Association between ambient black carbon and temperature and High-density lipoprotein

Yuhong Hu

Contents

1 Introduction

High-density lipoprotein (HDL) cholesterol is known as the 'good' cholesterol because it helps remove other forms of cholesterol from bloodstream. Higher levels of HDL cholesterol are associated with a lower risk of heart disease.

Air pollution is known to be one of the leading causes of cardiovascular disease. Emerging evidence suggests that particulate-mediated HDL dysfunction might be a novel mechanism linking air pollution exposure to adverse cardiovascular effects.However, few studies have evaluated the impact of traffic-related air pollution exposure (black carbon, nitrogen dioxide). Similarly, temperature has been linked to cardiovascular disease, but little is known about the underlying mechanisms.

Given the background, the project investigated the association between acute exposure to ambient black carbon (one source of traffic-related air pollution) and ambient temperature and HDL level. Additionally, to account for potential latency in HDL level changes, the project also assessed the association between 1-day and 2-day lag exposure and HDL level. Furthermore, the project also explored the effect modification by potential risk factors for the (lag) association of interest.

2 Method

2.1 Dataset and study population

For data disclosure issue,the project was based on part of the Veterans Administration Normative Aging Study (NAS), which could be used for class. NAS was a longitudinal study established in 1963, the study enrolled 2,280 men from the Greater Boston area, who were aged between 21 to 80 and were determined to be free of known chronic medical conditions by an initial health screening. Participants visited the study center repeatedly for physical examinations, blood pressure measurements, blood sample collection, and questionnaires approximately every four years. Blood samples were used for lipid analysis.

During the follow-up period, black carbon (BC) was measured at a central monitoring site located on the roof of Countway Library, Harvard Medical School, in Boston, MA, temperature was also collected. Single day lags were computed for these pollutants and meteorological variables from the same day of health visits and up to 2 days before the visit.

The dataset included a subset of 981 subjects with a total of 2483 observations. As NAS dataset was well-curated, so the dataset was pretty clean and less subject to missing data issue. For total of 31 variable, `RACE` had 29 missing observations, and `NEDUC` had 1 missing value. By looking at the summary statistics of each variable, we did not observe any implausible value for continuous variable or wrongly-coded value for categorical variable. Therefore, we simply excluded the observations with missing observations, as the missing data issue here is trivial. The final main dataset had 968 subjects with a total of 2453 observations from 1995 to 2011.

2.2 Data wrangling and transformation

We first create a new categorical variable `hdlcat` to divide the HDL level into 2 categories based on clinical recommendation. For easy plotting and modeling, we also level and label the categorical variable. For tabling summary statistics, we created label for key variables. In one of the result section, we would like to show the univariate association of interest by black carbon/temperature on the same day of each visit (`lag0`), black carbon/temperature on the previous 1 day of each visit (`lag1`), black carbon/temperature on the previous 2 days of each visit (`lag2`). However, the lag exposure measures were storaged in 'wide' format, thus, we transformed the original dataset (one subject per row) into 'long' dataset (one time per row, each subject may have muptiple rows) for easy plotting. The ambient black carbon and temperature exposure were the same for subjects who took hospital vist on the same day, thus, we had duplicates of the two exposure in the dataset. To plot time series of temperature and black carbon during following, we created dataset without duplicates for variables `BC24H` and `TEMPC24H`.Last but not least, the dataset produced without further usage was removed to keep environment tidy.

The packages used for EDA was skimr; The packages used for data wrangling and transformation were dplyr, tidyr; The packages used for visulization were ggplot2, ggpubr, table1, kableExtra.

3 Results

3.1 Descriptive statistics

The study included 968 subjects with a total of 2453 observations from 1995 to 2011. According to clinical cutoff, males with less than 40 mg/dL HDL was at risk of cardiovascular disease, and HDL level of 40 mg/dL or above was normal for males. Among 2453 visits, HDL levels were below this clinical cutoff for 729 visits. The Table 1 showed the basic summary statistics of key variables by the two levels of HDL. Participants' visits of normal HDL had lower `BMI`, lower daily `black carbon`, and lower daily `temperature`, compared to participants' visits with HDL below the cutoff. The proportion of `obeses` was higher in participants' visits with HDL at risk. The proportion of `diabetes` was lower in participants' visits of normal HDL. During the `cold season`, the HDL levels were more likely to be above the cutoff than during the `warm season`.

We also explore the temporal patterns of HDL levels, daily daily ambient black carbon and daily ambient temperature during follow-up (Figure 1). The scatterplots showed the seasonal pattern of the variables and the smooth indicated potential long-term temporal trends over study period. We did not observe a seasonal pattern of population HDL level, but the smooth curve suggested that the HDL level among the participants' visits increased from 1995 to 2000, remained stable from 2000 to 2005, and began fluctuated from 2005 to 2011. In general, the ambient black carbon level has decrease during follow-up, from 1995 to 2005, and remained stable afterwards.Besides, regular seasonal fluctuation was observed for ambient temperature during follow-up, shown by dots. No long-term trend of ambient temperature was observed.

