesis of assigning different toxicity to the different particles. Any choice is wrong (also the factors 2 and 0) as we yet do not know the real values. We have chosen the values applied in the NEEDS project.

Reviewer 55: Page 5889, line 5: Does the region “Europe (EU)” include the rather large parts of Turkey and Northern Africa that are included in DEHM domain 2?

Answer 55: The region Europe refers to the whole model domain 2 in DEHM, including Turkey, Russia, etc. However, results are not applied for Northern Africa covered by the model domain as this is a European study.

Reviewer 56: line 22-24: “However, the source-receptor relationships are non-linear due to the effects of atmospheric chemistry, and therefore the scenarios DK/1–10 and DK/all are not expected to be equal.” Unless I misunderstand the methodology, a large part if the difference between DK/all and DK/1-10 for “total N” is that the emission sources for NH3(g) and precursors for HNO3(g) are in different SNAP sectors, which leads to a double counting of NH4NO3(p) costs in the DK/1-10 case. NH3 is from agriculture and HNO3 (from NOx) largely from traffic and power plants. Since both NH3 and HNO3 are gases the formation of the “same” NH4NO3 particles will be counted as an effect both for the DK10 and DK7 (or DK1). Similarly, the ammonium part of ammonium sulphate could be counted for two different emission sectors. I would not call these effects “non-linear effects of the atmospheric chemistry” but rather an artefact of the way you count costs for different sectors (the NH4NO3 cost should have been split between two sectors rather than counted twice).

Answer 56: Yes – there could be some “double counting” since we examine the impacts of the emission sectors individually. As explained previously, it is important to answer the right questions in assessments since one scenario cannot be extrapolated into other scenarios. However, the non-linear atmospheric chemistry is the other main reason for not adding up to 100%. The most important factor in this being the NOx emissions eventually changing OH concentrations in the atmosphere and finally changing the whole chemistry – e.g. the reaction rates from primary emitted species to secondary species. This non-linearity cannot be disregarded in assessments.

Reviewer 57: Page 5890 line 25-p 5891 line 6: The figures given for YOLL and number of premature deaths “app. 49 000 in Europe and app. 8500 in Denmark”, “approximately 4600 premature deaths in Europe and approximately 800 premature deaths within Denmark” give the impression that these are relatively accurately estimated with little uncertainty. This needs to be revised to take into account the huge uncertainties in these estimates. Estimated ranges should be given instead of singular numbers. It is not trivial to estimate the uncertainty but it is very important! Uncertainties in emissions should be included and also the uncertainty in impacts of secondary/primary PM (and in impacts of other pollutants). Also an estimate of the uncertainty in the population exposure is needed (due to using a very coarse model resolution for urban populations, living relatively near large emission sources).

Answer 57: See answer 8. Furthermore, as we assign the word “app.” to every number, we clearly indicate that these numbers are associated with uncertainties.

Reviewer 58: Page 5891, Section 4.2/Table 4: Where do you count impacts of VOC emissions on ozone?

Answer 58: We count the impacts of VOC emission on ozone under “Total N” as explained in the table “Total N is the sum of the external costs of O^3 and NO_3^-“.

Reviewer 59: Page 5892, lines 3-16: The results regarding different emission sectors contributions to health impacts needs to be updated and compared to the published study by Griffiths (2011, Air Qual Atmos Health 4:189-197). The relative contributions of the different sectors are so crucially dependent on the assumptions regarding toxicity of secondary inorganic aerosol compared to primary particles that this must be discussed in more detail here. I suggest that estimated ranges of importance are given instead of singular numbers. Also the impact per emitted ton may be interesting since some sectors have very small emissions but could potentially still be of interest for emission reductions (results per ton emission could be put in Supplementary material). Based on the available scientific evidence there is no support for assuming equal health effects for all PM2.5 components. This assumption is unrealistic, according to current knowledge, and will lead to results that are very likely misleading for total health impact assessment (and if

applied in economic valuation of benefits from different potential emission reductions likely will lead to erroneous conclusions).

Answer 59: The reference to Griffiths (2011) has been included in the paper. We have chosen not to include the the impacts per tons emissions, since these are mainly interesting to decision makers and not in a scientific paper. The other comments have been addressed in previous answers.

Reviewer 59: lines 21-26: ”The difference in these numbers lies both in the difference in the emission areas (Europe or Northern Hemisphere) and in the inclusion of the natural emissions in the latter simulation assessing the impacts from the total air pollution levels as well as on the difference in taking the sum over the ten emission sectors (assuming linearity) or running all sectors simultaneously (assuming non-linearity).” I do not agree that running all sectors simultaneously means “assuming” non-linearity. If the model is run with all emissions non-linear chemistry is taken into account. I suggest you remove “(assuming non-linearity)”. However, I do not understand why you would choose to take “the sum over the ten emission sectors” for estimating total external costs in Europe. You will then double count ammonium nitrate (and part of the ammonium sulphate) as discussed above.

