

VARIABILITY IN THE FRACTION OF AMBIENT FINE PARTICULATE MATTER
IN INDOOR AIR AND IMPLICATIONS FOR AIR POLLUTION EPIDEMIOLOGY

By

NATASHA HODAS

A Dissertation submitted to the
Graduate School-New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Atmospheric Science

written under the direction of

Barbara Turpin

and approved by

New Brunswick, New Jersey

May, 2014

ABSTRACT OF THE DISSERTATION

Variability in the Fraction of Ambient Fine Particulate Matter in Indoor Air and Implications for Air Pollution Epidemiology

by NATASHA HODAS

Dissertation Director: Barbara Turpin

Exposure to ambient fine particulate matter ($PM_{2.5}$) is associated with multiple negative health outcomes. Studies investigating these associations commonly use $PM_{2.5}$ concentrations measured at outdoor, central-site monitors to estimate exposure. Because people spend the majority of time indoors, however, the variable efficiency with which ambient $PM_{2.5}$ penetrates and persists indoors is a source of error in epidemiologic analyses. This error generally results in an underestimation of health effects, hampering the detection of associations between ambient $PM_{2.5}$ exposures and the risk of health outcomes. To reduce this error, practical methods to model indoor concentrations of ambient $PM_{2.5}$ are needed.

This dissertation contributes to exposure science by advancing existing models of residential exposure to ambient $PM_{2.5}$ and by improving the robustness and accessibility of these tools. First, drivers of variability in the fraction of ambient $PM_{2.5}$ found indoors (F) are identified and the potential for this variability to explain observed heterogeneity in PM-mediated health-effect estimates is explored. Next, a physically-based mass-balance model and modeling tools that account for variability in human activity patterns (e.g. time spent in various indoor and outdoor environments) are used to compute ambient $PM_{2.5}$ exposures that account for the modification of $PM_{2.5}$ with outdoor-to-

indoor transport in order to explore whether the use of these refined exposure surrogates reduces error and bias in epidemiologic analyses. Subsequently, this outdoor-to-indoor transport model is evaluated and refined using measured indoor and outdoor $\text{PM}_{2.5}$ concentrations and air exchange rates, providing a practical and robust tool for reducing exposure misclassification in epidemiologic studies. Finally, the volatility basis set is used for the first time to study shifts in the gas-particle partitioning of ambient organics with transport indoors.

This dissertation provides guidance regarding measurements and data most critically needed to facilitate the prediction of refined exposure surrogates in large epidemiological studies and, thus, informs the design of future sampling campaigns and epidemiologic studies. It enables a better accounting of ambient particle penetration into and persistence in the indoor environment and constitutes an important advancement in the efforts to reduce exposure error in epidemiologic studies and to elucidate relationships between $\text{PM}_{2.5}$ exposure and adverse health outcomes.

ACKNOWLEDGEMENTS

My graduate education and research were funded by a Department of Education Graduate Assistance in Areas of National Need (GAANN) Fellowship, Environmental Protection Agency Science To Achieve Results (STAR) Fellowship agreement number FP-917336, and an Air Pollution Education and Research Grant (APERG) from the Air and Waste Management Association. Much of this research was done in collaboration with U.S. Environmental Protection Agency Cooperative Agreement CR-83407201-0.

I would like to thank all of those who have supported me during my time as a graduate student at Rutgers University and who have made this research possible. First, I would like to thank my advisors Drs. Barbara Turpin and Qingyu Meng. Dr. Turpin has provided me with invaluable guidance and support in my research and all aspects of my professional development. Dr. Meng has provided me with important statistical skills and has been a mentor throughout my graduate career. I would also like to thank the other members of my committee, Drs. Mark Miller and Ann Marie Carlton. Their expertise in the classroom and as research mentors has provided important support for this work. While not an official member of my committee, Dr. David Rich has been a valuable mentor and this work would not have been possible without his guidance. I would also like to acknowledge the other members of the EPA Coop: Melissa M. Lunden, Kelly Thevenet-Morrison, Pamela Ohman-Strickland, Lisa Baxter, Janet Burke, and Halûk Özkaynak. In addition, I thank my lab group and the faculty and staff of the Rutgers University Environmental Science Department for helping me to navigate all aspects of my graduate education.

My family and friends have provided support throughout my life and my time here. My parents (all four of them) have instilled in me a desire to learn and have always encouraged me to continuously reach for new goals. They and my sister have also always provided me with unconditional love and support. My husband, Paul Loikith, has been a pillar of strength. He has been there throughout the many ups and downs of life and can always find a way to make me smile. I also thank Virendra Ghate, Samantha Amato, and Timothy Heckler for providing their friendship, support, and comic relief. This dissertation is dedicated to the memory of Irving Hodas, who embodied the joy of learning and proved that you are never too old to explore particle physics and other scientific fields.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iv
Table of Contents	vi
List of Tables	xiii
List of Illustrations	xiv
Chapter 1: Introduction.....	1
1.1. Motivation.....	1
1.1.1 Ambient PM _{2.5} and Negative Health Outcomes.....	1
1.1.2 Exposure Error and Bias in Air Pollution Epidemiology.....	2
1.2. Background.....	4
1.2.1 Ambient PM _{2.5}	4
1.2.2 The Physics and Chemistry of Outdoor-to-Indoor Transport.....	6
1.2.3 Quantifying F	13
1.3. Hypotheses and Dissertation Objectives	15
1.4. Dissertation Overview	16
1.5. References	21
Chapter 2: Variability in the Fraction of Ambient Fine Particulate Matter Found Indoors and Observed Heterogeneity in Health Effect Estimates	29
2.1. Abstract	29
2.2. Introduction	30
2.3. Methods	32
2.3.1 Main Analysis.....	32

2.3.2 Sensitivity Analyses.....	37
2.4. Results	39
2.4.1 Main Analysis.....	39
2.4.2 Sensitivity Analyses.....	40
2.5. Discussion	42
2.5.1 Implications for Epidemiology.....	42
2.5.2 Refined Exposure Surrogates.....	43
2.6. Conclusions	47
2.7. References	47
Chapter 3: Refined Ambient PM _{2.5} Exposure Surrogates and the Risk of Myocardial Infarction	57
3.1. Abstract.....	57
3.2. Introduction.....	58
3.3. Methods.....	61
3.3.1. Study Population and Outcome Definition	61
3.3.2. Exposure Surrogates.....	62
3.3.3. Statistical Analyses.....	66
3.4. Results.....	70
3.4.1. Modeled Exposure Tier Analyses.....	70
3.4.2. AER Effect Modification Analyses.....	71
3.5 Discussion.....	73
3.6 Conclusions.....	78
3.7 References.....	79

Chapter 4: Toward Refined Estimates of Ambient PM _{2.5} Exposure: Evaluation of a Physical Outdoor-to-Indoor Transport Model.....	88
4.1. Abstract.....	88
4.2. Introduction.....	89
4.3. Methods.....	90
4.3.1. Modeled Indoor PM _{2.5} Concentrations.....	90
4.3.2. Model Evaluation.....	91
4.3.3. Attributing Model-Measurement Differences: Human Activities.....	92
4.3.4. Attributing Model-Measurement Differences: Indoor Sources of OC.....	94
4.3.5. Attributing Model-Measurement Differences: Uncertainty in OC Size Distributions.....	95
4.3.6. Attributing Model-Measurement Differences: Phase Changes of Ambient Organics.....	95
4.4. Results and Discussion.....	96
4.4.1. Initial Model.....	96
4.4.2. Model Refinements: Accounting for Human Activities.....	98
4.4.3. Accounting for Indoor Sources of OC.....	100
4.4.4. Variability and Uncertainty in OC Size Distributions.....	100
4.4.5. Shifts in the Gas-Particle Partitioning of Ambient OC.....	101
4.4.6. Further Recommendations for Epidemiologic Studies.....	102
4.5. Conclusions.....	104
4.6. References.....	105

Chapter 5: Shifts in the gas-particle partitioning of ambient organics with transport into the indoor environment.....	115
5.1. Abstract.....	115
5.2. Introduction.....	116
5.3. Methods.....	118
5.3.1. Overview.....	118
5.3.2. RIOPA Study Measurements.....	120
5.3.3. Volatility Distributions.....	120
5.3.4. Shifts in Gas-Particle Partitioning with Outdoor-to-Indoor Transport.....	121
5.3.5. Model-Measurement Closure.....	126
5.4. Results and Discussion.....	127
5.5. Conclusions.....	133
5.6. References.....	134
Chapter 6: Summary, Future Directions, and Implications.....	148
6.1. Summary and Concurrent Studies.....	148
6.2. Future Directions.....	153
6.2.1. Epidemiologic Analyses.....	153
6.2.2. Data and Measurements.....	156
6.3. Implications and Impacts.....	158
6.3.1. Implications for Future Studies.....	158
6.3.2. Implications for Exposure Mitigation Strategies and Policies.....	159
6.3.3. Broader Impacts.....	160
6.4. References.....	161

Appendix A: Supporting Information for Chapter 2.....	164
Appendix A References.....	165
Appendix A1. Queens, NY Size Distribution Measurements.....	166
Appendix A2. Fresno, CA Size Distribution Measurements.....	167
Appendix B: Supporting Information for Chapter 3.....	168
Appendix B References.....	173
Appendix B1. Input Seasonal Air Exchange Rate Distributions for the SHEDS Model used to Generate Tier 2A Exposure Estimates.....	174
Appendix B2. Input Parameters for the SHEDS Model used to Generate Tier 2A and Tier 3Exposure Estimates.....	175
Appendix B3. Mass Median Diameters and Associated Deposition Coefficients used for Particulate Sulfate, Nitrate, Elemental Carbon, and Organic Carbon in the Tier 2B Exposure Estimates	176
Appendix B4. Correlations of Measured PM _{2.5} Species Mass Fractions Across Central-Site Monitors.....	177
Appendix B5. Variables and simple linear regression results for variables identified on the RIOPA Baseline and Activity Questionnaires as being relevant to residential air exchange rate.....	178
Appendix B6. LBNL Infiltration Model Evaluation Multiple Linear Regression Results.....	181
Appendix B7. Modeled and Measured Air Exchange Rates.....	182
Appendix B8. Summary Statistics of Each Exposure Surrogate Tier.....	183
Appendix B9. Air Exchange Rate Summary Statistics by Tertile and Season...	186

Appendix B10. PM _{2.5} Species Mass Fractions and Ambient PM _{2.5} Concentrations.....	187
Appendix B11. Study Subjects by Monitoring-Site Community.....	188
Appendix B12. Relative Odds of a Transmural Infarction Calculated at the Zip-Code Level.....	189
Appendix B13. Study Population Characteristics by AER Tertile, Cool Season.....	190
Appendix B14. Study Population Characteristics by AER Tertile, Warm Season.....	191
Appendix C: Supporting Information for Chapter 4.....	192
Appendix C References.....	195
Appendix C1. Number of RIOPA Study Homes Included in Chapter 4 Analyses.....	199
Appendix C2. Summary Statistics of RIOPA Study Measurements.....	200
Appendix C3. Particle Size Distribution Literature Review.....	201
Appendix C4. Cumulative Distributions of Indoor Sulfate.....	203
Appendix C5. Robust Regression of Measured Indoor Organic Carbon on Measured Outdoor Organic Carbon.....	204
Appendix C6. Cumulative Distributions of Indoor Organic Carbon.....	205
Appendix C7. SAS Code for Chapter 4 Calculations: Elemental Carbon.....	206
Appendix C8. SAS Code for Chapter 4 Calculations: Sulfate.....	210
Appendix C9. SAS Code for Chapter 4 Calculations: Organic Carbon.....	216
Appendix D: Supporting Information for Chapter 5.....	225

Appendix D1. Volatility Distribution Parameters.....	225
Appendix D2. Aerosol Mass Spectrometer Factor Organic Aerosol Component Mass Fractions.....	226
Appendix D3. Distributions of the Change in Ambient Organic Aerosol Concentrations Association with Shifts in Gas-Particle Partitioing with Outdoor- to-Indoor Transport	227
Appendix D4. RIOPA-Region- and Season-Specific Aggregate Volatility Distributions Generated for the Chemically-Resolved, Organic Aerosol Component Calculations.....	228
Appendix D5. SAS Code for Chapter 5 Calculations: Indoor and outdoor Gas- Particle Partitioning for RIOPA Study Homes.....	229
Appendix D6. SAS Code for Chapter 5 Calculations: Multiple Linear Regression Analyses.....	276
Appendix D7. SAS Code for Chapter 5 Calculations: Average Conditions Volatility Basis Sets.....	283
Appendix D References.....	292

LIST OF TABLES

Table 2-1. Mass median diameter (MMD) and associated k_{dep} and P_{filter} values used for particulate soil, sulfate, nitrate, elemental carbon, and organic carbon.....	53
Table 3-1. Relative increase in odds of a transmural infarction associated with each IQR increase in $PM_{2.5}$ concentration, by exposure Tier.....	84
Table 3-2. Relative increase in odds of a transmural infarction associated with each IQR increase in $PM_{2.5}$ concentration, by exposure Tier. Z-score method.....	85
Table 3-3. Relative odds of transmural infarction associated with each interquartile range increase in $PM_{2.5}$ concentration, stratified by monitoring-site, in order of increasing median air exchange rate.....	86
Table 4-1. Ambient $PM_{2.5}$ species particle diameters and associated particle deposition loss rate coefficients, penetration efficiencies, central heating and air conditioning filter penetration efficiencies, and penetration efficiencies for homes with open windows.....	109
Table 4-2. Results of multiple linear regression analysis investigating the contribution of human activities to variability in model-measurement differences for sulfate.....	110
Table 5-1. Summary statistics of measured indoor and outdoor temperatures and organic aerosol concentrations for RIOPA study homes.....	140
Table 5-2. Multiple linear regression analyses investigating drivers of variability in changes in organic aerosol concentrations due to shifts in gas-particle partitioning.....	141

LIST OF ILLUSTRATIONS

Figure 2-1. Spatially-varying PM _{2.5} composition scenarios.....	54
Figure 2-2. Most frequently observed size distributions for major PM _{2.5} species (sulfate, nitrate, organic matter).....	55
Figure 2-3. Outdoor and indoor concentrations of ambient PM _{2.5} and the fraction of outdoor PM _{2.5} that penetrates and persists indoors (<i>F</i>) for the scenarios described in Figure 2-1 and Table 2-1.....	56
Figure 3-1. Relative odds of transmural infarction associated with each interquartile range increase in Tier 1 (central-site) PM _{2.5} concentration, stratified by air exchange rate tertile.....	87
Figure 4-1. Cumulative concentration distributions: measured indoor species and indoor species of ambient origin modeled with the initial model.....	111
Figure 4-2. Indoor PM _{2.5} species concentrations modeled with the initial model and measured.....	112
Figure 4-3. Indoor PM _{2.5} species concentrations modeled with the refined model and measured.....	113
Figure 4-4. Cumulative distributions of measured and modeled indoor organic carbon.....	114
Figure 5-1. Volatility distributions for the average conditions measured outside of RIOPA homes.....	142
Figure 5-2. Change in organic aerosol mass concentrations due to changes in gas-particle partitioning with outdoor-to-indoor transport by enthalpy of vaporization assumption.....	143

Figure 5-3. Change in organic aerosol mass concentrations due to changes in gas-particle partitioning associated with only indoor-outdoor temperature differences and indoor-outdoor differences in organic aerosol loading.....	144
Figure 5-4. Change in organic aerosol mass concentrations due to changes in gas-particle partitioning with outdoor-to-indoor transport for the three geographically and climatically diverse urban regions.....	145
Figure 5-5. Change in organic aerosol mass concentrations due to changes in gas-particle partitioning with outdoor-to-indoor transport assuming that organic aerosol can be represented as a mixture of factor-analysis components.....	146
Figure 5-6. Comparison of model-measurement differences and calculated shifts in gas-particle partitioning of ambient organics with outdoor-to-indoor Transport.....	147

Chapter 1. Introduction

1.1 Motivation

1.1.1 Ambient PM_{2.5} and Negative Health Outcomes

Chronic and acute exposures to particulate air pollution are associated with multiple negative health effects including airway inflammation, aggravation of asthma, myocardial infarction (MI), pulmonary disease, and cancer.¹ Recent attention has focused on fine particulate matter (PM_{2.5}), defined as particles with aerodynamics diameters smaller than or equal to 2.5 μm , because such particles penetrate efficiently into the air exchange regions of the lung² and estimates of risk of morbidity or mortality associated with exposure to PM_{2.5} are greater than those for PM₁₀, which includes larger, coarse-mode particles.¹ Notably, fine and coarse-mode particles are derived from different sources and formation mechanisms and, as a result, they have distinctly different physiochemical chemical properties. The fine mode, for example, is formed through combustion, nucleation, and gas-to-particle conversion processes, while the coarse mode is formed through mechanical processes (e.g. the breakup of larger material by construction activity or wind).¹ Due to these differences between PM_{2.5} and PM_{10-2.5}, the factors that influence human exposure to these pollutants, the health-effects associated with these exposures, and the strategies implemented to mitigate these exposure differ.¹ This dissertation is focused on improving estimates of ambient PM_{2.5} exposure and aims to increase understanding of the health effects associated with that exposure.