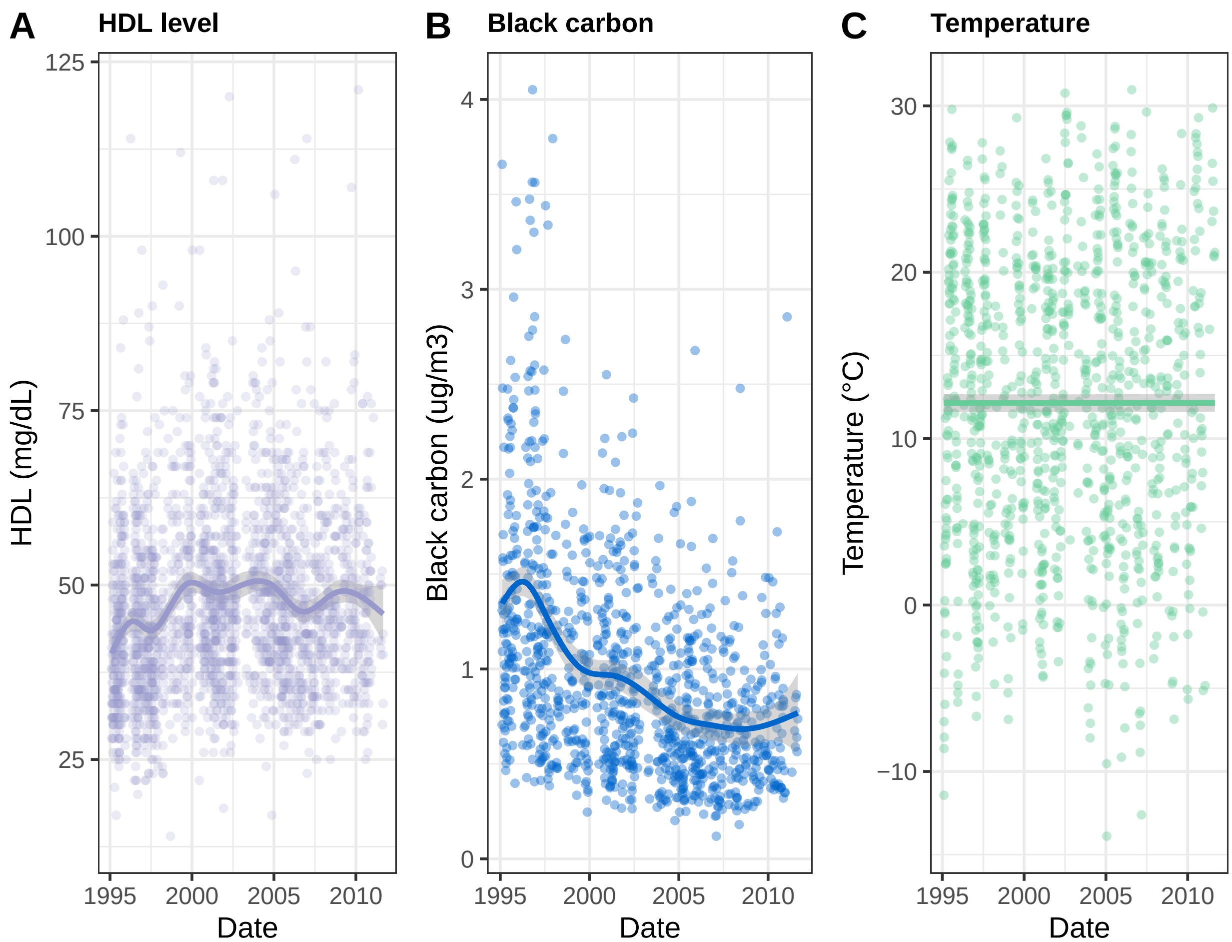

Figure 1: Temporal trends of daily HDL, daily ambient black carbon and temperature (1995-2011)

To further explore whether some demographic and physiological variables were associated with HDL levels, we displayed violin plots with boxplots (categorical variables) and scatterplots with linear regression lines (continuous variables) to show the distribution of HDL by different levels of potential risk factors, and to identify risk factors of HDL among NAS population (Figure 2). For categorical variables, the different shapes of violin were observed for different levels of Statin use, diabetes status, and race. The heterogeneity of distribution of HDL was most appreciable for different races. We observed that the distribution of continuous HDL levels was narrower for Statin users and diabetes compared to non-Statin-users and non-diabetes. However, we did not observe distinct patterns of distribution of HDL for within season, smoking status, days of the week. For continuous variables, age was slightly positively associated with HDL ($r = 0.092$, p-value < 0.05), and BMI was moderately negatively associated with HDL ($r = -0.27$, p-value < 0.05).

3.2 Univariate relationship between daily (lag) ambient black carbon, temperature and HDL

The distribution of HDL was not normal with right skewness (Figure 3).The normality was achieved by log transformation of HDL (Figure 3).Therefore, we used log(HDL) as outcome measure to assess the relationship of interest in this study.

The effect and lag effects of ambient black carbon on the HDL levels was shown in Figure 4 and Figure 5. The pattern of associations were similar at lag0 and lag1, which were different at lag2. In terms of the overall association, same day black carbon (lag0) , previous day black carbon (lag1), and previous two days black carbon (lag2) seemed to inversely and

Table 1: Summary Statistics of (repeated) measurement of characterstics by HDL level

