Table S2. Sensitivity analyses of the associations between indoor total volatile organic compound concentrations and urinary 2-MHA and 3,4-MHA

Analysis	Metabolite	Overall	Male	Female	*p* for interaction
Full sample					
Main analysis	2-MHA	0.041 (0.029, 0.052)	0.040 (0.024, 0.055)	0.040 (0.021, 0.059)	0.990
	3,4-MHA	0.040 (0.027, 0.052)	0.037 (0.021, 0.053)	0.041 (0.020, 0.062)	0.728
Household cluster-robust SE	2-MHA	0.041 (0.017, 0.065)	0.040 (0.015, 0.064)	0.040 (0.005, 0.075)	0.993
	3,4-MHA	0.040 (0.013, 0.066)	0.037 (0.009, 0.065)	0.041 (0.005, 0.078)	0.802
Excluding air–urine interval ≥1 month	2-MHA	0.040 (0.028, 0.052)	0.039 (0.023, 0.054)	0.041 (0.021, 0.060)	0.911
	3,4-MHA	0.039 (0.027, 0.052)	0.036 (0.020, 0.052)	0.042 (0.020, 0.063)	0.671
Excluding TVOC >99th percentile	2-MHA	0.056 (0.032, 0.080)	0.053 (0.014, 0.093)	0.056 (0.026, 0.086)	0.877
	3,4-MHA	0.058 (0.032, 0.084)	0.049 (0.008, 0.090)	0.063 (0.030, 0.095)	0.531
Non-employed participants					
Main analysis	2-MHA	0.036 (0.014, 0.058)	0.110 (0.049, 0.170)	0.024 (0.001, 0.048)	0.016
	3,4-MHA	0.045 (0.022, 0.069)	0.130 (0.067, 0.193)	0.031 (0.006, 0.057)	0.008
Household cluster-robust SE	2-MHA	0.036 (−0.006, 0.077)	0.110 (0.017, 0.203)	0.024 (−0.010, 0.058)	0.106
	3,4-MHA	0.045 (0.000, 0.091)	0.130 (0.024, 0.235)	0.031 (−0.002, 0.065)	0.102
Excluding air–urine interval ≥1 month	2-MHA	0.035 (0.012, 0.058)	0.105 (0.043, 0.166)	0.024 (−0.001, 0.048)	0.024
	3,4-MHA	0.044 (0.020, 0.068)	0.125 (0.061, 0.189)	0.030 (0.004, 0.056)	0.011
Excluding TVOC >99th percentile	2-MHA	0.058 (0.023, 0.094)	0.069 (0.004, 0.134)	0.057 (0.015, 0.099)	0.868
	3,4-MHA	0.071 (0.034, 0.108)	0.080 (0.012, 0.148)	0.067 (0.021, 0.112)	0.900

Values are β (95% confidence interval [CI]) for log-transformed, creatinine-adjusted metabolite concentrations per interquartile range (IQR) increase in indoor total volatile organic compound (TVOC) (IQR, 200.0 μg/m³ in the full sample and 179.8 μg/m³ among non-employed participants, as in Figure 2 and Figure S2). All models were unweighted and adjusted for the covariates listed in the Methods; overall models were additionally adjusted for sex.

Main analysis: linear regression with model-based standard errors (as reported in the main text). Household cluster-robust SE: the same models with heteroskedasticity- and cluster-robust (CR1) standard errors, using the household as the cluster. Excluding air–urine interval ≥1 month: participants whose urine was collected 1 month or more after indoor air sampling were excluded (76 of the 1,859 participants, 75 of whom had complete data for the regression models). Excluding TVOC >99th percentile: participants with indoor TVOC >1,777.5 μg/m³ were excluded (18 of the 1,859 participants, 16 of whom had complete data for the regression models).

Regression analyses included participants with complete outcome and covariate data. Numbers of participants (overall; males/females), given as 2-MHA–3,4-MHA: full sample, main and cluster-robust analyses 1,759–1,760 (796/963–964), excluding air–urine interval ≥1 month 1,684–1,685 (763/921–922), excluding TVOC >99th percentile 1,743–1,744 (787/956–957); non-employed participants, main and cluster-robust analyses 703–704 (243/460–461), excluding air–urine interval ≥1 month 670–671 (232/438–439), excluding TVOC >99th percentile 700–701 (242/458–459). Non-employed participants were those classified in the Korea National Health and Nutrition Examination Survey occupation variable as having no current employment (including homemakers and students).

2-MHA, 2-methylhippuric acid; 3,4-MHA, 3- and 4-methylhippuric acid; SE, standard error.

p for interaction is for the sex-by-TVOC product term in the corresponding model. Among non-employed participants, heteroskedasticity-robust standard errors without clustering yielded similar results to the household cluster-robust analysis (*p* for interaction = 0.112 for 2-MHA and 0.105 for 3,4-MHA).