Increases in exposure to ambient PM_{2.5} are associated with increased morbidity and mortality. A recent meta-analysis, for example, reported a 2.5% increase in the risk of MI with each 10 $\mu\text{g}/\text{m}^3$ increase in the ambient PM_{2.5} concentration in the 24 hours

preceding the MI.³ Each 10 $\mu\text{g}/\text{m}^3$ increase in the annual average ambient $\text{PM}_{2.5}$ concentration was associated with a 10.9 and a 20.8% increase in all-cause mortality for participants in an American Cancer Society study and the Harvard Six Cities Study, respectively.⁴ Research has also shown associations between certain health effects and $\text{PM}_{2.5}$ derived from specific sources or of specific chemical makeup. Multiple studies have reported increased risks of adverse health outcomes when ambient $\text{PM}_{2.5}$ is enriched in primary combustion tracers⁵⁻⁸ and with residential proximity to a major roadway.⁹⁻¹¹ Other studies have provided evidence for an increased risk of morbidity and mortality when $\text{PM}_{2.5}$ is enriched in chemical species formed through atmospheric chemistry (i.e. secondary $\text{PM}_{2.5}$).¹²⁻¹⁵

1.1.2 Exposure Error and Bias in Air Pollution Epidemiology

The vast majority of studies investigating relationships between ambient $\text{PM}_{2.5}$ exposure and negative health outcomes use $\text{PM}_{2.5}$ concentrations measured at outdoor, central-site (community-scale) monitors as surrogates for human exposure to ambient $\text{PM}_{2.5}$. However, this exposure is largely dependent on human activity patterns. More specifically, exposure is a function of the time spent in various microenvironments (e.g. in the home, at a restaurant, in a vehicle) and the ambient $\text{PM}_{2.5}$ concentrations in each of those microenvironments. The use of $\text{PM}_{2.5}$ concentrations measured at outdoor, stationary monitors to estimate exposure does not take human activity patterns into account. Notably, people spend the majority of their time indoors (85-90%), and most of that time in their homes.¹⁶ While ambient $\text{PM}_{2.5}$ concentrations can be higher at other microenvironments (e.g. at a bus stop), the large amount of time spent in the residence makes it an important venue for exposure to ambient $\text{PM}_{2.5}$.¹⁶

The use of central-site $\text{PM}_{2.5}$ concentrations as exposure surrogates inherently assumes that indoor and outdoor ambient $\text{PM}_{2.5}$ concentrations and compositions are well correlated. As is discussed in detail below, however, the fraction of ambient $\text{PM}_{2.5}$ that penetrates and persists indoors (F) varies within and across homes^{17,18} and is different for different chemical components of the $\text{PM}_{2.5}$ mixture.¹⁹⁻²² This variability has been identified as a source of exposure error in epidemiologic studies that use central-site $\text{PM}_{2.5}$ concentrations as exposure surrogates. This error is likely to bias health effect estimates towards the null (i.e. result in an underestimation of health effect),^{23,24} hampering the detection of statistically significant associations between increased ambient $\text{PM}_{2.5}$ exposures and the risk of negative health outcomes. In order to reduce the exposure error and bias associated with variability in F , practical methods to predict indoor concentrations of ambient $\text{PM}_{2.5}$ in large epidemiologic studies are needed. The research presented in this dissertation is designed to address this need.

There is evidence that supports the hypothesis that the variability in F contributes to observed heterogeneity in health-effect estimates. Multiple studies have reported a lower risk of morbidity or mortality associated with ambient $\text{PM}_{2.5}$ in communities with a higher prevalence of central air conditioning (AC), compared to risk estimates among communities with lower AC prevalence.²⁵⁻²⁸ Notably, homes with central AC in use tend to have lower F values because indoor air is filtered as it is re-circulated, thus increasing particle losses indoors.²⁹⁻³¹ When F values are lower, the difference between central-site $\text{PM}_{2.5}$ concentrations and actual ambient $\text{PM}_{2.5}$ exposure is greater, resulting in proportionally more exposure misclassification and larger bias towards the null (i.e. a greater underestimation of effect), likely contributing to the lower effect estimates

observed for homes with AC use. Work conducted as part of this dissertation and concurrent work conducted by Chen C. et al.^{32,33} and Chen R. et al.³⁴ provide more comprehensive estimates of F and use these to explore the hypothesis that variability in factors that influence F contributes to exposure error and bias in epidemiologic studies.

A better accounting of ambient particle penetration and persistence in the indoor environment is needed to reduce exposure error in epidemiologic studies and to further elucidate relationships between $PM_{2.5}$ exposure and adverse health outcomes. Toward this goal, this dissertation advances existing methods to predict indoor concentrations of ambient $PM_{2.5}$, increases the robustness and accessibility of these modeling tools, and explores the ways in which accounting for (and not accounting for) variability in F influences health-effect estimates derived from epidemiologic studies. This research also provides insights regarding the measurements and data most critical to the prediction of residential ambient $PM_{2.5}$ exposures and, thus, will inform the design of future sampling campaigns and epidemiologic studies.

1.2 Background

1.2.1 Ambient $PM_{2.5}$

Ambient $PM_{2.5}$ is a complex mixture of locally- and regionally-generated pollutants. Regionally-generated $PM_{2.5}$ is formed through atmospheric processing of gas-phase pollutants during transport. For example, sulfur dioxide emitted from coal-fired power plants in the Ohio River Valley is oxidized during transport, resulting in ubiquitous and abundant sulfate aerosol in the eastern United States. The majority of this sulfate is formed through aqueous-phase oxidation reactions in cloud droplets, while a lesser amount can be attributed to the homogenous gas-phase oxidation of sulfur

dioxide.^{35,36} Particle-phase nitrate is formed in the atmosphere through gas-particle partitioning following the uptake of ammonia by nitric acid. Like sulfate, secondary organic aerosol (SOA) is formed through both gas- and aqueous-phase oxidation processes (SOA_{gas} and SOA_{aq}).³⁶⁻³⁸ Both SOA_{gas} and SOA_{aq} formation processes begin with the gas-phase oxidation of volatile organic compounds emitted from anthropogenic and biogenic sources. SOA_{gas} is formed when these oxidation reactions result in products with low enough vapor pressures that they partition to existing organic matter.^{36,39} In the case of SOA_{aq} , smaller, more water-soluble gas-phase oxidation products partition into cloud and fog droplets or the liquid water associated with atmospheric aerosols. They then undergo further oxidation in the aqueous phase, resulting in lower volatility products that remain in the particle phase following the evaporation of water from the droplet or wet particle.^{37,38} Locally-generated $\text{PM}_{2.5}$ is dominated by primary emissions from local sources. Concentrations of regionally-generated, secondary $\text{PM}_{2.5}$ tend to be more uniformly distributed across an air shed, while primary $\text{PM}_{2.5}$ concentrations are enhanced in close proximity to sources (i.e. on urban and neighborhood scales)¹ and diluted with distance downwind.

The size distribution and chemical composition of ambient $\text{PM}_{2.5}$ are functions of PM sources and formation mechanisms. As a result, $\text{PM}_{2.5}$ concentrations and characteristics vary temporally and spatially with source mix, as well as with meteorological factors such as precipitation, boundary layer mixing height, wind direction and actinic flux.^{1,39} As is discussed in detail below, particle size and the volatility of particle-phase components are important determinants of F .^{20,21,40,41} Modeling tools that account for variability in F with particle size and composition are

crucial for improving ambient $PM_{2.5}$ exposure estimates. That research need is addressed in this dissertation

1.2.2 The Physics and Chemistry of Outdoor-to-Indoor Transport

The fraction of ambient $PM_{2.5}$ that penetrates into and persist in indoor air (F) is a function of residential air exchange rate (AER), the efficiency with which ambient particles penetrate across the building envelope (P), the rate of indoor particle losses due to deposition in indoor air (k_{dep}), and, for semi-volatile species, losses or gains in mass resulting from phase changes in indoor air.^{19,20,40-42} In the following paragraphs, the current understanding and outstanding research needs for each of these determinants of F are discussed.

Air Exchange Rate – The rate at which air in a home is exchanged with outdoor air is the sum of two processes: (1) leakage through cracks in the building shell driven by indoor-outdoor pressure differences and (2) air flow through open doors/windows.^{43,44} Air exchange rates are commonly measured using a tracer gas method, in which a non-toxic, inert gas (e.g. sulfur hexafluoride, a perfluorocarbon tracer) is released at a known emission rate and is collected with a passive sampler. Because the emission rate of the tracer gas is known, the measured concentration in the home serves as an indicator of the rate at which air in the home is exchanged with outdoor air.^{43,45,46} Notably, only a small fraction of AER studies have included a large number of homes. Pandian et al.^{47,48} aggregated AER measurements from approximately 100 studies and generated summary statistics and frequency distributions of AERs for this subset of U.S. homes. Yamamoto et al.⁴⁹ conducted 593 AER measurements in about 100 homes in each of three climatically and geographically diverse regions of the United States (Houston, TX, Los

Angeles County, CA, and Elizabeth, NJ) as part of the Relationships of Indoor, Outdoor, and Personal Air (RIOPA) study. These studies have demonstrated that AERs vary spatially with housing characteristics (e.g. construction material, home age) and both spatially and temporally with meteorological conditions and home ventilation conditions (e.g. wind speed, indoor-outdoor temperature differences, opening/closing of windows, use of heating or cooling systems). This heterogeneity in AER contributes to geographic and temporal variability in F . Notably, the homes for which AER measurements have been conducted are limited in their spatial and temporal extent and are not statistically representative of the complete United States Housing stock. Generalizing these measurements to the broader population of U.S. homes is further complicated by the fact that, in addition to home construction characteristics, AER varies with the meteorological conditions and the human activities related to home ventilation discussed above.

Due to their practical applications for the weatherization and energy-efficiency auditing industries, measurements of the effective leakage area of homes (ELA; the sum of all cracks and spaces through which air can flow)⁴⁴ are more common than AER measurements. However, like AER, ELA measurements have not been conducted for a representative sample of U.S. homes.⁴⁴ Chan et al.⁴⁴ utilized a database of about 70,000 air leakage measurements to explore the relationship between ELA and common housing characteristics with the objective of developing a statistical model that could be used to calculate ELA for the broader population of homes. Calculations of ELA require data readily available from the American Housing Survey and the U.S. Census.⁴⁴ Chan et al.⁴⁴ demonstrated that home age and floor area are the most significant predictors of ELA and that model-measurement agreement was maximized when separate models were used

depending on household poverty status because home leakiness varies with household income (with lower income homes tending to be leakier).

Modeling tools have been developed to predict AER from ELA and such tools provide a means to calculate AER for a broader sample of homes than those for which AER measurements are available. The Lawrence Berkeley National Laboratory (LBNL) Infiltration model, for example, predicts AER distributions for single-family, closed homes (i.e. windows and doors closed) based on normalized leakage area (ELA normalized by floor area), certain house characteristics, and meteorological conditions.^{50,51} The data required to generate AER distributions with the LBNL Infiltration model are readily available from the U.S. Census (www.census.gov), the American Housing Survey (www.census.gov/housing/ahs), and the National Climate Data Center (www.ncdc.noaa.gov). The LBNL Infiltration model is an important tool for generating AER distributions for the purpose of modeling indoor concentrations of ambient PM_{2.5}. This dissertation advances the LBNL Infiltration model by exploring the extent to which certain human activities and housing characteristics contribute to model-measurement disagreement and by elucidating refinements to the LBNL Infiltration model. These analyses informed the development of a refined version of the LBNL Infiltration model, by Dr. Melissa M. Lunden, which was applied in Chapter 3 of this dissertation to calculate refined estimates of exposure to ambient PM_{2.5} in an epidemiologic study investigating the relationship between ambient PM_{2.5} exposure and myocardial infarction.

Particle Penetration Efficiency – Particle penetration efficiency (P) describes the fraction of outdoor-generated particles that remain suspended following their infiltration

through the building envelope.^{41,43} Particle losses to the sides of building cracks during infiltration are the result of Brownian diffusion, gravitational settling, interception and impaction.⁵² Brownian diffusion dominates particle losses for particles with diameters smaller than about 0.1 μm .³⁶ For larger particles, interception, impaction and gravitational settling are the dominant loss processes.³⁶ Penetration efficiency describes the fraction of particles that are not removed by the combination of these mechanisms during passage across the building shell. Sampling and laboratory studies have measured P for particles of various sizes.⁵²⁻⁵⁴ Such studies have demonstrated that the penetration efficiency curve is relatively constant across the accumulation mode, but increases rapidly for larger (super-micron) and smaller (ultrafine) particles due to size-dependence of the particle loss processes described above.^{40,41} Values of P for particles in the size range considered in this work have been observed to range from 0.6 to 1.0.⁵⁵

Values of P , and thus F , vary with building ventilation conditions and the characteristics of cracks in the building shell.^{52,55} While it is known from laboratory studies that complex crack geometry and increased roughness of crack surfaces result in lower P values, the characteristics of cracks in buildings are not well known and are likely to be highly varied within and across buildings.^{52,55} Furthermore, human activities contribute to variability in P . For example, P values approach 1.0 for homes with open windows. As a result, there is substantial uncertainty in P . While this dissertation does not reduce this uncertainty, sensitivity analyses were conducted to evaluate the effect of this uncertainty on modeled F values and adjustments were made for the effect of certain human activities on P (e.g., opening windows) in the model of outdoor-to-indoor transport applied herein.

Depositional Losses - The same physical loss mechanisms that govern penetration are also important drivers of particle deposition in the indoor environment.^{17,40,42,54} Deposition loss rates (k_{dep}) are a strong function of particle size. Depositional losses in residences have been determined empirically from simultaneous indoor and outdoor PM_{2.5} concentration measurements.^{22,53,54,56} For examples, Ozkaynak et al.⁵³ measured indoor and outdoor PM_{2.5} concentrations at 60 residences in Riverside, CA. Values of k_{dep} and P were then calculated with non-linear least squares regression. A major limitation of the majority of studies aimed at measuring P and k_{dep} is that because both P and k_{dep} are unknown, the individual effects of each of these variables on indoor ambient PM_{2.5} concentrations cannot be separated. In other words, a single, unique value for P or k_{dep} can only be determined if one of the variables is known.^{54,57} Thatcher et al.⁵⁴ took specific measures in their experimental methods to separate the effects of P and k_{dep} . Simultaneous indoor and outdoor PM concentrations were measured at an unoccupied home and values of k_{dep} for sulfate, nitrate, and organic and elemental carbon were solved for using a transient solution to the mass balance equation. Particle concentration measurements were performed under two very distinct conditions in order to separate the effects of P and k_{dep} . First, particle levels were artificially elevated through resuspension activities. Because indoor concentrations are forced to relatively high values through this process, changes in indoor concentrations are driven by depositional losses and increases in indoor particle concentrations due to penetration from outside is negligible. Under the second scenario, which was implemented to reduce the influence of depositional losses, recovery of particle concentrations was measured after indoor particles concentrations were reduced to near zero through pressurization of the home with HEPA filtered air.

Riley et al.⁴⁰ combined a physical particle deposition model with a polynomial fit to empirically determined k_{dep} values from studies like those described above to provide k_{dep} values for a wide range of particle sizes.

These studies verify that depositional losses depend strongly on particle size. However, it is not well understood how depositional losses vary geographically and temporally with heterogeneity in PM_{2.5} sources, formation mechanisms, and chemical composition (all of which are related to particle size). This dissertation explores the extent to which heterogeneity in these factors contributes to variability in k_{dep} and, thus, to variability in F . As part of this dissertation, a comprehensive assessment of PM_{2.5} species size distributions is conducted in order to provide robust estimates of particle-size dependent model inputs such as k_{dep} .

Gas-Particle Partitioning - In addition to the physical losses discussed above, shifts in the gas-particle partitioning of semi-volatile ambient PM_{2.5} species can occur with outdoor-to-indoor transport due to indoor-outdoor differences in temperature and environmental characteristics such as the availability of surface area or particulate organic matter for sorption.^{20,58,59} Recent studies have demonstrated that such phase changes can have a substantial impact on F .^{19,20,22,42} Sarnat et al.²² estimated F as the indoor-outdoor ratios of PM_{2.5} concentrations measured inside and outside 17 Los Angeles area homes during periods when indoor sources were absent. Median F for black carbon (BC), which is non-volatile, was 0.84. The median F for ammonium nitrate, which is semi-volatile, was 0.18. When the authors compared the results from Los Angeles to other study areas, they found that F for PM_{2.5} mass increased as the mass fraction of outdoor ammonium nitrate decreased. In other words, a greater fraction of

outdoor-generated PM_{2.5} penetrates into and persists in indoor air in regions where concentrations of semi-volatile nitrate are lower. Lunden et al.²⁰ measured ammonium nitrate PM_{2.5} inside and outside of an unoccupied home in Clovis, CA. The dramatic differences between indoor and outdoor particulate nitrate concentrations could not be explained by only losses with penetration through the building envelope and physical deposition indoors. When a mass balance model that included penetration efficiency, physical deposition, *and* evaporative losses was applied, however, modeled and measured indoor ammonium nitrate showed good agreement.²⁰

Shifts in gas-particle partitioning with outdoor-to-indoor transport have also been observed for semi-volatile organic compounds (SVOCs).^{21,22,58,59} Lunden et al.²¹ observed a lower *F* value for organic carbon (OC; *F*=0.47) than for elemental carbon (EC; *F*=0.61) in an *unoccupied* home. The greater OC losses were attributed to the depletion of gas-phase organics as they sorbed to indoor sources followed by shifts in the partitioning of SVOCs from the particle phase towards the gas phase in order to reach a new equilibrium. In contrast, it was concluded that, in *occupied* homes, incoming organics from outdoors can shift from the gas phase toward the particle phase as they sorb to particulate organic matter emitted by indoor sources.⁵⁸⁻⁶⁰

The gas-particle partitioning of SVOCs in the *atmosphere* has been parameterized with a partitioning coefficient, K_p ⁶¹ and, as an extension of this concept, mapped into a volatility basis set (VBS).⁶² K_p describes the ratio of mass of a specific SVOC in the particle phase (normalized by the total particle mass) to the mass in the gas phase. In the rare case that the particle composition and the properties of a partitioning organic compound are known, the gas-particle partitioning of that compound can be predicted

from the temperature, the compound's subcooled liquid vapor pressure, and the concentration and properties of the particulate OC.⁶¹ However, because the atmosphere contains thousands of organics, atmospheric models rely on parameterizations of this process. The VBS treats ambient organics as a distribution of compounds binned by their volatilities and is used to describe the gas-particle partitioning of organics in the atmosphere as a function of temperature and organic aerosol loading.⁶²

Organics account for 20 to 90% of ambient PM_{2.5} mass.⁶³ It is accepted that phase changes of organics impact F . However, changes in the gas-particle partitioning of total ambient PM_{2.5} organic matter with outdoor-to-indoor transport have not been quantitatively predicted. This dissertation expands current knowledge of ambient organic PM_{2.5} exposure in part by investigating the extent to which variability in indoor ambient OC concentrations can be attributed to shifts gas-particle partitioning and by exploring the VBS as a tool for modeling the partitioning behavior of ambient organics with outdoor-to-indoor transport.