	At risk (<40)	Normal (>=40)	Overall
	(N=729)	(N=1724)	(N=2453)
AGE (years)			
Mean (SD)	71.7 (7.64)	73.3 (7.31)	72.9 (7.44)
Median [Min, Max]	72.0 [51.0, 96.0]	73.0 [49.0, 97.0]	73.0 [49.0, 97.0]
BMI (kg/m2)			
Mean (SD)	29.3 (4.49)	27.3 (3.62)	27.9 (4.00)
Median [Min, Max]	28.5 [19.4, 52.6]	27.0 [16.7, 43.9]	27.4 [16.7, 52.6]
BMI CATEGORY			
Underweight	0 (0%)	6 (0.3%)	6 (0.2%)
Normal	456 (62.6%)	1369 (79.4%)	1825 (74.4%)
Overweight	0 (0%)	0 (0%)	0 (0%)
Obese	273 (37.4%)	349 (20.2%)	622 (25.4%)
RACE			
White	722 (99.0%)	1677 (97.3%)	2399 (97.8%)
Black	4 (0.5%)	33 (1.9%)	37 (1.5%)
Hispanic White	3 (0.4%)	10 (0.6%)	13 (0.5%)
Hispanic Black	0 (0%)	4 (0.2%)	4 (0.2%)
American Indian	0 (0%)	0 (0%)	0 (0%)
STATIN USE			
No	448 (61.5%)	1041 (60.4%)	1489 (60.7%)
Yes	281 (38.5%)	683 (39.6%)	964 (39.3%)
DIABETE			
No	578 (79.3%)	1540 (89.3%)	2118 (86.3%)
Yes	151 (20.7%)	184 (10.7%)	335 (13.7%)
SEASON			
Cold	287 (39.4%)	775 (45.0%)	1062 (43.3%)
Warm	442 (60.6%)	949 (55.0%)	1391 (56.7%)
SMOKING STATUS			
Never	209 (28.7%)	519 (30.1%)	728 (29.7%)
Current	26 (3.6%)	71 (4.1%)	97 (4.0%)
Former	494 (67.8%)	1134 (65.8%)	1628 (66.4%)
SAME DAY BLACK CARBON (ug/m3)			
Mean (SD)	1.08 (0.636)	0.956 (0.552)	0.993 (0.581)
Median [Min, Max]	0.953 [0.119, 4.05]	0.824 [0.180, 3.56]	0.854 [0.119, 4.05]
SAME DAY TEMPERATURE (°C)			
Mean (SD)	13.3 (9.14)	12.6 (8.54)	12.8 (8.73)
Median [Min, Max]	14.0 [-11.4, 31.0]	12.8 [-13.9, 31.0]	13.3 [-13.9, 31.0]