Answer 59: The sum of the cost for the ten emission sectors are only included to illustrate the non-linearity of the atmospheric chemistry and the possible error to be made when assuming linearity. This is also explained in the text.

Reviewer 60: Page 5892, line 27-p5893 line 6: The part about estimated costs needs to be updated to show the huge uncertainties in line with the comment above about the health effect uncertainties.

Answer 60: We think we have already addressed this comment in previous answers.

Reviewer 61: Page 5893, lines 6-10: Same comment as above about the “non-linear atmospheric chemistry effect”. Is not a large part of the 15 (or 19)% higher cost in the “Sum1-10” cases compared to “all” calculations a consequence of counting the “same” ammonium nitrate for two different emission sectors?

Answer 61: No – the main effect is from the non-linear atmospheric chemistry. The non-linear effect is not very big when considering only smaller emission areas as Denmark. However, if e.g. Germany was considered very large non-linear effects would be seen, as big countries also have big impacts on their own background concentrations and changing emissions alters the chemical regimes in a larger region.

Reviewer 62: Page 5894, line 20: “significant impact on human health as secondary particles (ExternE, 1997)” This old ExternE report is a strange reference for the statement about ammonium sulphate and nitrate. If you can find some (preferably recent) scientific paper showing significant human health impacts of secondary particles refer to it instead. Otherwise the part about significant impact on human health should be removed or reformulated to something like: ‘for which some human health impacts cannot be excluded’ or ‘that may have some impact on human health’.

Answer 62: The reference to ExternE, 1997, has been removed here, as the section is just discussing the results.

Reviewer 63: lines 24-25: “The mass of ammonium (NH_4^+) must be included in the total particle mass associated with these particles” (and the sentences before about the external costs of NH_3 emissions). Please provide some arguments for why sulphuric acid particles become more toxic by being (partly) neutralized by ammonia. If secondary inorganic particles have some direct health effects (which is doubtful at ambient concentrations) I would assume that it would be mostly due to the acidity of the particles and

this will decrease when NH3 is taken up in the particles. If this is the case the NH3 emissions could potentially even decrease the health impacts of sulphuric acid aerosol.

Answer 63: We refer to the answers 4 and 5. As NH_4^+ is a part of the total PM2.5, there is no argumentation for disregarding this species.

Reviewer 64: lines 26-28: "According to WHO (2006), it is currently not possible to precisely quantify the contributions of different chemical components of PM, or PM from different sources, to the health effects caused by exposure to PM." The WHO report is now seven years old, so maybe it should no longer be considered "current", but I agree that it is probably still impossible to "precisely quantify the contributions of different components of PM" but that is exactly what Brandt et al. have done in this study. The work assumes that all different components have precisely the same impact. And the calculated health impacts are presented in tables transformed into very detailed external cost figures, for very different emission sectors, without proper indication of the impossibility to quantify the contributions from different components. In fact, this is the weakest point of this paper and one reason for not accepting it for publication in ACP.

Answer 64: We refer again to the answers 4 and 5.

Reviewer 65: Page 5895 line24-p 5896 line5: The Danish contributions to the health related external costs need to be given with proper uncertainty ranges, taking into account the possibility that secondary ammonium particles may be essentially harmless (as discussed above).

Answer 65: We see no scientific reasons for disregarding the impacts from ammonium particles as explained before. It is true that some particles might not be toxic, but this has to be regarded as scientific hypotheses.

Reviewer 66: Page 5896, line 7: Just as pointed out about Section 2.2 above emission details are needed (preferably in a Supplement). Are 2008 emissions available for the full model domain(s)? And are they consistent with the 2000 emissions (that is, are they prepared using equivalent methodology).
Answer 66: As noted before, we do not consider it important to include tables to show the amount of the emissions. This paper is showing a methodology, which is demonstrated using some scenarios. It is not am assessment report.

Reviewer 67: Lines 10-13: Domestic wood burning is a hot topic in many parts of Europe so it is interesting to see the large change in Danish emissions from 2000 to 2008. Some reference(s) are needed about the remarkably large increase in wood use in Denmark. An explanation why there has been such a large increase would be interesting.

Answer 67: The Danish emission inventory based on domestic wood burning is made via 2500 questionaires send out to wood stove owners about their usage of wood. The emissions we have used for Denmark is from the officially emission estimations made at NERI, Denmark in 2008.

Reviewer 68: Lines 14-20: As mentioned earlier, a more useful investigation of the range of uncertainty is needed (with zero impact of secondary inorganic aerosol (SIA) as one end of the uncertainty range). Also, I think this part should not be called a "sensitivity study". Possibly the whole paper could be called a sensitivity study (and be quite useful as such) if results are provided for a proper range of possible health effects of different components. The usage of the "equal effect assumption" should not be considered a "base case" since it is probably at least as extreme as assuming zero health impact of SIA. The arguments

from WHO (2006) and others indeed indicate that it is not possible to decide on a base case in the sense of a best guess.