1.2.3 Quantifying F

Multiple methods have been used to quantify the fraction of ambient PM_{2.5} that penetrates and persists indoors. In homes for which indoor sources of ambient PM_{2.5} are not present, F is commonly estimated as the ratio of ambient PM_{2.5} concentrations measured indoors to those measured outdoors.⁵⁵ Sarnat et al.,²² for example, utilized indoor and outdoor PM_{2.5} concentration measurements from overnight periods (i.e. times when home occupants were sleeping and indoor PM_{2.5} sources were minimized) to calculate F values for ambient PM_{2.5} and PM_{2.5} species. Note that this method provides an upper bound for F and could be biased high for species that are emitted or formed

indoors (e.g. EC and OC). Statistical methods, in which measured indoor $\text{PM}_{2.5}$ concentrations are regressed on measured outdoor concentrations have been used for homes or time periods for which indoor sources were present to estimate the fraction of indoor $\text{PM}_{2.5}$ that could be attributed to outdoor sources.^{18,59,64} The intercept of the resulting regression model indicates an average indoor source strength, while the slope of the regression model is an average value of F .^{18,64} A regression method that down-weights outliers (e.g. robust regression) has the advantage of reducing the influence of indoor sources of ambient $\text{PM}_{2.5}$ when estimating F because such outliers are likely indicative of a strong indoor source.¹⁸ These statistical methods assume that the indoor environment is well mixed and that indoor and outdoor sources are independent.¹⁸

Mass balance models can also be used to calculate the indoor concentration of ambient $\text{PM}_{2.5}$.^{55,65,66} Alzona et al.⁶⁵ and Koutrakis et al.⁶⁶ demonstrated that a simple, single-compartment mass balance model could be used to predict indoor concentrations and residential exposures to ambient $\text{PM}_{2.5}$. Hering et al.⁴² expanded on this work by pairing modeled AERs with central site $\text{PM}_{2.5}$ concentration data to estimate indoor concentrations of ambient $\text{PM}_{2.5}$ for an unoccupied home. Indoor concentrations were calculated with a forward-stepping solution to a single compartment mass balance equation. Evaporative losses of ammonium nitrate were modeled based on the results of Lunden et al.²⁰ discussed above. Because species-resolved size distribution data were not available, a constant k_{dep} value of 0.2 h^{-1} was assumed for all species. However, as noted above, k_{dep} is highly particle-size dependent and is likely to vary with chemical composition. Region- and species-specific characterization of $\text{PM}_{2.5}$ particle size constitutes a major gap in current knowledge and is needed to accurately account for

ambient particle penetration and persistence in the indoor environment. Species- and size-resolved characterization of $PM_{2.5}$ in different regions of the United States will allow for more accurate estimates of penetration and loss rate coefficients and, subsequently, the indoor concentration of outdoor-generated particles; these are crucial inputs for next-generation $PM_{2.5}$ exposure and health studies. As noted above, this dissertation provides estimates of such model inputs. The Hering et al.⁴² model laid the framework for estimating indoor concentrations of outdoor-generated $PM_{2.5}$ by making use of PM species data. The research presented in this dissertation expands on and refines this model by addressing the research needs discussed above.

1.3 Hypotheses and Dissertation Objectives

A central theme of this dissertation is to develop and implement a simple, practical method that can be applied in large epidemiologic studies to predict residential exposures to ambient $PM_{2.5}$. It is hypothesized that exposure surrogates that account for the effects of outdoor-to-indoor transport will provide more accurate estimates of ambient $PM_{2.5}$ exposures than central-site $PM_{2.5}$ concentrations. Thus, accounting for variability in F in ambient $PM_{2.5}$ exposure estimates will reduce exposure misclassification and will result in less bias in health effect estimates and smaller confidence intervals in epidemiologic studies. For this reason, this research described in this dissertation:

1. Identifies major drivers of spatial and temporal variability in the fraction of ambient $PM_{2.5}$ found in indoor air and evaluates whether observed heterogeneity in health-effect estimates can be attributed, in part, to variability in these factors.
2. Compares health-effect estimates derived from epidemiologic analyses in which variability in F is and is not accounted for in estimates of ambient $PM_{2.5}$ exposure and

discusses circumstances and epidemiologic study designs for which variability in F is more and less likely to be a major contributor to exposure error.

3. Identifies data and measurements most critical to the prediction of residential exposures to ambient $\text{PM}_{2.5}$ and, thus, informs the design of future sampling and epidemiologic studies.

4. Explores the magnitude and direction of shifts in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport.

1.4 Dissertation Overview

This research described herein contributes to the field of exposure science by advancing existing models of residential exposure to ambient $\text{PM}_{2.5}$ and by improving the robustness and accessibility of these tools. This work begins by examining drivers of geographic and spatial variability in F and identifying the potential for this variability to explain, at least in part, the heterogeneity in PM-mediated health effect estimates (Chapter 2). Next, (1) a physically-based mass balance model and (2) modeling tools that account for variability in human activity patterns (e.g. time spend in various indoor and outdoor environments) are used to compute ambient $\text{PM}_{2.5}$ exposures that account for the modification of $\text{PM}_{2.5}$ with outdoor-to-indoor transport. These refined exposure surrogates are used to calculate the risk of myocardial infarction associated with ambient $\text{PM}_{2.5}$ exposure and to explore whether the use of these refined exposure surrogates reduces exposure error and bias in epidemiologic analyses (Chapter 3). Subsequently, in Chapter 4, this outdoor-to-indoor transport model is validated and refined using data from the RIOPA study, providing a practical and robust tool for reducing exposure misclassification in epidemiologic studies. The work of Chapter 4 identified shifts in gas-

particle partitioning as a remaining source of exposure error and thus led to the work of Chapter 5. For the first time, the volatility basis set is applied to study shifts in the gas-particle partitioning of ambient organics with transport into the indoor environment (Chapter 5). Finally, the research contained in this dissertation provides guidance regarding measurements and data most critically needed to facilitate the prediction of refined exposure surrogates in large epidemiological studies and, thus, contributes to the design of future sampling campaigns and epidemiologic studies. This work enables a better accounting of ambient particle penetration into and persistence in the indoor environment and constitutes an important advancement in the efforts to reduce exposure error in epidemiologic studies and to elucidate relationships between $PM_{2.5}$ exposure and adverse health outcomes.

In Chapter 2, F was modeled using a mass balance approach for several scenarios across which heterogeneity in effect estimates has been observed: with geographic location of residence, residential roadway proximity, socioeconomic status, and central air conditioning use. Calculated values of F are higher in close proximity to primary combustion sources (e.g. proximity to traffic) and for lower income homes. F is lower when $PM_{2.5}$ is enriched in nitrate and with central air conditioning use. As a result, exposure error resulting from variability in F will be greatest when these factors have high temporal and/or spatial variability. The circumstances for which F is lower in these calculations correspond to circumstances for which lower effect estimates have been observed in epidemiological studies and higher F values correspond to higher effect estimates. These results suggest that variability in exposure misclassification resulting

from variability in F is a possible contributor to heterogeneity in PM-mediated health effect estimates.

The hypotheses generated in Chapter 2 were tested in Chapter 3. Using a case-crossover study design and conditional logistic regression, the relative odds of transmural (full-wall) myocardial infarction (MI) calculated using exposure surrogates that account for human activity patterns and the indoor transport of ambient $PM_{2.5}$ were compared with those calculated using central-site $PM_{2.5}$ concentrations to estimate exposure to $PM_{2.5}$ of outdoor origin (exposure to *ambient* $PM_{2.5}$). Because variability in human activity and indoor $PM_{2.5}$ transport contributes exposure error in epidemiologic analyses when central-site concentrations are used as exposure surrogates, surrogates that account for this variability are referred to as "refined" surrogates. As an alternative analysis, whether the relative odds of transmural MI associated with increases in ambient $PM_{2.5}$ is modified by residential air exchange rate (AER), a variable that influences the fraction of ambient $PM_{2.5}$ that penetrates and persists indoors, was evaluated. Use of refined exposure surrogates did not result in larger health effect estimates (ORs = 1.10 - 1.11 with each interquartile range increase.), narrower confidence intervals, or better model fits compared to the analysis that used central-site $PM_{2.5}$. However, evidence for heterogeneity in the relative odds of transmural MI with residential AER (effect-modification) was observed, with residents of homes with higher AERs having larger ORs than homes in lower AER tertiles. For the level of exposure-estimate refinement considered here, these findings add support to the use of central-site $PM_{2.5}$ concentrations for epidemiological studies that employ similar case-crossover study designs. In such designs, each subject serves as his or her own matched control. Thus, exposure error

related to factors that vary spatially or across subjects should only minimally impact effect estimates. It is possible that factors that influence the fraction of ambient $PM_{2.5}$ in indoor air (e.g., AER) could possibly bias health effects estimates in study designs for which a spatio-temporal comparison of exposure effects across subjects is conducted.

The epidemiologic modeling in Chapter 3 was conducted by Dr. David Rich and Kelly Thevenet-Morrison from the University of Rochester School of Medicine and Dentistry and Dr. Pamela Ohman-Strickland from the University of Medicine and Dentistry of New Jersey. Modeling with the Stochastic Human Exposure and Dose Simulation (SHEDS) model to produce Tier 2a exposure estimates was conducted by Dr. Lisa Baxter and colleagues Drs. Janet Burke and Halûk Özkaynak at the United States Environmental Protection Agency National Exposure Research Laboratory. My contribution includes the determination of the mass-balance model inputs that were used by Dr. Melissa Lunden of Lawrence Berkeley National Laboratory to generate the refined exposure estimates that accounted for the indoor transport of ambient $PM_{2.5}$, as well as the analysis and interpretation of the epidemiologic-study results and manuscript preparation.

Chapter 4 provides a partial validation of the exposure estimates used in Chapter 3, while also providing new insights that are used to refine the outdoor-to-indoor transport model. Toward the goal of providing practical methods to model indoor concentrations of ambient $PM_{2.5}$, the mechanistic outdoor-to-indoor transport model was evaluated and refined using measured indoor and outdoor $PM_{2.5}$ species concentrations and air exchange rates from the Relationships of Indoor, Outdoor, and Personal Air Study. Chapter 4 presents model evaluation results, discusses what data are most critical

to prediction of residential exposures at the individual-subject and populations levels, and makes recommendations for the application of the model in epidemiologic studies. This study demonstrates that not accounting for certain human activities (air conditioning and heating use, opening windows) leads to bias in predicted residential PM_{2.5} exposures at the individual-subject level, but not the population level. The analyses presented also provide quantitative evidence that shifts in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport contribute significantly to variability in indoor ambient organic carbon concentrations and suggest that methods to account for these shifts will further improve the accuracy of outdoor-to-indoor transport models.

Chapter 5 presents the first study to address the need for a method to predict shifts in the gas-particle partitioning of total ambient organics with transport into the indoor environment. Here, the change in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport was calculated for 167 homes using measured temperatures, particulate organic matter concentrations, and published organic aerosol (OA) volatility basis sets (VBS). To evaluate the sensitivity of these calculations to uncertainties in the thermodynamic properties of ambient OA, partitioning shifts were calculated assuming enthalpies of vaporization (ΔH_{vap}) of 100 and 50 kJ/mol. Volatility distributions constructed for OA components derived from factor analysis of aerosol mass spectra were used in an alternative analysis. Partitioning shifts were sensitive to ΔH_{vap} assumptions and resulted in changes in indoor concentrations of ambient OA of 11 - 27%, on average, depending on the assumed ΔH_{vap} and whether OA was treated as a single entity or as a mixture of chemically-distinct OA components. Hydrocarbon-like OA was most sensitive to outdoor-to-indoor changes in temperature and OA loading.

Outdoor-to-indoor transport did not induce shifts in partitioning for low-volatility oxygenated OA. The calculations indicate that phase changes are important determinants of residential ambient OA exposure. To account for these phase changes in predictive models, indoor temperatures and estimates (or distributions) of indoor OA emission rates are needed.

Conclusions, future directions, and the broader impacts of this research are discussed in Chapter 6. This work identifies the factors that drive variability in the fraction of ambient $PM_{2.5}$ in indoor air, as well as the geographic locations, seasons, and temporal and spatial scales for which outdoor-to-indoor transport is (and is not) a substantial source of exposure error in PM epidemiology. Modeling tools to address this error are presented and the data and measurements most critical to predicting residential ambient $PM_{2.5}$ exposures are highlighted. In addition, this research evaluates which epidemiologic study designs are (or are not) robust to the type of error introduced by indoor transport of ambient $PM_{2.5}$. This information will be useful in the design of measurement strategies for future exposure and health studies that make use of refined exposure surrogates. Future applications of the tools presented here will help to elucidate relationships between $PM_{2.5}$ exposure and adverse health outcomes and have the potential to aid in the development of strategies to mitigate this exposure.

1.5 References

1. U.S. EPA. Integrated Science Assessment for Particulate Matter (Final Report). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-08/139F, **2009**.
2. Samet, J. M.; Graff, D.; Berntsen, J.; Ghio, A.J.; Huang, Y.C.T.; Devlin, R., A comparison of studies on the effects of controlled exposure to fine, coarse, and ultrafine

ambient particulate matter from a single location. *Inhalation Toxicology* **2007**, 19, 29 - 32.

3. Mustafic, H.; Jabre, P.; Caussin, C.; Murad, M. H.; Escolano, S.; Tafflet, M.; Perier, M.C.; Marijon, E.; Vernerey, D.; Empana, J. P.; Jouven, X., Main air pollutants and myocardial infarction: a systematic review and meta-analysis. *Journal of the American Medical Association* **2012**, 307: 713 - 721.

4. Eftim, S. E.; Samet, J. M.; Janes, H.; McDermott, A.; Dominici, F., Fine particulate matter and mortality: A comparison of the Six Cities and American Cancer Society cohorts with a medicare cohort. *Epidemiology* **2008**, 19, 209 - 216.

5. Franklin, M.; Zeka, A.; Schwartz, J., The role of particle composition on the association between PM_{2.5} and mortality. *Epidemiology* **2008**, 19, 680 - 689.

6. Bell, M. L.; Ebisu, K.; Peng, R. D.; Samet, J. M.; Dominici, F., Hospital admissions and chemical composition of fine particle air pollution. *American Journal Respiratory and Critical Care* **2009**, 179, 1115 - 1120.

7. Peng, R. D.; Bell, M. L.; Geyh, A. S.; McDermott, A.; Zeger, S. L.; Samet, J. M.; Dominici, F., Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. *Environmental Health Perspectives* **2009**, 117, 957 - 963.

8. Ito, K.; Mathes, R.; Ross, Z.; Nadas, A.; Thurston, G.; Matte, T., Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. *Environmental Health Perspectives* **2011**, 119, 467 - 473.

9. Hoek, G.; Brunekreef, B.; Goldbohm, S.; Fischer, P.; van den Brandt, P. A., Association between mortality and indicators of traffic-related air pollution in the Netherlands. *Lancet* **2002**, 360, 1203 - 1209.

10. Gauderman, W. J.; Avol, E.; Lurmann, F.; Kuenzli, N.; Gilliland, F.; Peters, J.; McConnell, R., Childhood asthma and exposure to traffic and nitrogen dioxide. *Epidemiology* **2005**, 16, 737 - 743.

11. Brauer, M.; Lencar, C.; Tamburic, L.; Koehoorn, M.; Demers, P.; Karr, C., A cohort study of traffic-related air pollution impacts on birth outcomes. *Environmental Health Perspectives* **2008**, 116, 680 - 686.

12. Dockery, D. W.; Pope III, A. C.; Xu, X.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; Ferris, B. J.; Speizer, F. E., An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine* **1993**, 329, 1753 - 1759.

13. Pope III, A. C.; Burnnett, R. T.; Thun, M. J.; Calle, E. E.; Krewski, D.; Ito, K.; Thurston, G. D., Lung cancer, cardiopulmonary mortality and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association* **2002**, 287, 1132 - 1141.
14. Krewski, D.; Burnnett, R.; Goldberg, M.; Hoover, B. K.; Siemiatcki, J.; Jerrett, M.; Abrahamowicz, M.; White, W., Overview of the reanalysis of the Harvard Six Cities study and the American Cancer Society study of particulate air pollution and mortality. *Journal of Toxicology and Environmental Health* **2003**, 66, 1507 - 1552.
15. Rich, D. Q.; Özkaynak, H.; Crooks, J.; Baxter, L.; Burke, J.; Ohman-Strickland, P.; Thevenet-Morrison, K.; Kipen, H. M.; Zhang, J.; Kostis, J. B.; Lunden, M.; Hodas, N.; Turpin, B., The triggering of myocardial infarction by fine particles is enhanced when particles are enriched in secondary species. *Environmental Science and Technology* **2013**, in press.
16. Klepeis, N. E.; Nelson, W. C.; Ott, W. R.; Robinson, J. P.; Tsang, A. M.; Switzer, P.; Behar, J. V.; Stephen, C. H.; Engelman, W. H., The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *Journal of Exposure Science and Environmental Epidemiology* **2001** 11, 231 - 252.
17. Long, C. M.; Suh, H. H.; Catalano, P. J.; Koutrakis, P., Using time- and size-resolved particulate data to quantify indoor penetration and deposition behavior. *Environmental Science and Technology* **2001**, 35, 2089 - 2099.
18. Meng, Q. Y.; Turpin, B. J.; Polidori, A.; Lee, J. H.; Weisel, C. P.; Morandi, M.; Winer, A.; Zhang, J., PM_{2.5} of ambient origin: estimates and exposure errors relevant to PM epidemiology. *Environmental Science and Technology* **2005**, 39, 5105 - 5112.
19. Lunden, M. M.; Thatcher, T. L.; Hering, S. V.; Brown, N. J., The use of time- and chemically-resolved particulate data to characterize the infiltration of outdoor PM_{2.5} into a residence in the San Joaquin Valley. *Environmental Science and Technology* **2003**, 37, 4724 -4732.
20. Lunden, M. M.; Revzan, K. L.; Fischer, M. L.; Thatcher, T. L.; Littlejohn, D.; Hering, S. C.; Brown, N. J., The transformation of outdoor ammonium nitrate aerosols in the indoor environment. *Atmospheric Environment* **2003**, 37, 5633 - 5644.
21. Lunden, M. M.; Kirchstetter, T. W.; Thatcher, T. L.; Hering, S. V.; Brown, N. J., Factors affecting the indoor concentration of carbonaceous aerosols of outdoor origin. *Atmospheric Environment* **2008**, 42, 5660 - 5671.
22. Sarnat, S. E.; Coull, B. A.; Ruiz, P. A.; Koutrakis, P.; Suh, H. H., The influences of ambient particle composition and size on particle infiltration in Los Angeles, CA, Residences. *Journal of the Air and Waste Management Association* **2006**, 56, 186 - 196.