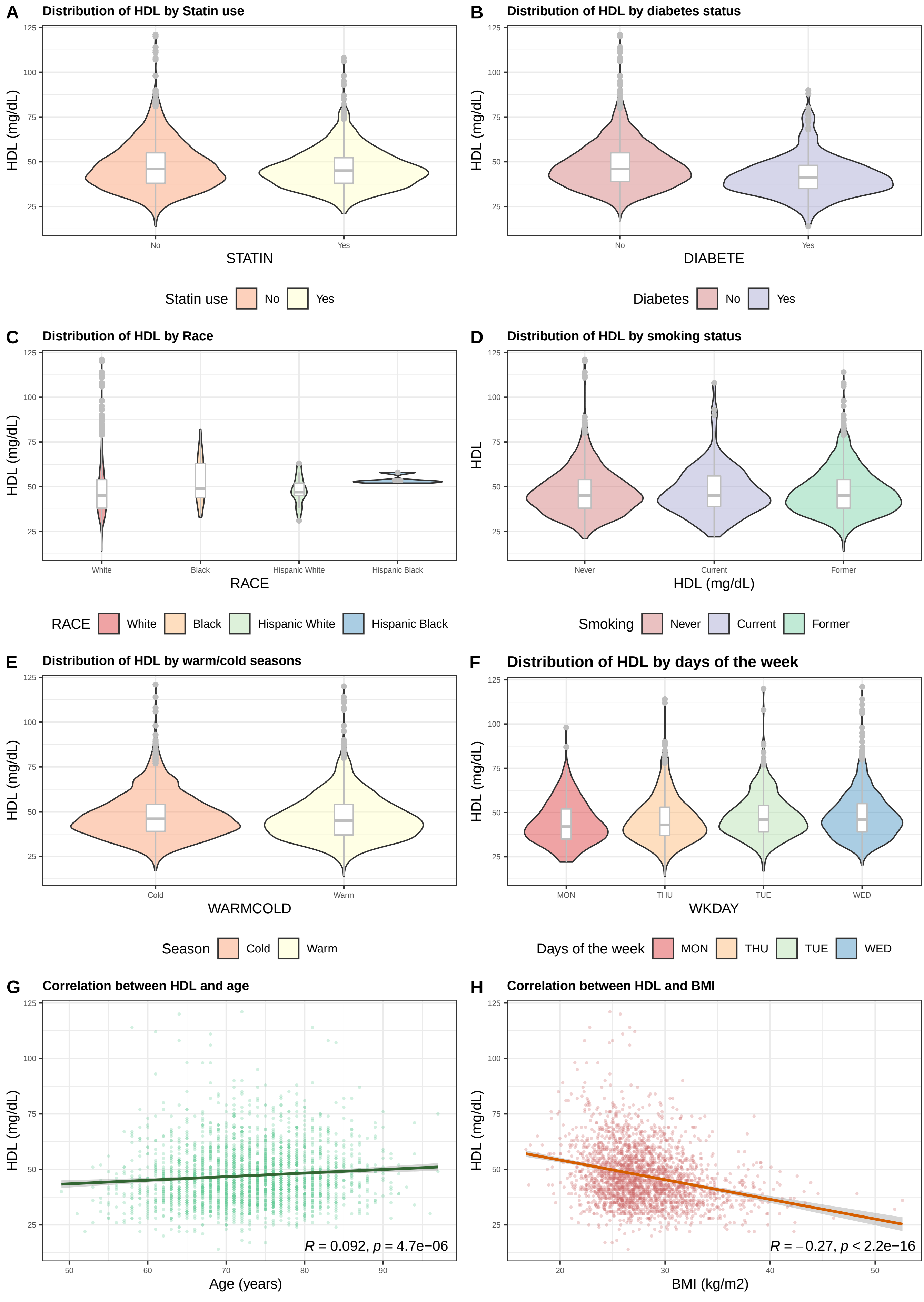

Figure 2: Distribution of HDL by different levels of risk factors

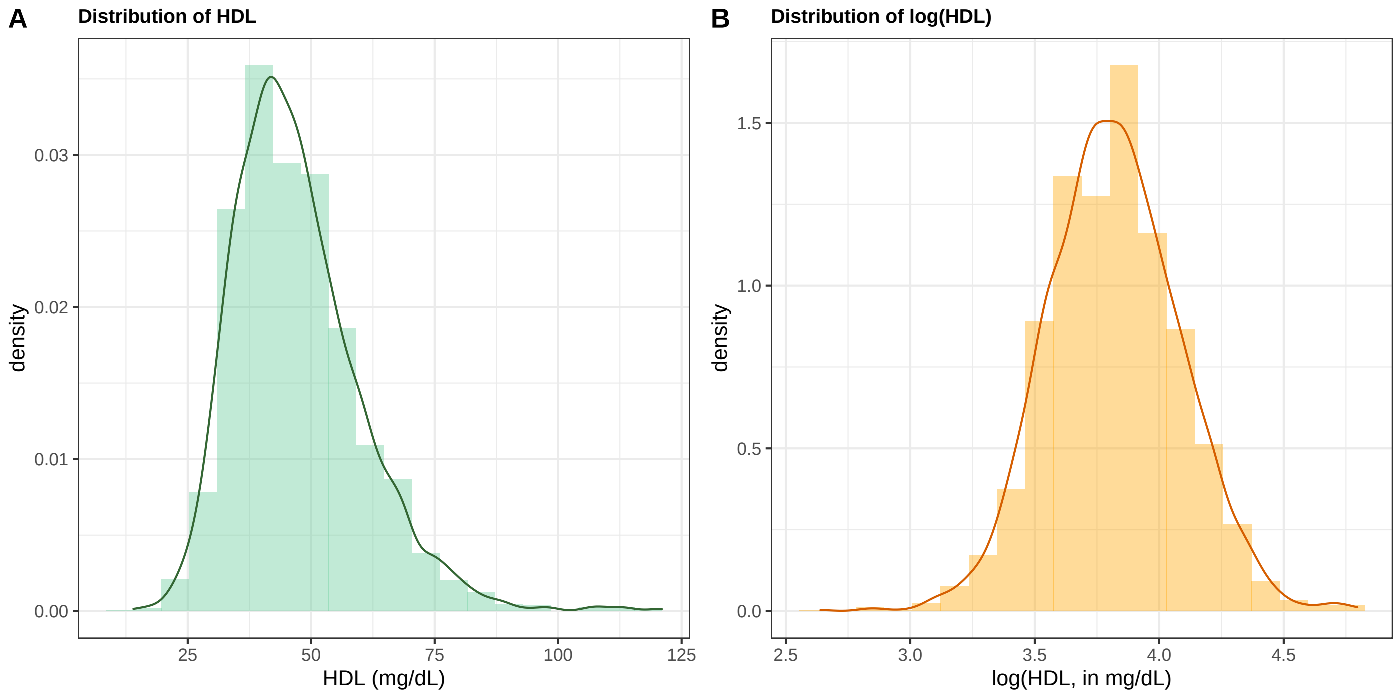

Figure 3: Distribution of HDL and log(HDL)

linearly associated with HDL level, if ignoring the wide confidence interval due to the sparse data for observations with black carbon concentrations above 2 ug/m^3. The magnitude of the overall associations were similar across the lags. Effect modification could be shown by different slopes of the subgroup association (Figure 5).Generally, in groups of diabetes, non-Statin-uses, and -relatively lower age, their subgroup associations were similar to the overall associations across lags. Effect heterogeneity of diabetes was observed for the association of interest at lag2, where null association was found for diabetes. For easier interpretation, we dichotomized the continuous age using median age (73yrs). Among people with greater age, the association of interest was stronger than people with age below median age at lag0 and lag1, whereas we found stronger negative association among people below the median age at lag2. No clear patterns of effect heterogeneity were shown for Statin use (mixed result). We did not observe effect heterogeneity by BMI categories, as was displayed by the same slopes for subgroup association within the same lag.

The effect and lag effects of temperature on the HDL levels was shown in Figure 6 and Figure 7. The pattern of associations were similar across three lags. In terms of overall association, the relationship was inverted U-shape, with average HDL increasing below around 10 °C and decreasing between temperatures 10 and 30 °C. The magnitude of the overall association was similar across the lags. Effect modification could be shown by different slopes of the subgroup association (Figure 7). Generally, in groups of diabetes, non-Statin-uses, and -relatively lower age, their subgroup associations were similar to the overall associations across lags. Effect heterogeneity of diabetes was observed, where null association was found for diabetes for all three lags assesed. For people using Statin, there was an inverse association between HDL and temperature. However, among people without Statin usage, the relationship was inverted U-shape, with average HDL increasing below around 10 °C and decreasing between temperatures 10 and 30 °C. Among people with age greater than 73 yrs, there was an inverse association between HDL and temperature. However, among people with age lower than median, the relationship was inverted U-shape, with average HDL increasing below around 10 °C and decreasing between temperatures 10 and 30 °C. No clear pattern were shown for BMI.

4 Conclusion

The relationship between same-day and lagged temperature and HDL levels was inverted U-shape, with average HDL increasing below around 10 °C and decreasing between temperatures 10 and 30 °C. In contrast, we found inverse and linear association between same-day and lagged exposure of ambient black carbon and HDL levels, this was further supported by the result of linear mixed effect model where we did detect a significant effect of same-day black carbon and its two lags on HDL (See Appendix). Furthermore, diabetes, Statin use, and age seemed to modify the association between daily temperature/black carbon and HDL levels across all the lags assessed, suggesting it is important to assess the

