**Interactive comment on** “Cost effective determination of vehicle emission factors using on-road measurements” **by N. Hudda et al.**

Anonymous Referee #3

Received and published: 10 September 2012

The manuscript is well written, with a good methodological approach and of interest for its immediate application for air quality management. Results on real-world emission factors and rates from separated vehicle categories are of primary interest for the air quality community. This manuscript has the additional value of evaluating the impact of environmental policy for HDV fleet. This kind of studies are urgently needed in Europe, where beside the drawback of CRT systems for NO2, the EURO standards did not reduce real-world primary NO2 emission per single light-vehicle.

I would like authors to address the following points:

The increase of NO2/NOx ratio is likely due to the CRT retrofit, but the strategy of CARB was also that of banning old vehicles. Do authors have sufficient data to distinguish between these two phenomena when evaluating their impact on single pollutants?

Table S2 should also include precipitation data and some estimate of traffic congestion found during each route.

Section 3.1. The first sentence does not seem true for CO and PB-PAH or at least the statement can not be observed in the plots. The same for line 20 of the same page

The results presented seem to suggest that NO2 can not be longer used as reliable indicator of traffic PM emissions, given the impact of retrofit systems and (in Europe) the change of NO2/NOx primary emissions. Can authors discuss if, basing on their study, BC is a more reliable air quality metric?

In the Conclusions, the words “diesel” and “gasoline –powered” should be added before HDV and LDV, since fuel separation is not so clear in rest of countries.

Fig S4 should be updated also with LDV. Fig. 5. Does 110 label refer to both stretches or only the southern? If so, please add the northern stretch as well.

I generally agree with the comments of other reviewers, mostly concerning:

The evaluation of this technology for the European case, where LDV are more separated between diesel and gasoline engines, with respect to US.

More details are needed about the background concentrations subtracted. Authors mentioned they used the first percentile, but it is not clear on what time series. Were data from the same freeway link assembled together or divided per time of the day? Background concentrations likely vary from one hour to another. Representativeness of the data used must be discussed.