23. Zeger, S. L.; Thomas, D.; Dominici, F.; Samet, J. M.; Schwartz, J.; Dockery, D.; Cohen, A., Exposure measurement error in time-series studies of air pollution: Concepts and consequences. *Environmental Health Perspectives* **2000**, 108, 419 - 426.
24. Bateson, T. F.; Coull, B. A; Hubbell, B.; Kazuhiko, I.; Jerrett, M.; Lumley, T.; Thomas, D.; Vedal, S.; Ross, M., Panel discussion review: session 3 – issue involved in interpretation of epidemiological analyses – statistical modeling. *Journal of Exposure Analysis and Environmental Epidemiology* **2007**, 17, S90 - S96.
25. Janssen, N. A.; Schwartz, H. J.; Zanobetti, A.; Suh, H. H., Air conditioning and source-specific particles as modifiers of the effect of PM10 on hospital admissions for heart and lung disease. *Environmental Health Perspectives* **2002**, 110, 43 - 49.
26. Zeka, A.; Zanobetti, A.; Schwartz, J., Short-term effects of particulate matter on cause specific mortality: effects of lags and modification by city-specific characteristics. *Occupational and Environmental Medicine* **2005** 62, 718 - 725.
27. Franklin, M.; , A.; Schwartz, J., Association between PM2.5 and all-cause and specific cause mortality in 27 US communities. *Journal of Exposire Science Environmental Epidemiology* **2007**, 17, 279 - 287.
28. Bell, M. L.; Ebisu, K.; Peng, R. D.; Dominici, F., Adverse health effects of particulate air pollution: modification by air conditioning. *Epidemiology* **2009**, 19, 680 - 689.
29. Thornburg, J.W.; Ensor, D. S.; Rodes, C. E.; Lawless, P. A.; Sparks, L. E.; Mosely, R. B., Penetration of particles into buildings and associated physical factors, Part I: model development and computer simulations. *Aerosol Science and Technology* **2001**, 34, 284 - 296.
30. Waring, M. S.; Siegel, J. A., Particle loading rates for HVAC filters, heat exchangers, and ducts. *Indoor Air* **2008**, 18, 209 - 224.
31. Stephens, B.; Siegel, J. A., Comparison of test methods for determining the particle removal efficiency of filters in residential and light-commercial central HVAC systems. *Aerosol Science and Technology* **2012**, 46, 504 - 513.
32. Chen, C.; Zhao, B.; Weschler, C. J., Assessing the influence of indoor exposure to “outdoor ozone” on the relationship between ozone and short-term mortality in U.S. communities. *Environmental Health Perspectives* **2012**, 120, 235 - 240.
33. Chen, C.; Zhao, B.; Weschler, C. J., Indoor exposure to "outdoor PM10": assessing its influence on the relationship between PM10 and short-term mortality in U.S. cities. *Epidemiology* **2012**, 23, 870 - 878.

34. Chen, R.; Zhou, B.; Kan, H.; Zhao, B., Associations of particulate air pollution and daily mortality in 16 Chinese cities: An improved effect estimate after accounting for the indoor exposure to particles of outdoor origin. *Environmental Pollution* **2013**, 182, 278 - 282.
35. Hering, S. V.; Friedlander, S. K.; Origins of aerosol sulfur size distributions in the Los Angeles basin. *Atmospheric Environment* **1982**, 16, 2647 - 2656.
36. Finlayson-Pitts, B. J.; Pitts Jr., J. N., Chemistry of the Upper and Lower Atmosphere. Academic Press: San Diego, CA, **2000**.
37. Blando, J. D.; Turpin, B. J., Secondary organic aerosol formation in cloud and fog droplets: a literature evaluation of plausibility. *Atmospheric Environment* **2000**, 34, 1623 - 1632.
38. Ervens, B.; Turpin, B. J.; Weber, R. J., Secondary organic aerosol formation in cloud droplets and aqueous particles (aqSOA): a review of laboratory, field, and model studies. *Atmospheric Chemistry and Physics* **2011**, 11, 11069 - 11102.
39. Seinfeld, J. H.; Pandis, S. N., From air pollution to climate change. John Wiley & Sons: New York, NY, **1998**.
40. Riley, W. J.; McKone, T. E.; Lai, A. C. K.; Nazaroff, W. W., Indoor particulate matter of outdoor origin: importance of size-dependent removal mechanisms. *Environmental Science and Technology* **2002**, 36, 200 - 207.
41. Nazaroff, W. W., Indoor particle dynamics. *Indoor Air* **2004**, 14, 175 - 183.
42. Hering, S. V.; Lunden, M. M.; Thatcher, T. L.; Kirchstetter, T. W.; Brown, N. J., Using regional data and building leakage to assess indoor concentrations of particles of outdoor origin. *Aerosol Science and Technology* **2007**, 41, 639 - 654.
43. U.S. EPA (Environmental Protection Agency). (2011) Exposure factors handbook **2011** edition. National Center for Environmental Assessment, Washington, DC; EPA/600-09/052F. Available from the National Technical Information Service, Springfield, VA, and online at <http://www.epa.gov/ncea/efh>.
44. Chan, W. R.; Nazaroff, W. W.; Price, P. N.; Sohn, M. D.; Gadgil, A. J., Analyzing a database of residential air leakage in the United States. *Atmospheric Environment* **2005**, 39, 3445 - 3455.
45. Dietz, R. N.; Cote, E. A., Air infiltration measurements in a home using a convenient perfluorocarbon tracer technique. *Environment International* **1982**, 8, 419 - 433.

46. Dietz, R. N.; Goodrich, R. W.; Cote, E. A.; Wiesner, R. F., Detailed description and performance of a passive perfluorocarbon tracer system for building ventilation and air exchange measurements. *Measured air leakage of buildings ASTM STP* **1986**, 904, 203 - 264.
47. Pandian, M. D.; Ott, W. R.; Behar, J. V., Residential air exchange rates for use in indoor air and exposure modeling studies. *Journal of Exposure Analysis and Environmental Epidemiology* **1993**, 3, 407 - 416.
48. Pandian, M. D.; Behar, J. V.; Ott, W. R.; Wallace, L. A.; Wilson, A. L.; Colome S. D.; Koontz, M., Correcting errors in the nationwide data base of residential air exchange rates. *Journal of Exposure Analysis and Environmental Epidemiology* **1998**, 8, 577 - 586.
49. Yamamoto, N.; Shendell, D. G.; A. M. Winer; Zhang, J., Residential air exchange rates in three major US metropolitan areas: results from the Relationships Among Indoor Outdoor and Personal Air study 1999-2001. *Indoor Air* **2010**, 20, 85 - 90.
50. Sherman, M. H.; Grimsrud, D. T., Measurement of infiltration using fan pressurization and weather data. Lawrence Berkeley National Laboratory Report. LBNL-10852, Berkeley, CA, **1980**.
51. Sherman, M. H.; Dickerhoff, D. J., Airtightness of US dwellings. *ASHRAE Transactions* **1998**, 104, 1359 - 1367.
52. Liu, D. L.; Nazaroff, W. W., Modeling particle penetration across building envelopes. *Atmospheric Environment* **2001**, 35, 4451 - 4462.
53. Ozkaynak, H. J.; Xue, R.; Weker, D.; Butler, P.; Koutrakis, P.; Spengler, J., The particle TEAM (PTEAM) study: analysis of the data (final report, volume III), U.S. Environmental Protection Agency, Washington, DC, **1996**.
54. Thatcher, T. L.; Lunden, M. M.; Revzan, K. L.; Sextro, R. G.; Brown, N. J., A concentration rebound method for measuring particle penetration and deposition in the indoor environment. *Aerosol Science and Technology* **2003**, 37, 847 - 864.
55. Chen, C.; Zhao, B., Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. *Atmospheric Environment* **2011**, 45, 275 - 288.
56. Meng, Q. Y.; Turpin, B. J.; Lee, J. H.; Polidori, A.; Weisel, C. P.; Morandi, M.; Colome, S.; Zhang, J.; Stock, T.; Winer, A., How does infiltration behavior modify the composition of ambient PM_{2.5} in indoor spaces? An analysis of RIOPA data. *Environment Science and Technology* **2007**, 41, 7315 - 7321.

57. Bennett, D. H.; Koutrakis, P., Determining the infiltration of outdoor particles in the indoor environment using a dynamic model. *Journal of Aerosol Science* **2006**, 37, 766 - 785.
58. Naumova, Y. Y.; Offenberg, J. H.; Eisenreich, S. J.; Meng, Q. Y.; Polidori, A.; Turpin B. J.; Weisel, C. P.; Morandi, M. T.; Colome, S. D.; Stock, T. H.; Winer, A. M.; Alimokhtari, S.; Kwon, J.; Maberti, S.; Shendell, D.; Jones, J.; Farrar, C., Gas/particle distribution of polycyclic aromatic hydrocarbons in coupled indoor/outdoor atmospheres. *Atmospheric Environment* **2003**, 37, 703 - 719.
59. Polidori, A.; Turpin, B.; Meng, Q. Y.; Lee, J. H.; Weisel, C.; Morandi, M.; Colome, S.; Stock, T.; Winer, A.; Zhang, J.; Kwon, J.; Alimokhtari, S.; Shendell, D.; Jones, J.; Farrar, C.; Maberti, S., Fine organic particulate matter dominates indoor-generated PM_{2.5} in RIOPA homes. *Journal of Exposure Science and Environmental Epidemiology* **2006**, 16, 321 - 331.
60. Shi, S., Zhao, B., Comparison of the predicted concentration of outdoor originated indoor Polycyclic Aromatic Hydrocarbons between a kinetic partition model and a linear instantaneous model for gas-particle partition. *Atmospheric Environment* **2012**, 59, 93-101.
61. Pankow, J. F., An absorption model of gas/particle partitioning of organic compounds in the atmosphere. *Atmospheric Environment* **1994**, 28, 185 - 188.
62. Donahue, N. M.; Robinson, A. L.; Stanier, C. O.; Pandis, S. N., Coupled partitioning, dilution, and chemical aging of semivolatile organics. *Environmental Science and Technology* **2006**, 40, 2635 - 2643.
63. Jimenez, J. L.; Canagaratna, M. R.; Donahue, N. M.; Prevot, A. S. H.; Zhang, Q.; Kroll, J. H.; DeCarlo, P. F.; Allan, J. D.; Coe, H.; Ng, N. L.; Aiken, A. C.; Docherty, K. S.; Ulbrich, I. M.; Grieshop, A. P.; Robinson, A. L.; Duplissy, J.; Smith, J. D.; Wilson, K. R.; Lanz, V. A.; Hueglin, C.; Sun, Y. L.; Tian, J.; Laaksonen, A.; Raatikainen, T.; Rautiainen, J.; Vaattovaara, P.; Ehn, M.; Kulmala, M.; Tomlinson, J. M.; Collins, D. R.; Cubison, M. J.; Dunlea, E. J.; Huffman, J. A.; Onasch, T. B.; Alfarra, M. R.; Williams, P. I.; Bower, K.; Kondo, Y.; Schneider, J.; Drewnick, F.; Borrmann, S.; Weimer, S.; Demerjian, K.; Salcedo, D.; Cottrell, L.; Griffin, R.; Takami, A.; Miyoshi, T.; Hatakeyama, S.; Shimono, A.; Sun, J. Y.; Zhang, Y. M.; Dzepina, K.; Kimmel, J. R.; Sueper, D.; Jayne, J. T.; Herndon, S. C.; Trimborn, A. M.; Williams, L. R.; Wood, E. C.; Middlebrook, A. M.; Kolb, C. E.; Baltensperger, U.; Worsnop, D. R., Evolution of organic aerosol in the atmosphere. *Science* **2009**, 326, 1525 - 1529.
64. Ott, W.; Wallace, L.; Mage, D., Predicting particulate (PM₁₀) personal exposure distributions using a random component superposition statistical model. *Journal of the Air and Waste Management Association* **2000**, 50, 1390 - 1406.

65. Alzona, J.; Cohen, B. L.; Jow, R. H. N.; Frohlinger, J. O., Indoor-outdoor relationships for airborne particulate matter of outdoor origin. *Atmospheric Environment* **1978**, 13, 55 - 60.
66. Koutrakis, P.; Briggs, S. L. K.; Leaderer, B. P., Source apportionment of indoor aerosols in Suffolk and Onondaga Counties, New York. *Environmental Science and Technology* **1992**, 26, 521 - 527.

Chapter 2. Variability in the Fraction of Ambient Fine Particulate Matter Found Indoors and Observed Heterogeneity in Health Effect Estimates

Material in this chapter has been published previously as:

Hodas, N.; Meng, Q. Y.; Lunden, M. M.; Rich, D. Q.; Özkaynak, H.; Baxter, L. K.; Zhang, Q.; Turpin, B. J., Variability in the fraction of ambient fine particulate matter found indoors and observed heterogeneity in health effect estimates. *J. Exposure Sci. Environ. Epidemiol.* **2012**, 22, 448 - 454.

2.1 Abstract

Exposure to ambient (outdoor-generated) fine particulate matter (PM_{2.5}) occurs predominantly indoors. The variable efficiency with which ambient PM_{2.5} penetrates and persists indoors is a source of exposure error in air pollution epidemiology and could contribute to observed temporal and spatial heterogeneity in health effect estimates. Using a mass balance approach, F was modeled for several scenarios across which heterogeneity in effect estimates has been observed: with geographic location of residence, residential roadway proximity, socioeconomic status, and central air conditioning use. Calculated values of F are higher in close proximity to primary combustion sources (e.g. proximity to traffic) and for lower income homes. F is lower when PM_{2.5} is enriched in nitrate and with central air conditioning use. As a result, exposure error resulting from variability in F will be greatest when these factors have high temporal and/or spatial variability. The circumstances for which F is lower in these calculations correspond to circumstances for which lower effect estimates have been observed in epidemiological studies and higher F values correspond to higher effect estimates. These results suggest that variability in exposure misclassification resulting

from variability in F is a possible contributor to heterogeneity in PM-mediated health effect estimates.

2.2 Introduction

Multiple epidemiologic studies have reported spatial and temporal heterogeneity in PM-mediated health effect estimates. Notably, larger effects have been observed for the eastern U.S. compared to the western U.S.¹⁻⁴ Several studies report larger risks of adverse health outcomes when PM_{2.5} is enriched in primary combustion tracers⁴⁻⁷ and with proximity to roadway.⁸⁻¹¹ Low socioeconomic status has also been identified as a predictor of susceptibility to PM-related health effects.¹² Various factors have been explored to explain this variability. Bell et al.⁵ concluded that as much as 37% of spatial and seasonal heterogeneity in relative risk of cardiovascular hospital admissions could be explained by variability in exposure to specific PM_{2.5} species. Sacks et al.¹² identified disparities in access to health care as one possible contributor to variability in effect estimates by socioeconomic status. This work explores whether variability in the fraction of ambient PM_{2.5} that penetrates into and persists in the home is also a possible contributor to observed heterogeneity in fine-PM-mediated health effect estimates.

Central site PM_{2.5} is commonly used as a surrogate for exposure to ambient (outdoor-generated) PM_{2.5} in epidemiology. However, people spend the majority of their time indoors (85-90%), and most of that time in their homes.¹³ As a result, exposure to ambient PM_{2.5} mostly occurs in the indoor environment and, specifically, within the residence. Importantly, the fraction of ambient PM_{2.5} that penetrates and persists indoors (F) varies temporally¹⁴ and spatially¹⁵ and is different for different components of the PM_{2.5} mixture.¹⁶ Exposure metrics that rely on central site concentrations do not account

for this variability, nor do they account for changes in $PM_{2.5}$ properties (i.e., composition and size distribution) that result from outdoor-to-indoor transport.¹⁵ Variability in F has been identified as a potential source of non-Berksonian exposure error that could lead to a downward bias and underestimation of effects derived from epidemiological analyses¹⁷⁻¹⁹ when central-site $PM_{2.5}$ is used as an estimate of ambient $PM_{2.5}$ exposure. Several studies are underway which will apply refined exposure surrogates that account for the effects of indoor transport to health studies. This paper articulates the motivation for those studies.

There is evidence that F can influence health effect estimates in epidemiological studies. Several studies show a reduced risk of mortality or morbidity associated with PM when the prevalence of air conditioning (AC) is higher.^{2,20-22} Bell et al.²² reported a 43% decrease in risk of cardiovascular hospitalization associated with each $10 \mu\text{g}/\text{m}^3$ increase in $PM_{2.5}$ for every additional 20% of households with central AC. Central AC prevalence explained 17% of between-community variability in $PM_{2.5}$ effect estimates. The use of AC increases particle losses indoors and, therefore, decreases F and exposure to ambient $PM_{2.5}$.^{23,24} Further, homes with AC in use are more likely to have their windows closed and, therefore, have lower air exchange rates (and F) compared to homes with open windows.²⁵ Certainly, the use of AC varies seasonally and geographically, leading to heterogeneity in F across regions of the U.S. and across seasons.