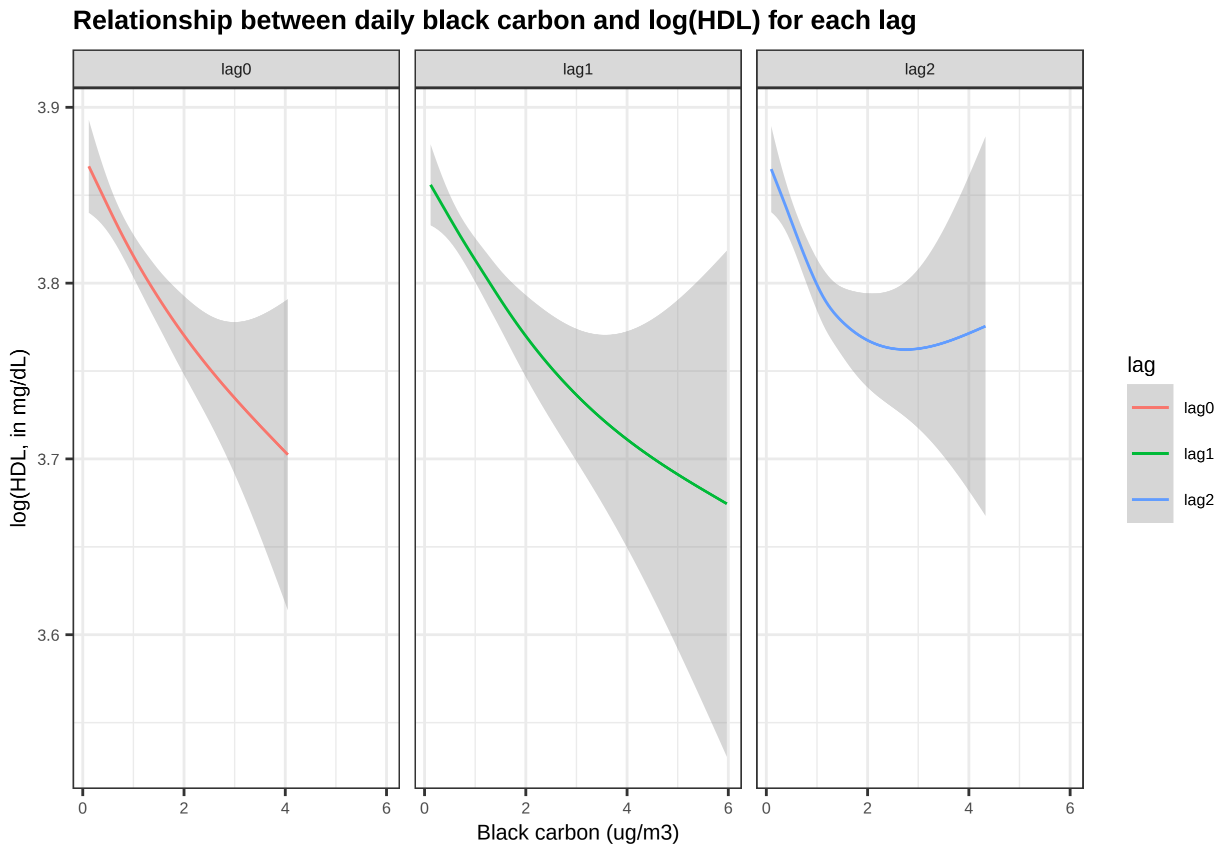

Figure 4: Overall relationship between daily ambient black carbon and log(HDL) for each lag

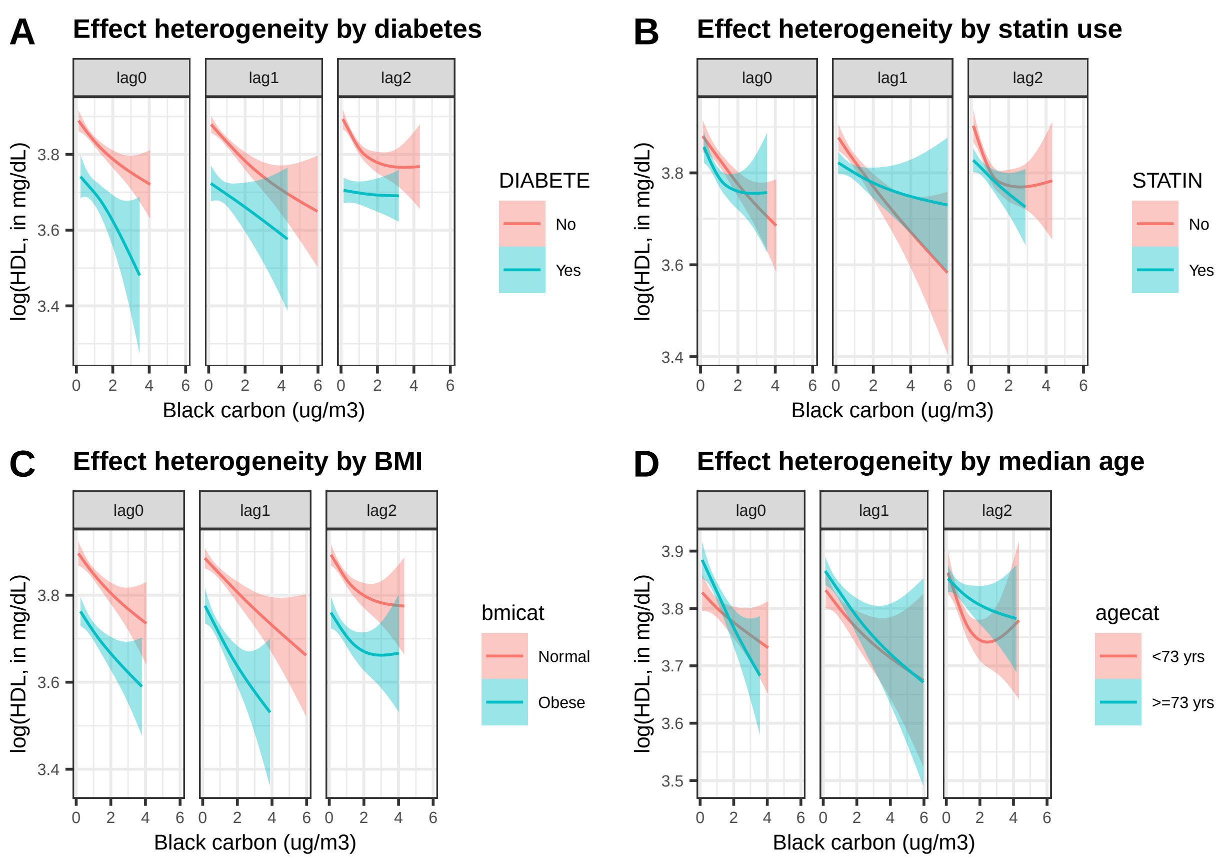

Figure 5: Relationship between daily ambient black carbon and log(HDL) for each lag by different risk factors