Answer 68: We refer to previous answers addressing this comment.

Reviewer 69: Section 5 Discussions and overall conclusions This section is very long and a lot of it is just more or less identical repetition of things from earlier sections (especially Section 4 Results and discussion). Thus, many of my earlier comments are relevant also for Section 5. I think that Section 5 can be substantially reduced since it does not add much new discussion. Below I only comment on a few things in Section 5.

Page 5896, lines 25-27: "The system represents an attempt to apply state-of-the-art science, models, data, and methodologies in every link of the impact pathway chain." This is a commendable goal, but still the EVA model system largely use, rather preliminary, ExternE ERFs from the 1990s, which is not exactly state-of-the-art science in 2013. The ExternE199X ERFs should no longer be used for valuation of air pollution impacts in scientific studies, at least not without a critical reevaluation. It would be a good idea to work on updating the ERFs taking into account health studies from the last decades in future work.

Answer 69: As stated in answer 35, the ERF's has been extensively reviewed in 2011 and this information is now included in the paper.

Reviewer 70: Reviewer Page 5897, lines 2-4 (and at other places in the paper) "The EVA system was run for different scenarios, assessing the human health impacts and associated external costs from the ten main emission sectors in Europe and Denmark" The "ten main emission sectors in Europe" should include shipping emissions but I see no discussion of this important sector in the manuscript. In SNAP it is usually included in sector 8 but I guess that you have excluded it completely from this paper. Am I wrong? If it is excluded I suggest that you add some discussion about it (to put into relation to other sources).

Answer 70: It is correct that we have only taken land based emissions into account in this paper. The ship traffic is handled in a companion paper.

Reviewer 71: lines 6-9: "From this discussion, we concluded that with our present knowledge we are not able to distinguish between the impacts from different particle types and therefore the toxicity of the particles is handled equally." As mentioned several times before this is not a very reasonable assumption, based on available evidence in the health impact literature.

Answer 71: We think we have already answered this comment and we simple disagree with the reviewer.

Reviewer 72: Lines 18-27: Probably all of this can be skipped (already discussed in section 4). Otherwise it needs to be updated to show what happens with the "order-of-importance" if SIA has negligible health effects.

Answer 72: We consider this section as the overall discussions and conclusions, and think it is necessary to summarize the key points in the paper.

Reviewer 73: Page 5898, lines 14-16: "From the results in this study, we conclude that not only the impacts on nature should be taken into account when regulating the emissions of ammonia. Also impacts on human health should be considered." The problem is that the impacts on human health of particulate ammonium are unknown (and possibly negligible). This means that it is very difficult to take this aspect into account. It

is extremely important to regulate ammonia emissions because of the severe eco-system effects so health impact arguments are not really needed to motivate this.

Lines 21-24: “The related external costs found in this work can be used to directly compare the contributions from the different emission sectors, potentially as a basis for decision making on regulation and emission reduction.” This could lead to seriously wrong decisions. The assumption that all PM2.5 mass is equally toxic could lead to a focus on reducing secondary particulate pre-cursor emissions (instead of almost certainly much more toxic primary PM components). Reducing SIA is good for ecosystem reasons (acidification, eutrophication) but will likely do very little in terms of reducing health effects. A further complicating factor is that SIA impact on climate is cooling, while a substantial part of the primary particles (soot/BC) contributes to heating.

Lines 24-26: “This study shows that the major visible and already highly regulated emission sources (e.g., power plants and road traffic) do not always constitute the most significant problems related to human health.” This is only shown by using the unrealistic assumption of equal impact of all PM2.5 mass. The majority of air pollution health effect papers from the last five or ten years clearly point out road traffic as the most important problem!

Page 5899, lines 6-10: “However, these compounds commonly share the same sources as the compounds included in this study and the health effects are likely to be included in our calculations due to their correlations with the included compounds, since the exposure-response functions used correlates the PM2.5 concentrations with the total health impacts.” This is highly questionable. We know very little about this (but many studies show associations between health impacts and metals at least). It may be partially true for the PAH, POPs, metals and dioxins, that usually contribute relatively little to PM2.5 mass but it does not make sense for secondary organic aerosol (SOA) that may contribute a lot to the particle mass. Substantial amounts of SOA may come from anthropogenic emissions of so-called intermediate volatility organic compounds (IVOC), see e.g., Shrivastava et al. (2008, J. Geophys. Res., 113, D18301). Since the DEHM model does not include SOA an important part of PM2.5 is underestimated (mostly from combustion sources). Biogenic SOA may also contribute much to PM2.5 but this could be considered mostly a natural source of PM and thereby less important for models aiming at aiding in emission reduction decisions (BSOA concentrations are indirectly affected by anthropogenic emissions through variations in oxidant concentrations in the atmosphere that affects the BSOA formation rate).

lines 13-15: “and that the economic valuation of the health impacts has been conservative, the overall results in this work can also be considered conservative.” I strongly disagree with this. The overall results cannot be considered conservative. Griffiths (2011) showed that the assumption of equal impact of all PM2.5 mass could lead to about three times larger estimates of health impacts (and thereby costs) than if (as toxicological evidence suggests) particulate toxicity resides mainly in the primary fraction.