There are additional factors besides AC that have a dramatic impact on F . F depends on the rate at which air within the building is exchanged with outdoor air (air exchange rate; AER), the efficiency of particle penetration across the building envelope, depositional losses in indoor air, and, for semi-volatile species, particle losses or gains due to phase changes. These parameters vary temporally and spatially with factors such

as meteorological conditions, building characteristics, human activities, and PM_{2.5} source mix.^{26,27}

This paper explores whether variability in F could contribute to the observed heterogeneity in effect estimates with geographic location of residence, residential proximity to roadways, air conditioning use, and socioeconomic status. It is hypothesized that regions and circumstances for which lower effect estimates have been observed in epidemiological studies (i.e., with AC use and residence in the western U.S.) have lower F values and that regions and circumstances with higher effect estimates (i.e., residence in the eastern U.S., close proximity to roadways, and low socioeconomic status) have higher F values. While other factors undoubtedly also contribute to differences in effect estimates, it should be noted that when F is lower, the non-differential exposure misclassification that results when central-site PM_{2.5} concentrations are used to estimate exposure is greater, resulting in larger bias towards the null and underestimation of effects.

2.3 Methods

2.3.1 Main Analysis

To identify situations for which variations in F may contribute substantially to exposure error, a mass balance approach was used to model F under a variety of scenarios related to the conditions under which heterogeneity in effect estimates has been observed (i.e., with geographic location of residence, residential roadway proximity, socioeconomic status, and air conditioning use). Specifically, F was modeled for two homes, a “typical” home and a low-income home, under a variety of conditions: location

in the eastern U.S. and the western U.S., in close proximity and further from a roadway, and with and without central AC in use.

For non-volatile species, the mass balance model describes the concentration of PM_{2.5} species j in indoor air ($C_{in,j}$) as a function of its outdoor concentration ($C_{out,j}$), residential air exchange rate (AER), the efficiency of particle penetration across the building envelope (P_j), and the rate of indoor loss by deposition ($k_{dep,j}$):

$$\frac{dC_{in,j}}{dt} = C_{out}(P_j \times AER) - C_{in,j}(k_{dep,j} + AER) \quad (1a)$$

For nitrate, which is semi-volatile and can undergo phase changes, the model includes the rate of indoor loss by evaporation (k_{evap,NO_3^-}):

$$\frac{dC_{in,NO_3^-}}{dt} = C_{out,NO_3^-}(P_j \times AER) - C_{in,NO_3^-}(k_{dep,NO_3^-} + AER + k_{evap,NO_3^-}) \quad (1b)$$

The steady state solution to Equations (1a) and (1b) was used in the calculations below to solve for $C_{in,j}$:

$$C_{in,j} = C_{out,j} \times \frac{(P_j \times AER)}{(AER + k_{dep,j})} \quad (2a)$$

$$C_{in,NO_3^-} = C_{out,NO_3^-} \times \frac{(P \times AER)}{(AER + k_{dep,NO_3^-} + k_{evap})} \quad (2b)$$

Summing $C_{in,j}$ over all major particle species provides the indoor concentration of ambient PM_{2.5}, since no indoor sources are included in the model. Without indoor sources, F is the ratio of the PM_{2.5} concentration indoors to outdoors:

$$F = \frac{C_{in}}{C_{out}} \quad (3)$$

A version of the model used in this analysis has been previously validated with real-time particle species and AER measurements made inside and outside of an unoccupied (closed) house in California.^{28,29} Calculated indoor concentrations captured

well the attenuation of the outdoor concentrations and were highly correlated with indoor measurements ($R^2 = 0.8-0.93$).²⁹

Using this mass balance model, the indoor concentrations and composition of ambient $PM_{2.5}$ were predicted for several scenarios representative of conditions for which variability in effect estimates has been observed. First, $PM_{2.5}$ composition and size distributions were varied in calculations of F to capture variability across geographic location in the U.S. and with roadway proximity. For this analysis, the major contributors to $PM_{2.5}$ mass (sulfate, nitrate, elemental carbon (EC), organic carbon (OC), and soil) were considered and residential outdoor composition measurements from the Relationships of Indoor, Outdoor, and Personal Air (RIOPA) study³⁰ and species-specific size distributions from an Aerosol Mass Spectrometer (AMS)^{31,32} were used. Certainly disparities in health effects could result from differences in concentrations of other $PM_{2.5}$ species (e.g., metals), but these species comprise a very small fraction of total fine particle mass and will not affect F in a measureable way. The RIOPA OC measurements are artifact corrected and were converted to organic matter (OM) using a factor of 1.4.³⁰ It should be noted that in the RIOPA study, all major fine particle species were measured, with the exception of nitrate and water. For the calculations presented here, it was assumed that nitrate completes the species mass balance, providing an upper bound estimate for nitrate.

$PM_{2.5}$ composition for each scenario is shown in Figure 2-1. Scenario (a) represents a typical northeastern U.S. $PM_{2.5}$ composition; it is the mean composition measured outside Elizabeth, NJ homes during the RIOPA study.³⁰ Scenario (b) represents a high $PM_{2.5}$ episode in the southwestern U.S; it is enriched in nitrate. Scenario (b) has

the mean composition of the top 25th percentile PM_{2.5} mass concentration days from the Los Angeles County RIOPA measurements.³⁰ Scenario (c), which represents a northeastern U.S. near roadway scenario, is enriched in OC and EC based on the near-roadway measurements of Lena et al.³³

Particle composition is taken into account in the calculations of F through the use of species-specific deposition loss rates and by accounting for the semivolatile nature of nitrate. The value for $k_{\text{evap},\text{NO}_3^-}$ is from the work of Lunden et al.³⁴ and Hering et al.²⁹ and involves the temperature-dependent equilibrium constant for ammonium nitrate dissociation (Appendix A). Other PM components, including organic PM, were treated as non-volatile. Size-resolved k_{dep} values were fit based on species-resolved size distributions as described below. Thus, changes in PM_{2.5} composition result in changes in k_{dep} and changes in F .

Because particle size distributions are driven by formation mechanisms, ambient *species-specific* size distributions are less variable than size distributions for total particle mass (Appendices A1, A2). Therefore, a k_{dep} value was assigned for each of the major PM_{2.5} species based on species-resolved size distributions measured with the Aerodyne Aerosol Mass Spectrometer (AMS), which measures vacuum aerodynamic diameter.³⁵ For the northeastern United States composition scenarios, species-specific size distributions measured in Queens, NY in August 2001³¹ were used. Species-specific size distributions measured in Fresno, CA in January 2010³² were used to represent the southwestern U.S.

Hourly, campaign-averaged species-resolved size distributions (Appendix A) were examined to determine the number and frequency of modes, the mass median

diameter (MMD) and geometric standard deviation of each mode, and the fraction of total mass in each mode when distributions were bimodal. The most frequently observed distribution was chosen as representative of each species (Figure 2-2). For all composition scenarios, a mass median diameter of 2.0 μm was assumed for soil. It was also assumed that the smaller diameter mode observed for OC in New York was representative of EC (The AMS does not measure EC). For scenario (c), the near roadway scenario, both EC and OC were assumed to be entirely in the smaller diameter mode, as is likely for fresh, primary $\text{PM}_{2.5}$.

Values of k_{dep} for the MMD of each species were assigned using the deposition curve presented in Riley et al.,²⁶ which combines a physical particle deposition model with empirically determined k_{dep} values from several sampling studies. Results are summarized in Table 2-1. For all calculations and species, a constant P_j value of 0.8, the median of the range of measured P values reported in the literature for particles in the size range examined here ($P = 0.6 - 1.0$),³⁶ was assumed.

For the northeastern U.S. mean composition scenario, the effects of central air conditioning on F were also considered (scenario d). As noted above, air conditioning increases particle losses indoors by filtering re-circulated air. To account for this, a species-specific filter penetration efficiency term (P_{filter}) was added to the model based on the particle-size-resolved filter efficiency curve presented in Riley et al.²⁶ for residential buildings using the same MMDs as were used to assign k_{dep} values. Values of P_{filter} for each species are given in Table 2-1.

For each scenario described above, F was calculated for two AERs. AERs from the literature were used to represent typical U.S. homes and low-income residences,

respectively, in order to illustrate how F varies with the socioeconomic status of the residents. The median measured AER from an aggregation of AER measurement studies that spanned all climatic regions of the U.S. was used to represent typical U.S. housing stock (0.45 h^{-1}).³⁷ The geometric mean of AERs measured in 255 low-income homes (0.90 h^{-1})³⁷ was used to represent low-income residences. AER also varies within homes with meteorological conditions, season, and human activity patterns (e.g., opening windows).

2.3.2 Sensitivity Analyses

The sensitivity of the mass balance model to uncertainty in $\text{PM}_{2.5}$ species size distributions was evaluated. For scenario (a), F was calculated for total $\text{PM}_{2.5}$ by fitting k_{dep} values based on the size distribution for total $\text{PM}_{2.5}$ mass (i.e., the sum of the measured species). The total $\text{PM}_{2.5}$ mass size distribution is bimodal with a smaller mode MMD near $0.08 \text{ }\mu\text{m}$ and a larger mode MMD near $0.4 \text{ }\mu\text{m}$, which correspond to k_{dep} values of 0.05 and 0.07 h^{-1} , respectively. Because $\text{PM}_{2.5}$ mass is not broken down by species here, evaporative losses of nitrate with transport indoors are not accounted for in this calculation. To separate the effects of not accounting for the evaporative losses of nitrate and not accounting for the variability in k_{dep} across species, F was also recalculated for scenario (a) neglecting evaporative losses of nitrate. In this case, variability in k_{dep} across species is accounted for, but nitrate is treated as non-volatile.

The sensitivity of the model to variations in species-specific size distributions across season and location was also evaluated. Size- and species-resolved $\text{PM}_{2.5}$ concentration data are available for only a limited number of cities and, where they are available, are generally limited in their temporal scope. Thus, it is important to evaluate

how temporal and spatial variability in $\text{PM}_{2.5}$ species size distributions might contribute to uncertainty in the methods illustrated here. F was recalculated for scenarios (a) and (d) using k_{dep} and P_{filter} values fit based on the Fresno size distribution data in place of the New York size distribution data. F was also recalculated for scenarios (a) and (d) with k_{dep} and P_{filter} values fit using size distributions measured in New York during the winter.³¹ Winter size distributions were analyzed in the same manner as described above. During the winter sampling campaign, particulate sulfate, nitrate and organics had size distributions that could be characterized by a single mode with an MMD near $0.3\ \mu\text{m}$. EC and Soil MMDs were held constant at the values used in the main analysis. This calculation was done for both scenarios (a) and (d) because, for the particle size ranges considered here, the filter efficiency curve is more sensitive to changes in particle diameter than the deposition-rate curve.²⁶

To evaluate the sensitivity of the model to uncertainty in size distribution *measurement*, the MMD of sulfate and nitrate were varied between 0.5 and $0.7\ \mu\text{m}$ in the calculations of F . Previous studies have reported an MMD of $0.7\ \mu\text{m}$ for nitrate and sulfate size distributions when measurements are resolved to account for sampler collection efficiencies.^{e.g. 38,39} The impact of uncertainty in EC size distributions was also considered. Notably, the AMS does not measure EC size distributions. In the main analyses, it was assumed that EC was represented by the smaller mode of the bimodal OC size distribution measured in New York. Here, we recalculated F for scenario (a), assuming EC has the same bimodal distribution as OC, as might be expected for an air mass characterized by a mix of primary and regionally-transported, aged $\text{PM}_{2.5}$.

Uncertainty in particle size distributions also contributes to uncertainty in P_j because, like deposition, the loss mechanisms that govern penetration efficiency are particle-size dependent. P_j also varies with building ventilation conditions and the geometry of cracks in the building shell.^{36,40,41} While it is known from laboratory studies that complex crack geometry and increased roughness of crack surfaces result in lower P values, the characteristics of cracks in buildings are not well known and are likely to be highly varied within and across buildings.^{36,40} Furthermore, size-resolved measurements of P_j under a variety of ventilation conditions and building characteristics are limited. As a result, there is substantial uncertainty in P_j . Chen and Zhao³⁶ reported that P values measured in real buildings generally range from 0.6 to 1.0 for particles in the size range considered here. Williams et al.,⁴¹ however, reported a minimum P value of 0.11 measured during the Research Triangle Park Particulate Matter Panel Study. The effect of uncertainty in P_j was evaluated by calculating the change in F if P_j was 0.6, 1.0 and 0.11, rather than 0.8 for all species.

2.4 Results

2.4.1 Main Analysis

Outdoor-to-indoor transport reduced ambient $\text{PM}_{2.5}$ concentrations to roughly one-half of their outdoor values, on average, in the modeling scenarios, depending on AER, $\text{PM}_{2.5}$ composition, and air-conditioning use (Figure 2-3). Variations in F due to variations in $\text{PM}_{2.5}$ composition alone and AER alone were comparable. When considering variations in $\text{PM}_{2.5}$ composition alone (i.e., comparing scenarios (a) – (c) at a fixed AER), F changed by between 7 and 32%. When considering variations in AER alone (i.e. comparing each scenario across AERs), F changed by between 13 and 22%. It

was higher for the “near roadway” scenario and at the “low income” AER of 0.90 h^{-1} and it was lower for the “high nitrate” scenario and with air conditioning. Notably, F was nearly a factor of two greater for the “near roadway” scenario at the “low income” AER ($F=0.62$) than for the high nitrate scenario at the “typical” AER ($F=0.36$). This suggests heterogeneity in F between low income urban residents and higher income suburban residents in the southwestern U.S., for example. The semi-volatile nature of nitrate was responsible for the largest compositional effects. At an AER of 0.45 h^{-1} , ambient particulate nitrate indoors was only 12% of its outdoor concentration; in contrast F was 0.67 for sulfate. Increased particle losses due to the filtration of re-circulated air in central air conditioning systems had the largest overall impact on F . Compared to the natural ventilation scenario with the same $\text{PM}_{2.5}$ composition and size distribution, air conditioning decreased F by 34 and 35% at AERs of 0.45 and 0.90 h^{-1} , respectively.

Calculated values of F (Figure 2-3) are in agreement with the range of F values reported by several studies ($F= 0.32\text{-}0.8$).³⁶ Sarnat et al.⁴² and Lunden et al.⁴³ observed higher values of F for EC (0.84 and 0.64, respectively) compared to other species. This is consistent with the higher F values we calculated for EC (0.71, 0.75 for AERs of 0.45 and 0.90 h^{-1} , respectively) and the higher F values we calculated for the high carbonaceous aerosol (near roadway) scenario.

2.4.2 Sensitivity Analyses

F increased substantially when $\text{PM}_{2.5}$ mass was treated as a single entity, rather than a mix of separate species with distinct loss rates. Using k_{dep} values fit for total $\text{PM}_{2.5}$ mass, F values of 0.69 and 0.74 were obtained for AERs of 0.45 and 0.90 h^{-1} , respectively. This is an increase in F of 30 and 23% over the values obtained for scenario

(a) in the initial analysis. Not accounting for variability in k_{dep} across species and not accounting for the evaporative losses of nitrate both contributed to this increase in F . Accounting for variability in k_{dep} across species, but not evaporative losses of nitrate resulted in F values of 0.66 and 0.72 at AERs of 0.45 and 0.90 h^{-1} , respectively.

The model showed little sensitivity to changes in k_{dep} with variations in species-specific size distributions across season and location; it was more sensitive to changes in P_{filter} . The overall F value for scenario (a: NE composition) did not change when k_{dep} values fit based on the Fresno species-specific size distributions were used rather than the New York City size distributions. Similarly, F values changed by less than 2% when k_{dep} values were selected based on the winter New York size distributions rather than the summer values. Using Fresno size distributions in place of New York Size distributions to fit k_{dep} and P_{filter} for scenario (d) resulted in an increase in F of only about 3%, but using winter New York size distributions to fit k_{dep} and P_{filter} rather than summer values, resulted in an increase of F of 15 – 17%, depending on AER.

Varying the sulfate MMD between 0.7 and 0.5 μm changed F by less than 3% (F for sulfate is 0.62 assuming a MMD of 0.7 μm and 0.67 assuming a MMD of 0.5 μm , for an AER of 0.45 h^{-1}). Varying the MMD for nitrate over the same range had a negligible effect on F because evaporation, not deposition, is the dominant loss mechanism. Using k_{dep} values fit assuming EC was bimodal for scenario (a) had no effect on F .

F is sensitive to uncertainty in P . Varying P_j between 0.6 and 1.0, the range reported by Chen and Zhao et al.,³⁶ changed F by as much as 25% compared to our calculations that assumed a P_j value of 0.8 for all species. A P value of 0.11, the minimum reported by Williams et al.,⁴¹ reduced F by about 86%.

2.5 Discussion

2.5.1 Implications for Epidemiology

The results of this study suggest that variability in the fraction of ambient $\text{PM}_{2.5}$ that penetrates into and persists in indoor air is a possible contributor to the observed heterogeneity in PM-mediated health effect estimates. As hypothesized, lower values of F correspond to the conditions for which lower PM-mediated effect estimates have been observed. The “scrubbing” of nitrate from indoor-transported aerosol is a source of non-differential exposure misclassification when outdoor $\text{PM}_{2.5}$ concentrations are used as exposure surrogates. As a result, there is greater exposure misclassification in regions where nitrate is a large and variable fraction of the $\text{PM}_{2.5}$ mass. There are large geographic differences in $\text{PM}_{2.5}$ composition across the United States, with nitrate comprising a larger and more variable fraction of ambient $\text{PM}_{2.5}$ mass in the western United States compared to the eastern United States.⁴⁴ Thus, the lower F value calculated for the southwestern U.S., high nitrate scenario is consistent with the lower risk estimates for PM-associated morbidity/mortality in the western U.S.¹⁻⁴ Similarly, F was greatly reduced in these calculations by the addition of a term to account for filtration losses in air conditioning systems. This is consistent with the lower effect estimates observed for communities with a higher prevalence of central AC.^{2,20-22} Variability in air conditioning use with season and meteorological conditions may also contribute to temporal heterogeneity in effect estimates.