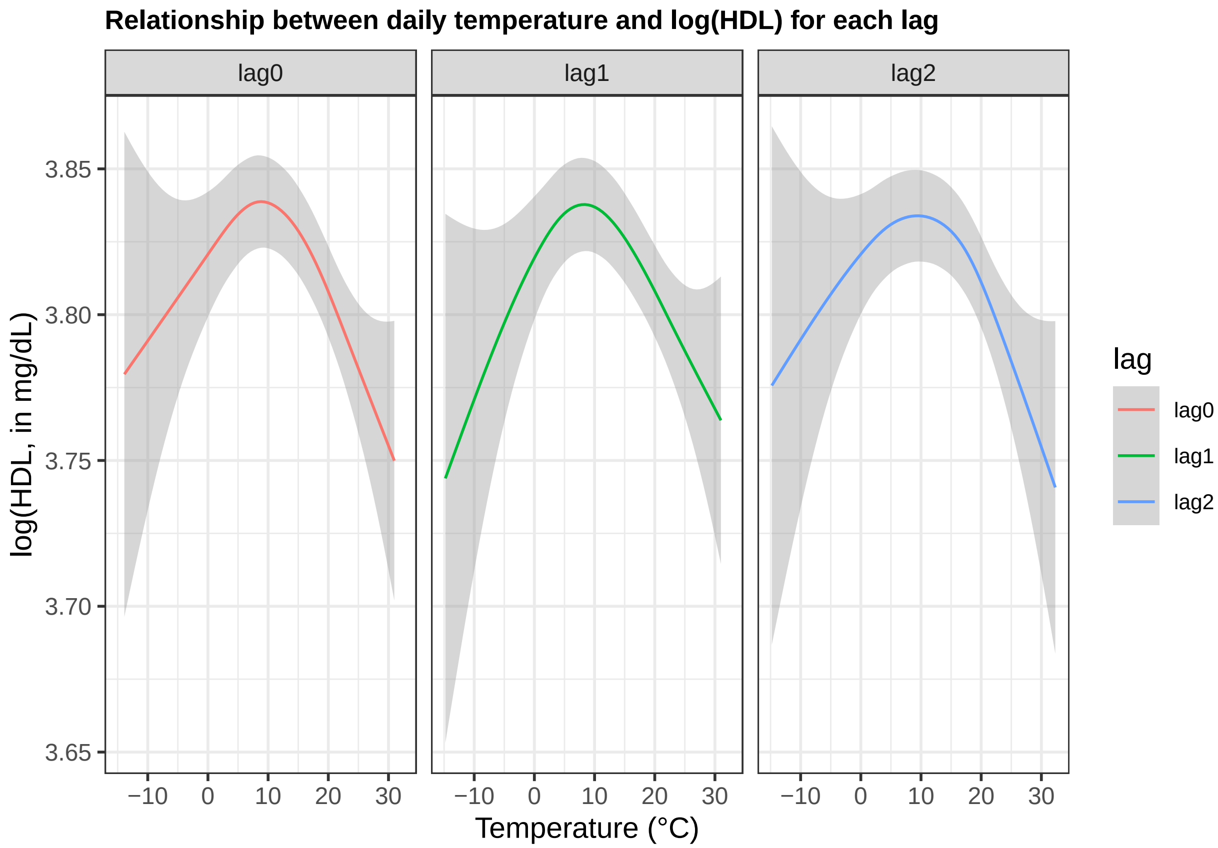

Figure 6: Overall relationship between daily temperature and log(HDL) for each lag