Lines 18-20: ”The main uncertainties in the integrated model system are associated with the emissions (which have an uncertainty of ±30% on annual basis)” References are missing for the uncertainty range for emissions. I assume that the uncertainties in emissions varies substantially between different emission sectors and at least for the residential heating sector the uncertainty is likely larger than ±30%.

Answer 73: As these are all opinions and comments that we have answered previously, we do not see the necessity to answer further.

Reviewer 74: Page 5901, lines 4-5: “The emissions are all linked in the chemical composition of the atmosphere via non-linear chemical processes.” This is a good example of a sentence that can be skipped in Section 5. The mentioning of “non-linear chemical processes” etc. occurs a bit too many times in the paper.

Answer 74: It is a key point in the paper that the non-linear atmospheric chemistry cannot be disregarded when conducting emission scenarios. Therefore, we would like to keep the sentence.

Reviewer 75: Page 5901, lines 6-8: "The results in this work show that the integrated EVA model system can be used to answer relevant health-related socio-economic questions and can be used for ranking of environmental stressors by health impact" No, ranking of different emission sectors by health impact is not really possible (or at least not meaningful) before we have better ERFs for different sources.

Answer 75: We believe that we have already answered this comment, where we disagree with the reviewer. However, we do agree that much more research is needed to test the scientific hypotheses that some species (e.g. EC/BC, metals, PAH's, POPs, etc.), are more toxic than others.

Reviewer 76: Page 5913, Table 1. References for the different E-R coefficients and valuations in the table are needed. Where did all these values come from? Original scientific literature sources should be provided. If numbers are corrected/updated/modified for usage at different years or locations (USA to Europe, PM10 to PM2.5 etc) information about this must be given (such thing are very good to put in Supplementary material).

Answer 76: As previous mentioned the CRF have been reviewed in 2011 – a reference to work will be given in the paper.

Reviewer 77: The "Restricted activity days (PM)" lines are strange (what do the negative values mean and why are there four lines?):
= 8.4E-4 days/µgm−3 (adults)
−3.46E-5 days/µgm−3 (adults)
−2.47E-4 days/µgm−3 (adults>65)
−8.42E-5 days/µgm−3 (adults)

Answer 77: We have chosen to give the CFRs which are corrected by subtracting RAD due to other reasons as in the original reference. This will be explained in detail in the revised version of the paper.

Reviewer 77: "Infant mortality (PM)" but the text says SO2 rather than PM for infant mortality.

Answer 77: The infant mortality is related to PM and not SO2. We cannot find anywhere in the text where infant mortality should be associated with SO2.

Reviewer 78: Page 5915-5918, Tables 3-6. All these tables give the impression of very small uncertainties, which is completely unreasonable. Considering the discussion about uncertainty ranges above all tables need to be updated to show ranges of cases/costs instead of single numbers.

Answer 78: As this paper contains a demonstration of the methodology in the EVA system, we do not agree that uncertainty ranges should be calculated. This could be included in a future paper as the whole discussion of proper uncertainty ranges is a very large issue.

Reviewer 79: Page 5916, Table 4. Provide an explanation for PM2.5 in this table (I guess it means primary PM2.5?).

Answer 79: It is already stated in the table legend that it means the primary part of $PM_{2.5}$.

TECHNICAL CORRECTIONS

Page 5873, line 28: change “live times” to “lifetimes” (or “atmospheric residence times”).

Page 5885, line 27: “Lim et al. 2013” should be “Lim et al., 2012”. Also erroneous in the reference list.

Page 5886, line 1; “Raaschou-Nielsen et al., 2005” should be “Raaschou-Nielsen et al., 2002”.

Page 5876 line 21, page 5887 line 8, and page 5902: References: There are two different Andersen et al., 2007 references, need to be denominated 2007a and 2007b.

Page 5892 line 10: change “be discussion” to “be discussed”

Page 5905, line 5: I could not find the ExternE, 1997 reference at www.externe.info, the web address needs updating

Page 5922, Figure 4: The colour scales are much too small in the figure! They must be changed to much larger size (they should be readable on a printout of the paper).

caption, line 4: change CO {ppm] to CO [ppm]

Answer: All the suggestions above have been corrected in the manuscript.