When F is larger and less variable, ambient $\text{PM}_{2.5}$ exposures are more similar to and more highly correlated with central-site $\text{PM}_{2.5}$ concentrations. As a result, biases not attributable to Berkson's fallacy are smaller. A higher F value was observed for the near

roadway scenario, which is consistent with several studies that report larger risks of adverse health outcomes when $\text{PM}_{2.5}$ is enriched in primary combustion tracers⁴⁻⁷ and with proximity to roadway.^{e.g. 8-11} While $\text{PM}_{2.5}$ components associated with primary combustion sources might be more toxic than others, these results suggest that fresh combustion particles also penetrate and persist indoors with higher efficiency than secondary sulfate due to their smaller depositional losses. Low socioeconomic status, a predictor of susceptibility to PM-related health effects,¹² affects F because low income residences tend to have higher AER values and are more likely to be in close proximity to busy roadways. The results of this study suggest that both factors enhance exposure to ambient $\text{PM}_{2.5}$. Note F values were as much as 22% greater at an AER of 0.90 h^{-1} compared to an AER of 0.45 h^{-1} for the scenarios studied. Disparity in exposure misclassification, in addition to factors such as access to health care, could contribute to differences in health effect estimates for these populations.

While the focus of this analysis was on spatial variability, parameters considered in this analysis also vary temporally. For example, AERs vary with meteorological conditions (e.g., wind, indoor-outdoor temperature difference) and home ventilation conditions (e.g., open windows versus closed windows). The physical and chemical properties of ambient $\text{PM}_{2.5}$ also vary temporally with variations in primary emissions, atmospheric stability, relative humidity, and photochemical activity. Certainly, AC use varies temporally with outdoor temperature. Thus, variability in F could contribute to exposure misclassification in time-series epidemiology, as well as to the observed geographic differences in PM-mediated health effect estimates.

2.5.2 Refined Exposure Surrogates

This study suggests that refined exposure surrogates that account for the outdoor-to-indoor transport of $\text{PM}_{2.5}$ could reduce exposure error in $\text{PM}_{2.5}$ epidemiology. Tools for modeling indoor concentrations of ambient $\text{PM}_{2.5}$ exist, are being refined, and are being incorporated into population exposure models such as the Stochastic Human Exposure and Dose (SHEDS) model.⁴⁵ Such models show promise for the development of exposure surrogates for epidemiology that account for modifications of ambient $\text{PM}_{2.5}$ with outdoor-to-indoor transport. This paper illustrates a method to predict the indoor concentration and composition of ambient $\text{PM}_{2.5}$ – a method that makes use of readily available data and is computationally inexpensive. It is proposed that this approach could be used to provide refined exposure surrogates for population-based epidemiologic analyses. In the following paragraphs, recommendations for the use of this method to generate refined exposure estimates are made and important refinements are discussed.

The sensitivity analysis presented above demonstrates the importance of accounting for species-specific losses in indoor air when predicting indoor concentrations of ambient $\text{PM}_{2.5}$. Treating ambient $\text{PM}_{2.5}$ as a single entity resulted in a substantial increase in F over the calculation that accounted for differences in depositional losses across species and phase changes of nitrate. Thus, the potential benefits from use of refined exposure surrogates may not be realized if F is calculated based on total $\text{PM}_{2.5}$ mass rather than $\text{PM}_{2.5}$ composition and species-resolved size distributions. Notably, particle species size distributions are not frequently measured. Most species size distribution measurements have been made during intensive sampling campaigns in remote locations for the purpose of visibility research. It is expected that broad generalizations can be made about species size distributions based on source proximity and atmospheric

chemistry. However, a systematic analysis backed by measurements has not been performed to support the application of the methods demonstrated here to other locations. An improved understanding of the variability of PM_{2.5} species size distributions would help facilitate the use of this model elsewhere.

The sensitivity analysis also illustrates the importance of accounting for phase changes of semivolatile species. F decreased by as much as 25% when the evaporative losses of nitrate were accounted for. A limitation of this method is that organic PM_{2.5} is treated as non-volatile, when in fact it is semi-volatile.⁴⁶ Sampling studies have demonstrated that phase changes of organics can impact F . Lunden et al.⁴³ concluded that a lower F for OC ($F=0.5$) than for EC ($F=0.61$) in an *unoccupied* home was due to evaporation of some particulate organic matter as organic gases sorbed to indoor surfaces. However, it has also been suggested that ambient organic matter shifts from the gas into the particle phase by sorption into indoor-generated PM in *occupied* homes.⁴⁷ This has been demonstrated for polycyclic aromatic hydrocarbons.⁴⁸ Accounting for phase changes of ambient organics with outdoor-to-indoor transport to further refine ambient PM_{2.5} exposure prediction is an area of future research.

Particle loss mechanisms are influenced by many factors in addition to particle size and composition, including home ventilation conditions, air flow characteristics, and building construction.^{16,23,2649,50} For example, P values are likely to differ for homes above and below the poverty line, with homes below the poverty line (which tend to be leakier)⁵¹ having higher P values. Thus, the disparity in F values for homes above and below the poverty line could be larger than demonstrated here. It should be noted that sampling from particle size distributions to generate distributions of k_{dep} , P , and P_{filter}

values (e.g. Monte Carlo methods) representative of each $\text{PM}_{2.5}$ species would better capture the distribution of $\text{PM}_{2.5}$ exposures over a wider range of conditions than the use of a single value for each $\text{PM}_{2.5}$ species, as done in this demonstration. This refinement is recommended when using the methods demonstrated here to generate exposure distributions. Distributions of these model parameters reflecting each $\text{PM}_{2.5}$ species size distribution (i.e. Figure 2-2) can be obtained using distributions of P , k_{dep} , and P_{filter} available for a range of particle sizes and under a variety of conditions.^{e.g. 16,26,50}

While AER values in this work were taken from the literature, AER distributions for a study population can be modeled. The Lawrence Berkeley National Laboratory (LBNL) infiltration model^{52,53} has the potential to provide AER estimates that could be used in mass balance models to generate refined estimates of exposure to ambient $\text{PM}_{2.5}$. Currently, the LBNL infiltration model predicts AER distributions for closed (i.e. windows and doors closed) housing stock using housing-characteristics data that are readily available from sources such as the Census and American Housing Survey and meteorological data available, for example, from the National Climate Data Center.^{51,52,54} The LBNL infiltration model must be adapted to account for natural ventilation (i.e. air flow through open windows/doors). This is an active area of research.^{e.g. 25} As is recommended for k_{dep} , P , and P_{filter} , sampling from the distribution of AERs generated with the LBNL Infiltration Model is recommended when calculating exposure distributions.

The method presented here does not account for exposure to ambient $\text{PM}_{2.5}$ in environments other than the home. In addition, the methods presented in this work are meant only to predict the indoor concentrations of ambient $\text{PM}_{2.5}$ in single family,

detached homes. While people do spend the majority of their time in their homes (near 70%), they also spend time in other indoor environments, outdoors and in transit.¹³ Exposure to ambient $PM_{2.5}$ in these environments also contributes to total ambient $PM_{2.5}$ exposure. Prediction of the concentrations of ambient $PM_{2.5}$ in these other environments and in multi-family residences is also needed.

2.6 Conclusions

Variability in the fraction of ambient $PM_{2.5}$ that penetrates and persists indoors can be substantial and may contribute to the heterogeneity in effect estimates of PM-related health outcomes. This work suggests that F is higher in close proximity to primary combustion aerosol (e.g. proximity to traffic) and with increased prevalence of poverty. F is lower when $PM_{2.5}$ is enriched in nitrate and with AC use. As a result, exposure error resulting from variability in F will be greatest when these factors have high temporal and/or spatial variability. Analyses of the relative sizes of the temporal and spatial errors in exposure estimates are needed in order to understand which types of epidemiological study designs are more, versus less affected by this type of exposure error. Certainly case-crossover designs, in which each subject acts as their own control, can avoid errors associated with location of residence, proximity to sources, socioeconomic status, and residential construction. The methods described here have the potential to reduce exposure misclassification for the study designs for which this variability is not inherently controlled for.

2.7 References

1. Dominici, F.; Peng, R. D.; Bell, M. L.; Pham, L.; McDermott, A.; Zeger S. L.; Samet, J. M., Fine particulate air pollution and hospital admissions for cardiovascular and respiratory diseases. *JAMA* **2006**, 295, 1127 - 1134.

2. Franklin, M.; Zeka, A.; Schwartz, J., Association between PM_{2.5} and all-cause and specific cause mortality in 27 US communities. *J. Exposure Sci. Environ. Epidemiol.* **2007**, 17, 279 - 287.
3. Bell, M. L.; Ebisu, K.; Peng, R. D.; Walker, J.; Samet, J. M.; Zeger, S. L.; Dominici, F., Seasonal and regional short-term effects of fine particles on hospital admissions in 202 US counties, 1999 -2008. *Am. J. Epidemiol.* **2008**, 168, 1301 - 1310.
4. Franklin, M.; Zeka, A.; Schwartz, J., The role of particle composition on the association between PM_{2.5} and mortality. *Epidemiology* **2008**, 19, 680 - 689.
5. Bell, M. L.; Ebisu, K.; Peng, R. D.; Samet, J. M.; Dominici, F., Hospital admissions and chemical composition of fine particle air pollution. *Am. J. Respir. Critic. Care Med.* **2009**, 179, 1115 - 1120.
6. Peng, R. D.; Bell, M. L.; Geyh, A. S.; McDermott, A.; Zeger, S. L.; Samet, J. M.; Dominici, F., Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. *Environ. Health Perspect.* **2009**, 117, 957 - 963.
7. Ito, K.; Mathes, R.; Ross, Z.; Nadas, A.; Thurston, G.; Matte, T., Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. *Environ. Health Perspect.* **2011**, 119, 467 - 473.
8. Hoek, G.; Brunekreef, B.; Goldbohm, S.; Fischer, P.; van den Brandt, P.A., Association between mortality and indicators of traffic-related air pollution in the Netherlands. *Lancet* **2002**, 360, 1203 - 1209.
9. Gauderman, W. J.; Avol, E.; Lurmann, F.; Kuenzli, N.; Gilliland, F.; Peters, J.; McConnell, R., Childhood asthma and exposure to traffic and nitrogen dioxide. *Epidemiology* **2005**, 16, 737 - 743.
10. McConnell, R.; Berhane, K.; Yao, L.; Jerrett, M.; Lurmann, F.; Gilliland, F.; Kunzli, N.; Gauderman, J.; Avol, E.; Thomas, D.; Peters, J., Traffic, susceptibility, and childhood asthma. *Environ. Health Perspect.* **2006**, 114, 766 - 772.
11. Brauer, M.; Lencar, C.; Tamburic, L.; Koehoorn, M.; Demers, P.; Karr, C., A cohort study of traffic-related air pollution impacts on birth outcomes. *Environ. Health Perspect.* **2008**, 116, 680 - 686.
12. Sacks, J. D.; Stanek, L. W.; Luben, T. J.; Johns, D. O.; Buckley, B. J.; Brown, J. S.; Ross, M., Particulate matter-induced health effects: who is susceptible? *Environ. Health Perspect.* **2011**, 119, 446 - 454.
13. Klepeis, N. E.; Nelson, W. C.; Ott, W. R.; Robinson, J. P.; Tsang, A. M.; Switzer, P.; Behar, J. V.; Stephen, C. H.; Engelmann, W. H., The National Human Activity Pattern

Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *J. Exposure Sci. Environ. Epidemiol.* **2001** 11, 231 - 252.

14. Long, C. M.; Suh, H. H.; Catalano, P. J.; Koutrakis, P., Using time- and size-resolved particulate data to quantify indoor penetration and deposition behavior. *Environ. Sci. Technol.* **2001**, 35, 2089 - 2099.

15. Meng, Q. Y.; Turpin, B. J.; Polidori, A.; Lee, J. H.; Weisel, C. P.; Morandi, M.; Winer, A.; Zhang, J., PM_{2.5} of ambient origin: estimates and exposure errors relevant to PM epidemiology. *Environ. Sci. Technol.* **2005**, 39, 5105 - 5112.

16. Nazaroff, W. W., Indoor particle dynamics. *Indoor Air* **2004**, 14, 175 - 183.

17. Dominici, F.; Zeger, S. L.; Samet, J. M., A measurement error model for time-series studies of air pollution and mortality. *Biostatistics* **2000**, 1, 157 - 175.

18. Zeger, S. L.; Thomas, D.; Dominici, F.; Samet, J. M.; Schwartz, J.; Dockery, D.; Cohen, A., Exposure measurement error in time-series studies of air pollution: Concepts and consequences. *Environ. Health Perspect.* **2000**, 108, 419 - 426.

19. U.S. EPA. Integrated Science Assessment for Particulate Matter (Final Report); U.S. Environmental Protection Agency (U.S. EPA), Washington, DC, EPA/600/R-08/139F, **2009**, 1-2228.

20. Janssen, N. A.; Schwartz, H. J.; Zanobetti, A.; Suh, H. H., Air conditioning and source-specific particles as modifiers of the effect of PM₁₀ on hospital admissions for heart and lung disease. *Environ. Health Perspect.* **2002**, 110, 43 - 49.

21. Zeka, A.; Zanobetti, A.; Schwartz, J., Short-term effects of particulate matter on cause specific mortality: effects of lags and modification by city-specific characteristics. *Occup. Environ. Med.* **2005** 62, 718 - 725.

22. Bell, M. L.; Ebisu, K.; Peng, R. D.; Dominici, F., Adverse health effects of particulate air pollution: modification by air conditioning. *Epidemiology* **2009**, 19, 680 - 689.

23. Thornburg, J.W.; Ensor, D. S.; Rodes, C. E.; Lawless, P. A.; Sparks, L. E.; Mosely, R. B., Penetration of particles into buildings and associated physical factors, Part I: model development and computer simulations. *Aerosol Sci. Technol.* **2001**, 34, 284 - 296.

24. Waring, M. S.; Siegel, J. A., Particle loading rates for HVAC filters, heat exchangers, and ducts. *Indoor Air* **2008**, 18, 209 - 224.

25. Breen, M. S.; Breen, M.; Williams, R. W.; Schultz, B. D., Predicting residential air exchange rates from questionnaires and meteorology: model evaluation in central North Carolina. *Environ. Sci. Technol.* **2010**, 44, 9349 - 9356.
26. Riley, W. J.; McKone, T. E.; Lai, A. C. K.; Nazaroff, W. W., Indoor particulate matter of outdoor origin: importance of size-dependent removal mechanisms. *Environ. Sci. Technol.* **2002**, 36, 200 - 207.
27. Meng, Q. Y.; Turpin, B. J.; Lee, J. H.; Polidori, A.; Weisel, C. P.; Morandi, M.; Colome, S.; Zhang, J.; Stock, T.; Winer, A., How does infiltration behavior modify the composition of ambient PM_{2.5} in indoor spaces? An analysis of RIOPA data. *Environ. Sci. Technol.* **2007**, 41, 7315 - 7321.
28. Lunden, M. M.; Thatcher, T. L.; Hering, S. V.; Brown, N. J., The use of time- and chemically-resolved particulate data to characterize the infiltration of outdoor PM_{2.5} into a residence in the San Joaquin Valley. *Environ. Sci. Technol.* **2003**, 37, 4724 -4732.
29. Hering, S. V.; Lunden, M. M.; Thatcher, T. L.; Kirchstetter, T. W.; Brown, N. J., Using regional data and building leakage to assess indoor concentrations of particles of outdoor origin. *Aerosol Sci. Technol.* **2007**, 41, 639 - 654.
30. Polidori, A.; Turpin, B.; Meng, Q. Y.; Lee, J. H.; Weisel, C.; Morandi, M.; Colome, S.; Stock, T.; Winer, A.; Zhang, J.; Kwon, J.; Alimokhtari, S.; Shendell, D.; Jones, J.; Farrar, C.; Maberti, S., Fine organic particulate matter dominates indoor-generated PM_{2.5} in RIOPA homes. *J. Exposure Sci. Environ. Epidemiol.* **2006**, 16, 321 - 331.
31. Drewnick, F.; Jayne, J. T.; Canagaratna, M.; Worsnop, D. R.; Demerjian, K. L., Measurement of ambient aerosol composition during the PMTACS-NY 2001 using an aerosol mass spectrometer. Part II: chemically speciated mass distributions. *Aerosol Sci. Technol.* **2004**, 38, 104 - 117.
32. Ge, X. L.; Zhang, Q.; Sun, Y. L.; Ruchi, C. R.; Setyan, A., Impacts of aqueous-phase processing on aerosol chemistry and size distributions in Fresno, California, during wintertime. *Environ. Chem.* **2012**, 221 - 235.
33. Lena, T. S.; Ochieng, S.; Carter, M.; Holguin-Veras, J.; Kinney, P. L., Elemental carbon and PM_{2.5} levels in an urban community heavily impacted by truck traffic. *Environ. Health Perspect.* **2002**, 110, 1009 - 1015.
34. Lunden, M. M.; Revzan, K. L.; Fischer, M. L.; Thatcher, T. L.; Littlejohn, D.; Hering, S. C.; Brown, N. J., The transformation of outdoor ammonium nitrate aerosols in the indoor environment. *Atmos. Environ.* **2003**, 37, 5633 - 5644.