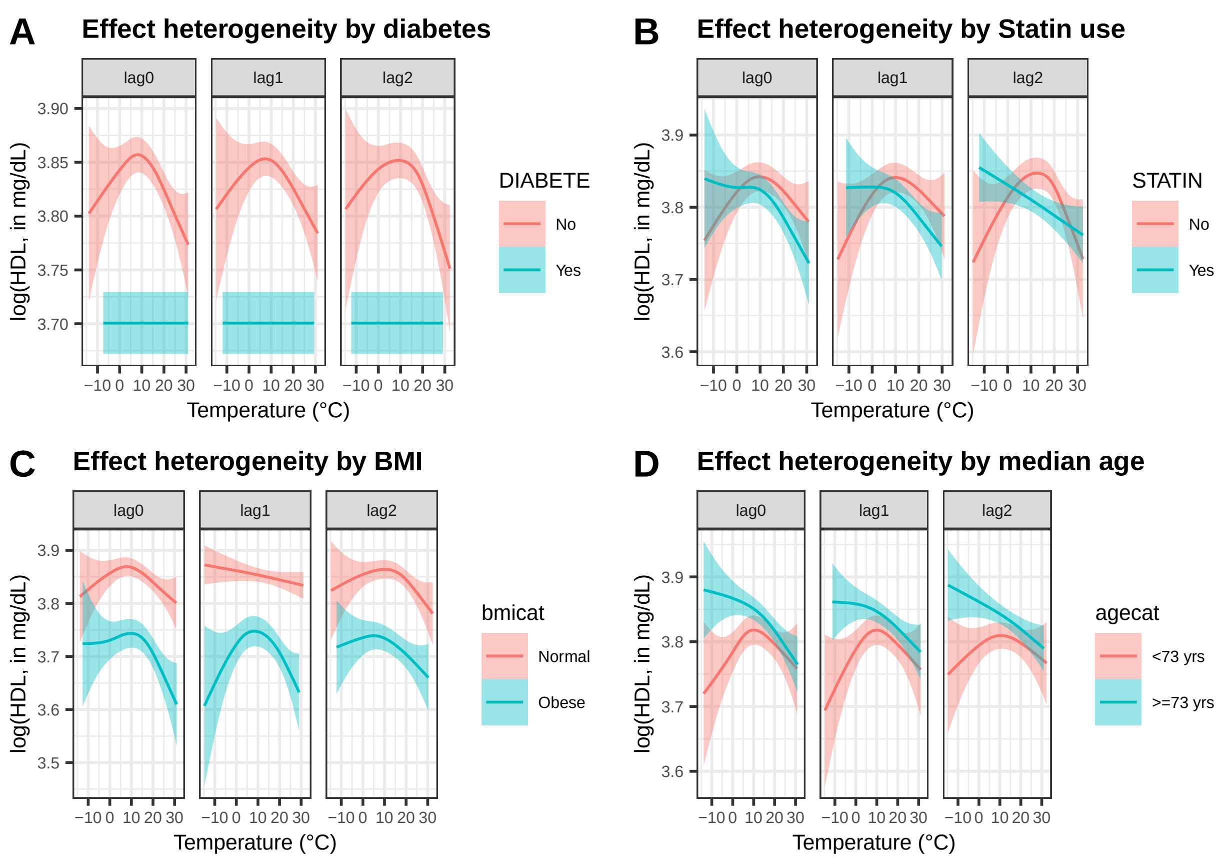

Figure 7: Relationship between daily temperature and log(HDL) for each lag by different risk factors

risk of cardiovascular disease mediated by temperature-HDL changes in susceptible populations to achieve better clinical intervention. To further understand the (lag) effect of acute exposure to black carbon and temperature on HDL-mediated cardiovascular diseases, mixed effect models and mediation analyses were needed.

5 Reference

1. Li J, Zhou C, Xu H, Brook RD, Liu S, Yi T, Wang Y, Feng B, Zhao M, Wang X, Zhao Q, Chen J, Song X, Wang T, Liu S, Zhang Y, Wu R, Gao J, Pan B, Pennathur S, Rajagopalan S, Huo Y, Zheng L, Huang W. Ambient Air Pollution Is Associated With HDL (High-Density Lipoprotein) Dysfunction in Healthy Adults. Arterioscler Thromb Vasc Biol. 2019 Mar;39(3):513-522. doi: 10.1161/ATVBAHA.118.311749. PMID: 30700134.

2. Halonen JI, Zanobetti A, Sparrow D, Vokonas PS, Schwartz J. Outdoor temperature is associated with serum HDL and LDL. Environ Res. 2011 Feb;111(2):281-7. doi: 10.1016/j.envres.2010.12.001. Epub 2010 Dec 18. PMID: 21172696; PMCID: PMC4437587.

3. Park SK, O'Neill MS, Vokonas PS, Sparrow D, Schwartz J. Effects of air pollution on heart rate variability: the VA normative aging study. Environ Health Perspect. 2005 Mar;113(3):304-9. doi: 10.1289/ehp.7447. PMID: 15743719; PMCID: PMC1253756.

4. https://www.mayoclinic.org/diseases-conditions/high-blood-cholesterol/in-depth/hdl-cholesterol/art-20046388

6 Appendix

6.1 Results based on linear mixed effect model

For further exploration of adjusted association of interest, we performed linear mixed effect model with the simplest distributed lag structure (simultaneous adjustment), to account for correlation between repeated measure and lag effects. For parsimony, we simply adjusted for covariates that we identified as major risk factors (diabetees, race, statin use, age, and BMI) in the NAS population in the previous section.

Significant effect and lag effects of black carbon level on HDL level were observed (Table 2 and Figure 8). For each lag, black carbon level was inversely associated with HDL. The decrease in HDL was more likely to be associated with the ambient black carbon level on the day of visit compared with the black carbon level on the days before the visits. Neither of the lag ambient temperature were significant predictors of the outcome (Table 3 and Figure 9). Linear effect might mask the curvature relationship, which was shown in result section. Mixed effect models with GAM was needed for future studies.

Table 2: Effect estimates of the association between black carbon and log(HDL) for each lag

	Value	Std.Error	t-value	p-value
BC24H	-0.0271123	0.0071104	-3.813071	0.0001429
BCD1	-0.0185845	0.0073608	-2.524792	0.0116806
BCD2	-0.0207110	0.0069974	-2.959813	0.0031273

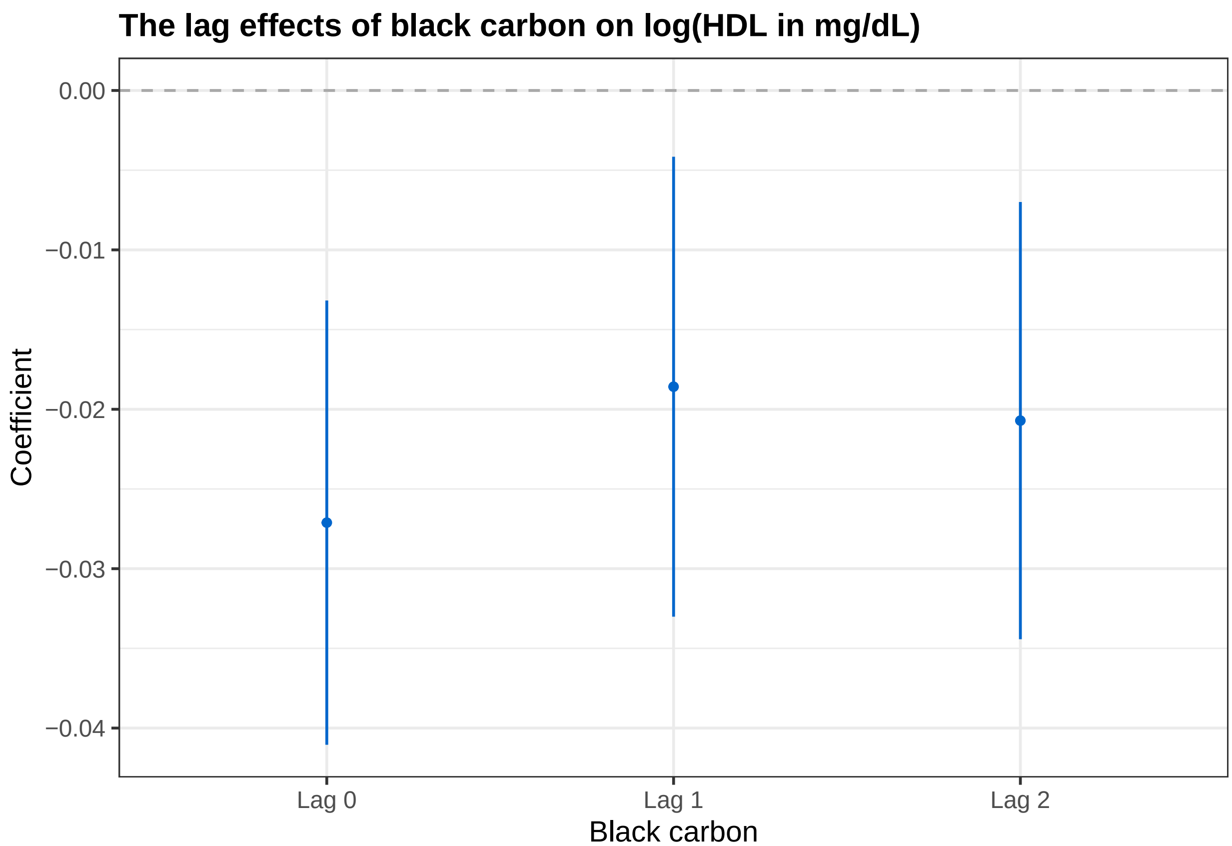

Figure 8: Lag effects for black carbon on log(HDL)

Table 3: Effect estimates of the association between temperature and log(HDL) for each lag

	Value	Std.Error	t-value	p-value
TEMPC24H	-0.0013909	0.0010084	-1.3792982	0.1680115
TMPCD1	-0.0001248	0.0013790	-0.0905062	0.9278972
TMPCD2	-0.0003733	0.0010409	-0.3586537	0.7199054

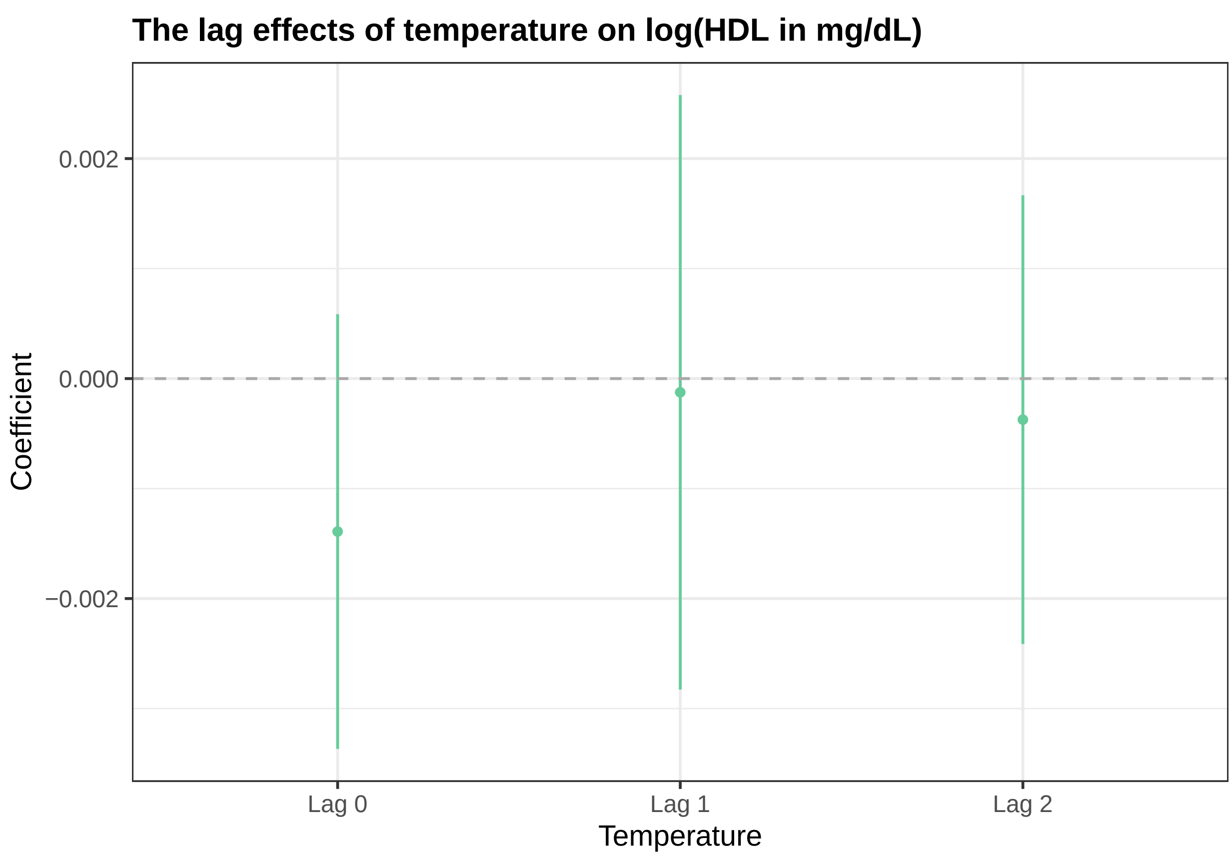

Figure 9: Lag effects for temperature on log(HDL)