35. DeCarlo, P.; Slowik, J. G.; Worsnop, D. R.; Davidovits, P.; Jimenez, J. L., Particle morphology and density characterization by combined mobility and aerodynamic diameter measurements. Part 1: theory. *Aerosol Sci. Technol.* **2004**, 38, 1185 - 1205.
36. Chen, C.; Zhao, B., Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. *Atmos. Environ.* **2011**, 45, 275 - 288.
37. US EPA.. Exposure Factors Handbook (Final Report). US Environmental Protection Agency: Washington, DC, EPA/600/P-95/002F a-c, **1997**.
38. John, W.; Wall, S. M.; Ondo, J. L.; Winklmayr, W., Modes in the size distributions of atmospheric inorganic aerosol. *Atmos. Environ.* **1990**, 24A, 2349 - 2359.
39. Pandis, S.; Davidson, C.; Robinson, A.; NARSTO EPA_SS_PITTSBURGH Gas Concentrations and PM Physical Properties Data. 2007. Available on-line (http://eosweb.larc.nasa.gov/PRODOCS/narsto/table_narsto.html) at the Atmospheric Science Data Center at NASA Langley Research Center, Hampton, Virginia, USA **2007**.
40. Liu, D. L.; Nazaroff, W. W., Modeling particle penetration across building envelopes. *Atmos. Environ.* **2001**, 35, 4451 - 4462.
41. Williams, R.; Suggs, J.; Rea, A.; Sheldon, R.; Rodes, C.; Thornburg, J., The Research Triangle Park particulate matter panel study: modeling ambient source contribution to personal and residential PM mass concentrations. *Atmos. Environ.* **2003**, 37, 5365 - 5378.
42. Sarnat, S. E.; Coull, B. A.; Ruiz, P. A.; Koutrakis, P.; Suh, H. H., The influences of ambient particle composition and size on particle infiltration in Los Angeles, CA, Residences. *J. Air Waste Manage. Assoc.* **2006**, 56, 186 - 196.
43. Lunden, M. M.; Kirchstetter, T. W.; Thatcher, T. L.; Hering, S. V.; Brown, N. J., Factors affecting the indoor concentration of carbonaceous aerosols of outdoor origin. *Atmos. Environ.* **2008**, 42, 5660 - 5671.
44. NARSTO.. Particulate Matter Science for Policy Makers: A NARSTO Assessment, P. McMurry, M. Shepherd, and J. Vickery, (eds). Cambridge University Press: Cambridge, England, **2004**.
45. Burke, J. M.; Zufall, M. J.; Ozkaynak, H., A population exposure model for particulate matter: case study results for PM_{2.5} in Philadelphia, PA. *J. Exposure Anal. Environ. Epidemiol.* **2001**, 11, 470 - 489.
46. Lipsky, E. M.; Robinson, A. L., Effects of dilution on fine particle mass and partitioning of semivolatile organics in diesel exhaust and wood smoke. *Environ. Sci. Technol.* **2006**, 40, 155 - 162.

47. Polidori, A.; Kwon, J.; Turpin, B. J.; Weisel, C., Source proximity and residential outdoor concentrations of PM_{2.5}, OC, EC, and PAHs. *J. Exposure Sci. Environ. Epidemiol.* **2009**, 20, 457 - 468.
48. Naumova, Y. Y.; Offenberg, J. H.; Eisenreich, S. J.; Meng, Q. Y.; Polidori, A.; Turpin B. J.; Weisel, C. P.; Morandi, M. T.; Colome, S. D.; Stock, T. H.; Winer, A. M.; Alimokhtari, S.; Kwon, J.; Maberti, S.; Shendell, D.; Jones, J.; Farrar, C., Gas/particle distribution of polycyclic aromatic hydrocarbons in coupled indoor/outdoor atmospheres. *Atmos. Environ.* **2003**, 37, 703 - 719.
49. Thatcher, T. L.; Lunden, M. M.; Revzan, K. L.; Sextro, R. G.; Brown, N. J., A concentration rebound method for measuring particle penetration and deposition in the indoor environment. *Aerosol Sci. Technol.* **2003**, 37, 847 - 864.
50. Wallace, L.; Emmerich, S. J.; Howard-Reed, C., Effect of central fans and in-duct filters on deposition rates of ultrafine and fine particles in an occupied townhouse. *Atmos. Environ.* **2004**, 38, 405 - 413.
51. Chan, W. R.; Nazaroff, W. W.; Price, P. N.; Sohn, M. D.; Gadgil, A. J., Analyzing a database of residential air leakage in the United States. *Atmos. Environ.* **2005**, 39, 3445 - 3455.
52. Sherman, M. H.; Grimsrud, D. T., Measurement of infiltration using fan pressurization and weather data. Lawrence Berkeley National Laboratory Report. LBNL-10852, Berkeley, CA, **1980**.
53. Sherman, M. H.; Dickerhoff, D. J., Airtightness of US dwellings. *ASHRAE Transactions* **1998**, 104, 1359 - 1367.
54. Persily, A.; Musser, A.; Emmerich, S. J., Modeled infiltration rate distributions for US housing. *Indoor Air* **2010**, 20, 473 - 485.

	Sulfate	Nitrate	Elemental Carbon	Organic Carbon	Soil
Queens, NY					
MMD ¹ (μm)	0.5	0.5	0.07	0.07, 0.4	2.0
kdep ² (h ⁻¹)	0.09	0.09	0.06	0.06, 0.07	0.58
Pfilter ²	0.60	0.60	0.85	0.85, 0.70	0.00
Fresno, CA					
MMD ³ (μm)	0.5	0.4	0.07	0.3	2.0
kdep ² (h ⁻¹)	0.09	0.07	0.06	0.06	0.58
Pfilter ²	0.60	0.70	0.85	0.78	0.00

¹ Drewnick et al. (2004)

² Riley et al. (2002)

³ Ge et al. (2012)

Table 2-1. Mass median diameter (MMD) and associated k_{dep} and P_{filter} values used for particulate soil, sulfate, nitrate, elemental carbon, and organic carbon. Values of k_{dep} and P_{filter} are from Riley et al. ²⁶

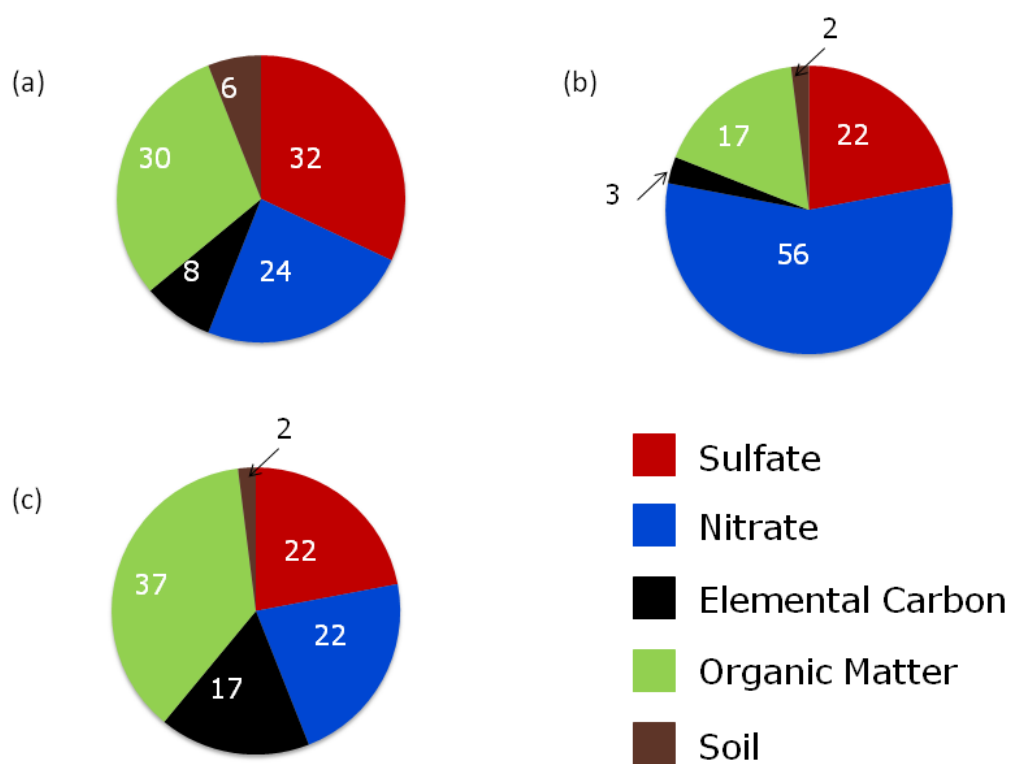


Figure 2-1. Three PM_{2.5} composition scenarios. (a) “northeastern mean composition”. (b) “southwestern high nitrate” composition (c) “near roadway” composition. Scenario (d) uses the “northeastern mean composition” with air conditioning.

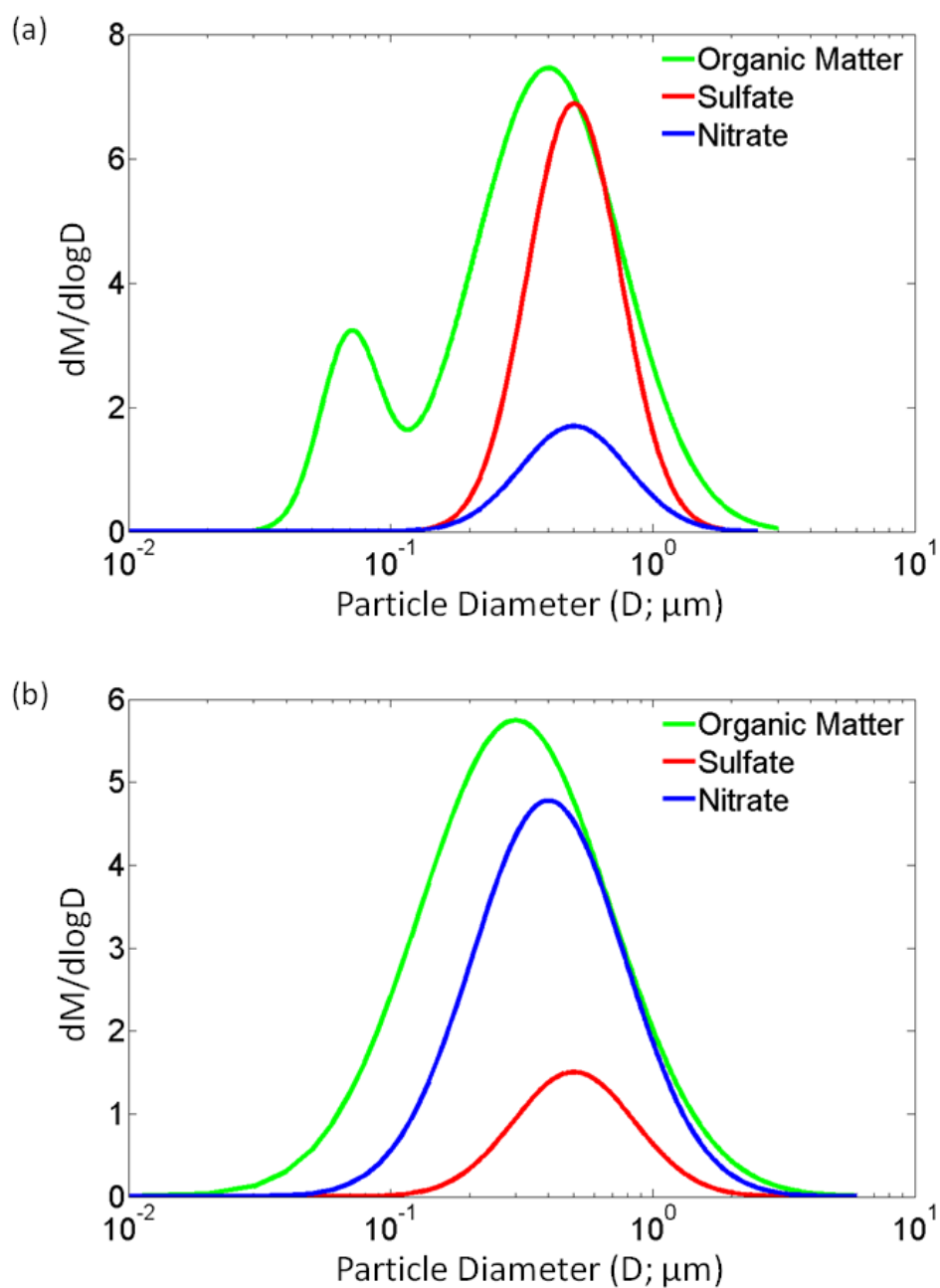


Figure 2-2. Most frequently observed size distributions for major PM_{2.5} species (sulfate, nitrate, organic matter). (a) in Queens, NY in August 2001 (Drewnick et al., 2004) and (b) Fresno, CA in January 2010 (Ge et al., 2012).

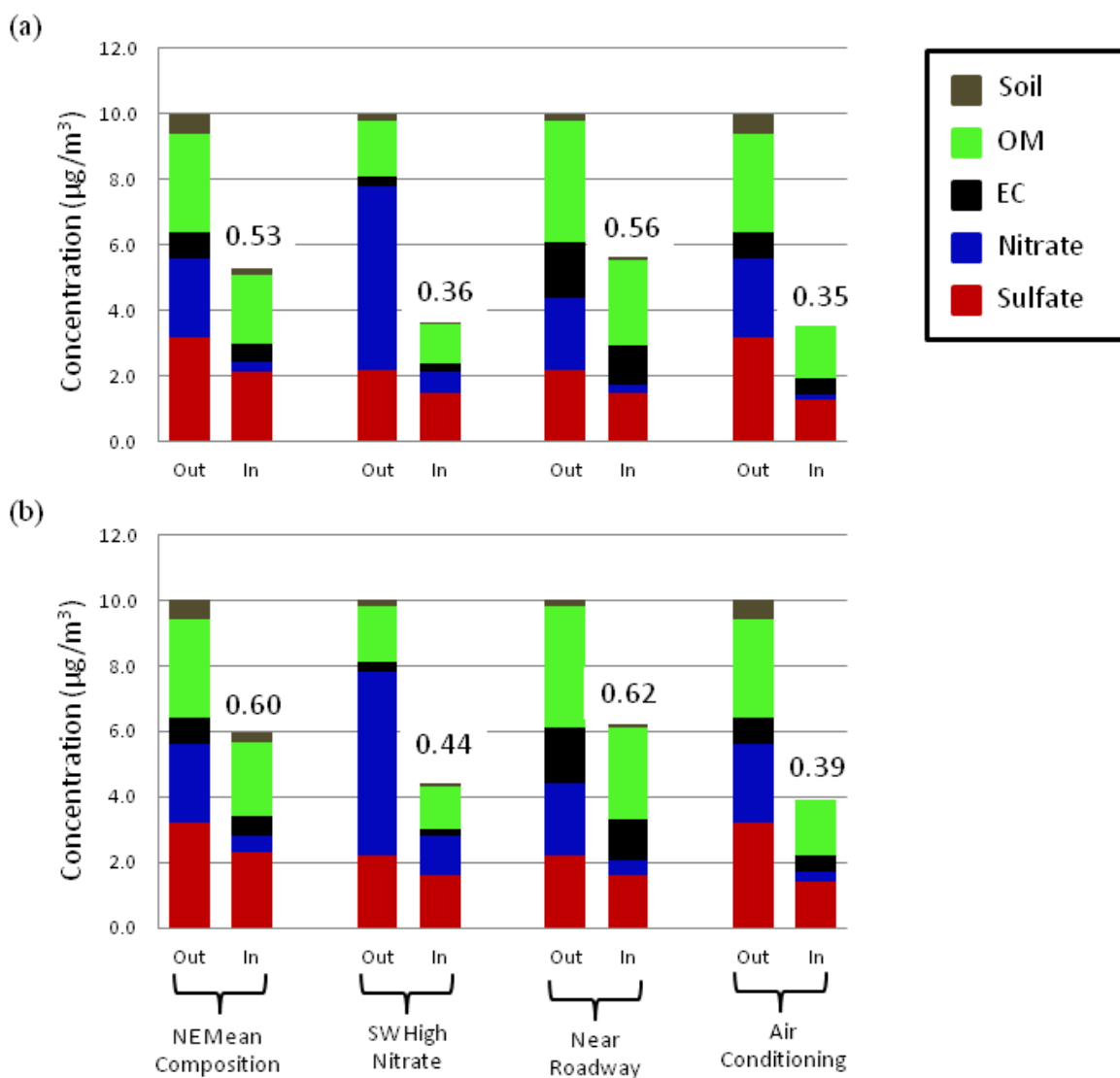


Figure 2-3. Outdoor and indoor concentrations of ambient $\text{PM}_{2.5}$ and the fraction of outdoor $\text{PM}_{2.5}$ that penetrates and persists indoors (F) for the scenarios described in Figure 2-1 and Table 2-1. Note the “NE mean composition” and “air conditioning” scenarios have the same composition and species-specific size distributions. Data labels are F values for each scenario. (a) typical residential U.S. air exchange rate of 0.45 h^{-1} and (b) air exchange rate for low income residence of 0.90 h^{-1} .³⁷

Chapter 3. Refined Ambient PM_{2.5} Exposure Surrogates and the Risk of Myocardial Infarction

Material in this chapter has been published previously as:

Hodas, N.; Turpin, B. J.; Lunden, M. M.; Baxter, L. K.; Özkaynak, H.; Burke, J.; Ohman-Strickland, P.; Thevenet-Morrison, K.; Rich, D. Q., Refined ambient PM_{2.5} exposure surrogates and the risk of myocardial infarction. *J. Exposure Sci. Environ. Epidemiol.* **2013**, 23, 573 - 580.

3.1 Abstract

Using a case-crossover study design and conditional logistic regression, the relative odds of transmural (full-wall) myocardial infarction (MI) calculated using exposure surrogates that account for human activity patterns and the indoor transport of ambient PM_{2.5} were compared with those calculated using central-site PM_{2.5} concentrations to estimate exposure to PM_{2.5} of outdoor origin (exposure to *ambient* PM_{2.5}). Because variability in human activity and indoor PM_{2.5} transport contributes exposure error in epidemiologic analyses when central-site concentrations are used as exposure surrogates, surrogates that account for this variability are referred to as "refined" surrogates. As an alternative analysis, whether the relative odds of transmural MI associated with increases in ambient PM_{2.5} is modified by residential air exchange rate (AER), a variable that influences the fraction of ambient PM_{2.5} that penetrates and persists indoors, was evaluated. Use of refined exposure surrogates did not result in larger health effect estimates (ORs = 1.10 - 1.11 with each interquartile range increase.), narrower confidence intervals, or better model fits compared to the analysis that used central-site PM_{2.5}. Evidence for heterogeneity in the relative odds of transmural MI with

residential AER (effect-modification) was observed, with residents of homes with higher AERs having larger ORs than homes in lower AER tertiles. For the level of exposure-estimate refinement considered here, these findings add support to the use of central-site PM_{2.5} concentrations for epidemiological studies that employ similar case-crossover study designs. In such designs, each subject serves as his or her own matched control. Thus, exposure error related to factors that vary spatially or across subjects should only minimally impact effect estimates. These findings also illustrate that variability in factors that influence the fraction of ambient PM_{2.5} in indoor air (e.g., AER) could possibly bias health effects estimates in study designs for which a spatio-temporal comparison of exposure effects across subjects is conducted.

3.2 Introduction

A recent meta-analysis, which reported a statistically significant 2.5% increase in the risk of myocardial infarction (MI) associated with each 10 µg/m³ increase in ambient (outdoor-generated) PM_{2.5} concentration lagged one day, concluded that acute increases in PM_{2.5} may trigger MI.¹ A previous study, which was included in this meta-analysis, reported an increased risk of transmural (full wall) MI, but not non-transmural (subendocardial) MI, associated with increased PM_{2.5} concentration in the 24 hours before emergency department admission for that infarction.² The work presented here builds on this "initial" analysis. In all of these studies, PM_{2.5} measured at one or more nearby (within 10 km) central-site monitors was used as a proxy for a subject's exposure to PM_{2.5} of outdoor origin (i.e., exposure to *ambient* PM_{2.5}). This likely resulted in exposure error due, in part, to proximity to local sources, human activity patterns (e.g., time spent in various locations) and temporal and spatial variability in the efficiency with

which ambient $PM_{2.5}$ penetrates into and persists in the indoor environment. While other air pollution studies have explored exposure refinements that account for spatial variability in ambient $PM_{2.5}$ due to local sources,^{e.g. 3-6} the variable effects of human activity patterns and ambient $PM_{2.5}$ losses with outdoor-to-indoor transport are largely unexplored.

The fraction of ambient $PM_{2.5}$ that penetrates and persists indoors (F) varies with multiple factors including particle size and chemical composition, housing characteristics (e.g., home age), meteorological conditions (e.g., wind speed and temperature),^{7,8} and human activities (e.g., opening windows or using air conditioning).⁹ Variability in the time spent in various locations (e.g., outdoors, indoors, or in a vehicle) also influences personal exposure to $PM_{2.5}$ of outdoor origin due to spatial variability in both outdoor $PM_{2.5}$ concentrations and the indoor transport of ambient $PM_{2.5}$. This exposure error is likely a combination of Berkson and classical errors, which would bias effect estimates towards the null and/or inflate variances,^{10,11} hampering the detection of statistically significant associations between increased ambient $PM_{2.5}$ exposures and the risk of MI. Therefore, ambient $PM_{2.5}$ exposure surrogates that account for these factors could offer improvement over the direct use of central-site monitor $PM_{2.5}$ concentrations in air pollution epidemiology studies.

Exposure errors associated with variability in F and human activity patterns may modify ambient-PM-mediated health effect estimates. Multiple studies have reported a lower risk of morbidity or mortality associated with increases in $PM_{2.5}$ concentration in communities with a high prevalence of central air conditioning (AC), compared to risk estimates among communities with lower AC prevalence.¹²⁻¹⁶ Central AC use reduces F

because indoor air is filtered as it is re-circulated, thus increasing particle losses indoors.¹⁷⁻¹⁹ In Chapter 2, it is reported that conditions resulting in lower calculated values of F (due to spatial variability in $PM_{2.5}$ composition and/or residential AER) corresponded to circumstances under which lower effect estimates had been observed in previous epidemiological studies. It was concluded that exposure misclassification due to variability in F could partially explain this observed geographic heterogeneity in ambient-PM-mediated health effect estimates.²⁰

Using a case-crossover study design, herein the relative odds of transmural MI associated with increased ambient $PM_{2.5}$ exposure in the previous 24 hours were estimated using three different $PM_{2.5}$ exposure metrics that account for variability in human activity patterns and/or the indoor transport of ambient $PM_{2.5}$: (a) a stochastic human exposure model that simulates the ambient $PM_{2.5}$ concentration and time spent in each of several locations (i.e. outdoors, indoors, in a vehicle) to estimate population distributions of ambient $PM_{2.5}$ exposure, (b) a deterministic mass-balance model that estimates residential, indoor concentrations of ambient (outdoor-generated) $PM_{2.5}$ using a more refined treatment of residential air exchange rates (AERs) and $PM_{2.5}$ penetration and losses with indoor transport, and (c) a hybrid of these two models.²¹ As noted above, variability in human activity patterns and the indoor transport of ambient $PM_{2.5}$ can contribute to exposure error in epidemiologic analyses when central-site concentrations alone are used to estimate exposure to ambient $PM_{2.5}$ and, thus, the exposure surrogates that account for this variability are referred to as "refined" exposure surrogates in the following text. It was hypothesized that these refined ambient $PM_{2.5}$ exposure surrogates would have less non-differential exposure error (which tends to bias effect estimates

towards the null) and, thus, would result in larger health effect estimates, narrower confidence intervals, and better model fits compared to the analysis that used central-site $PM_{2.5}$ concentrations alone as surrogates for ambient $PM_{2.5}$ exposures. As an alternative analysis, whether the association between ambient $PM_{2.5}$ and transmural MI is modified by residential AER was evaluated. For this analysis, it was hypothesized that effect estimates would be smaller for low AERs because a smaller fraction of ambient $PM_{2.5}$ penetrates and persists indoors. Thus, at low AERs, the difference between central-site $PM_{2.5}$ concentrations and actual ambient $PM_{2.5}$ exposure is greater, resulting in proportionally more non-differential exposure misclassification and larger bias towards the null (i.e. greater underestimation of effect).

3.3 Methods

3.3.1 Study Population and Outcome Definition

The study population and definition of transmural infarction used in this study have been described previously.² Briefly, all unscheduled hospital admissions with a primary diagnosis of acute myocardial infarction (International Classification of Diseases 9th Revision [ICD-9] code 410.01, 410.11, 410.21, 410.31, 410.41, 410.51, 410.61, 410.71, 410.81, 410.91) were extracted from the Myocardial Infarction Data Acquisition System (MIDAS), a New-Jersey-wide database of hospital discharges and death certificate registrations.^{22,23} Only those patients who were admitted between January 2004 and December 2006, were ≥ 18 years of age, were residents of New Jersey at the time of their MI, and had no previous diagnosis of MI were included. These subjects ($n = 1563$) were primarily male (63%) and white (69%) and had a median age of 62. Only subjects who resided within 10 km of a central-site monitor at the time of their MI were

included in this study. This study was approved by the Institutional Review Boards of the University of Medicine and Dentistry of New Jersey and Rutgers, The State University of New Jersey and the University of Rochester Research Subjects Review Board. MIDAS was also approved by the New Jersey Department of Health and Senior Services Institutional Review Board.

3.3.2 Exposure Surrogates

Four different exposure surrogates generated from central-site monitor concentrations were used to estimate personal exposure to $PM_{2.5}$ of outdoor origin (i.e., exposure to ambient $PM_{2.5}$). Because a significantly increased relative odds of transmural MI associated with average $PM_{2.5}$ concentrations in the 24 hours preceding emergency department admission was observed in the initial analysis,² here, hourly ambient $PM_{2.5}$ exposures were computed and averaged over that 24 hour period for each exposure metric. Detailed descriptions of each exposure surrogate and comparisons between them are available elsewhere.²¹ In the following paragraphs, a brief description of each exposure metric is provided. The exposure surrogates are labeled based on their level of refinement and complexity, with higher-numbered Tiers corresponding to a greater degree of refinement.

Tier 1. Central-site $PM_{2.5}$ Concentrations. For Tier 1, hourly ambient $PM_{2.5}$ concentrations for the study period (January 2004 - December 2006) measured at 7 New Jersey Department of Environmental Protection monitors were retrieved from the United States Environmental Protection Agency website.²⁴ The zip code of each patient's residence at the time of MI was extracted from MIDAS and subjects were assigned 24 hour average $PM_{2.5}$ concentrations, for all case and control periods, from the monitor

closest to their residence.² Tier 1 exposure estimates varied temporally within and across central-site-monitor regions with ambient $PM_{2.5}$ concentrations. Because subjects residing within 10 km of the same monitor were assigned the same exposure value for a given 24-hour case or control period, there was no geographic variability in exposure estimates within that 10 km radius. Within a given case or control period, however, exposure estimates did vary across monitoring locations.

Tier 2a. SHEDS – In Tier 2a, the exposure-modifying effects of human activity patterns and the indoor transport of ambient $PM_{2.5}$ were taken into account using the Stochastic Human Exposure and Dose Simulation (SHEDS) model.²⁵ Distributions of ambient $PM_{2.5}$ exposures were generated for a simulated population representative of the study population. For each census tract within 10 km of a central-site monitor, 10,000 representative individuals were simulated by sampling from census-tract level demographic data (gender, age, and employment status) from the 2000 U.S. Census. For each simulated individual, a time series of human activity patterns was simulated using diary data from the Consolidated Human Activity Database²⁶ matched by age, gender, season, and day of week. Hourly central-site $PM_{2.5}$ concentrations (Tier 1) were used as inputs, and personal exposure to $PM_{2.5}$ of outdoor origin was calculated as a time-weighted average of the *ambient* $PM_{2.5}$ concentrations in each microenvironment (e.g. home, office, outdoors). Note, indoor $PM_{2.5}$ sources were set to zero to estimate the distribution of exposures to $PM_{2.5}$ of outdoor origin only (i.e., ambient $PM_{2.5}$ exposures) in each census tract. For residential microenvironments, SHEDS sampled from a representative distribution of housing types, AERs, particle penetration efficiencies, and indoor particle deposition rates (Appendices B1, B2). It should be emphasized that the

AERs used in this version of SHEDS vary seasonally, but not spatially within the study domain. From the distribution of ambient $PM_{2.5}$ exposures generated for each hour during the study period, the median was used to calculate 24 hour mean exposures for each case and control period. The 24 hour mean exposures calculated for each census tract were then averaged over the 10 km region surrounding each central-site monitor.

Tier 2b. The Aerosol Penetration and Persistence Model – Hourly $PM_{2.5}$ concentrations measured at the central-site monitors (Tier 1) were modified to account for the effects of outdoor-to-indoor transport using the Aerosol Penetration and Persistence (APP) model^{8,27,28} and the Lawrence Berkeley National Laboratory (LBNL) Infiltration model.^{29,30} The APP model is a deterministic mass balance model that predicts the indoor concentration of ambient $PM_{2.5}$ based on AER, outdoor $PM_{2.5}$ concentrations, the efficiency of particle penetration into the home, the rate of depositional losses in indoor air, and, for ammonium nitrate, phase changes in the indoor environment.^{8,27,28} The equations that form the APP model are provided in Chapter 2 (Equations 1a and 1b). In addition to accounting for the semi-volatile nature of ammonium nitrate, daily variations in particle chemical composition were taken into account through the use of particle-size-resolved deposition loss rates specific to the size distributions of the major $PM_{2.5}$ species (sulfate, nitrate, elemental carbon, and organic carbon; Appendix B3). Central-site $PM_{2.5}$ composition data from the EPA Speciation Trends Network (STN) is available for every third day and was downloaded from the US EPA website for this purpose.²⁴ For days without measurements, $PM_{2.5}$ species mass fractions were interpolated using a weighted average of the two nearest mass fraction measurements. Subjects were excluded if there was a period of more than nine days

between STN measurements for the case period or all control periods. Because speciation measurements were not available for all central-site-monitor locations, values for the New Brunswick monitoring station, which were most highly correlated with data from other monitors across the state (Appendix B4), were used. With this approach, particle losses indoors varied daily with variations in $PM_{2.5}$ composition. Note, however, that deposition loss rates did not vary spatially in this work.

AERs calculated with the LBNL Infiltration model, which was modified to include air flow through open windows (Appendix B), were used as inputs to the APP model. The LBNL infiltration model predicts AER for single-family homes based on normalized leakage rates (which describe the effective area of openings in the building shell through which air can flow, normalized by home floor area and a parameter accounting for building height and validated against measurements in 70,000 closed homes) and meteorological conditions.^{29,30} Meteorological data were gathered from four airports in New Jersey (Newark, Caldwell, Somerset, and Trenton) and subjects were assigned the weather data from the monitor nearest their residence at the time of MI. The normalized leakage area was calculated using a model resulting from a statistical analysis relating leakage to housing characteristics (home age, floor area)³¹ using census-tract level housing data from the 2000 U.S. Census and the American Housing Survey. Notably, the model used to calculate normalized leakage rate differs for homes above and below the poverty line because home leakiness varies with resident poverty status, with low-income homes tending to be leakier.³¹ Thus, variations in calculated AERs arise from temporal and spatial variability in meteorological conditions and with spatial variability in housing stock. Unlike Tier 2a, detailed human activity patterns are not accounted for in

this metric, but Tier 2b provides a more refined treatment of residential AER and $PM_{2.5}$ penetration and losses with indoor transport. Census-tract-level ambient $PM_{2.5}$ exposures were averaged over the 10 km area around each central-site monitor.

Tier 3. SHEDS and APP Hybrid - The final exposure metric combined the refined treatment of human activity patterns from Tier 2a, with the more temporally- and spatially-resolved estimates of residential AER from Tier 2b (but without variations in $PM_{2.5}$ deposition rates with variations in $PM_{2.5}$ composition). $PM_{2.5}$ exposures were estimated with SHEDS as described above, but using residential AERs estimated with the LBNL Infiltration model.

3.3.3 Statistical Analyses

Study Design: For each ambient $PM_{2.5}$ exposure surrogate (tier), the same time-stratified case-crossover design^{32,33} as in the initial analysis² was used to estimate the relative odds of a transmural infarction associated with increased exposure in the previous 24 hours. In this design, each patient contributed information both as a case during the period immediately before the MI, and as a matched control during times when a MI did not occur. Since each subject serves as their own control, factors that differ only across subjects are controlled by design. Case periods were defined as the 24 hour period before emergency department admission for MI. Control periods (3-4 per case depending on the number of days in the calendar month), defined as 24 hour periods in which no MI occurred, were matched to the case period by day of the week, time of day, year, and calendar month. Central-site $PM_{2.5}$ concentrations (Tier 1) and modeled ambient $PM_{2.5}$ exposures (Tier 2a, 2b, 3) corresponding to these case and control periods were then contrasted in the statistical analyses.

Modeled Exposure Tier Analyses: The same conditional logistic regression model as in the initial analysis,² stratified by study subject, was used to examine the multiplicative interaction between ambient PM_{2.5} exposure and transmural MI:

$$\log \left(\frac{P(Y_{ij}=1)}{1-P(Y_{ij}=1)} \right) = \alpha_i + PM_{ij} + f(Temp_{ij}; \gamma) \quad (1)$$

Case-control status (i.e., case period = 1, control period = 0) was regressed against the mean estimated ambient PM_{2.5} exposure in the 24 hour period before emergency department admission for the index infarction or the corresponding control period. A natural spline (3 degrees of freedom) of the mean apparent temperature,^{34,35} from the same 24 hour period, was also included to estimate each subjects' perceived ambient air temperature. Hourly temperature and relative humidity data used to calculate apparent temperature were gathered from the same airports as the data used to calculate AER in Tiers 2b and 3. The relative odds of transmural MI was estimated using each exposure surrogate (Tier 1, 2a, 2b, or 3) scaled to the Tier-specific interquartile range (IQR) increase in the ambient PM_{2.5} exposure. For each Tier, the odds ratio (OR), its 95% confidence interval, and its Akaike's Information Criterion (AIC) value, which was used to compare the fit of these non-nested models to Tier 1, are presented.

Whether the refined exposure estimates (Tiers 2a, 2b, 3) added explanatory power over the Tier 1 estimate was also examined. In other words, whether the refined exposure estimates provided supplementary exposure information beyond that accounted for in the Tier 1 estimates and whether including that information in effect-estimate calculations resulted in additional MI risk over that associated with the central-site PM_{2.5} concentrations (Tier 1 estimates) alone was evaluated. For each case and control time period, the Tier 1 exposure estimate and each of the refined exposure estimates were

converted to z-scores based on their respective means and standard deviations. The conditional logistic regression model described above was run again with the Tier 1 z-score and the z-score difference (e.g., the difference between the Tier 1 z-score and the Tier 2a z-score) as covariates. Z-scores were used in order to create scale- and location-invariant versions of the exposure metrics. Given that variables that differ only by scale and location may contribute equivalently to explaining a response in the context of linear modeling, entering the difference (between the refined and the original z-scores) into a linear model in addition to the original represents the additional contribution that the refined variable can make over the original in explaining the response in a linear model. The regression coefficient for the Tier 1 z-score, times the observed IQR, estimated the increase in log-odds of transmural infarction associated with each IQR increase in the Tier 1 $PM_{2.5}$ concentration, while the regression coefficient for the “z-score difference” provided an estimate of the additional increase in log-odds of a transmural infarction associated with each IQR increase in the refined $PM_{2.5}$ exposure estimate, independent of the Tier 1 $PM_{2.5}$ concentration. A significance test of the “z-score difference” regression coefficient provides a test of whether the refined Tier adds any statistically significant relative odds beyond what is provided by Tier 1. This same model was run separately for each refined metric (Tiers 2a, 2b, and 3).

AER Effect Modification Analyses: Whether residential AER alone, without the other components contributing to the refined exposure surrogates, modified the association between the Tier 1 exposure surrogate and transmural infarction was also explored. This was done because AER estimates have smaller uncertainties than the more expansive exposure models and are important predictors of the fraction of ambient $PM_{2.5}$

that penetrates and persists indoors. However, as explained below, this approach also differs from the main analysis in that it introduces a spatial comparison.

AERs from the Tier 2b exposure estimates were ranked into tertiles (high AER, middle AER, and low AER). See Appendix B7 for summary statistics of AERs in each tertile. The Tier 1 conditional logistic regression analysis was then re-run adding two interaction terms to the model, as well as indicator variables for AER. The base model is

$$\log \left(\frac{P(Y_{ij}=1)}{1-P(Y_{ij}=1)} \right) = \alpha_i + \beta_1 AER_{low,i} + \beta_2 AER_{mid,i} + \beta_3 AER_{low,i} \times PM_{ij} + \beta_4 AER_{mid,i} \times PM_{ij} + \beta_5 AER_{high,i} \times PM_{ij} + f(Temp_{ij}; \gamma) \quad (2)$$

where Y_{ij} equals one if the j^{th} period for the i^{th} subject is a case and zero if control. Further, $AER_{low,i}$, $AER_{mid,i}$ and $AER_{high,i}$ are indicator variables equal to one if subject i has a low, middle or high AER and zero otherwise. The term $f(Temp_{ij}; \gamma)$ represents the natural spline that is added to adjust for apparent temperature and α_i represents the sum of a random intercept for subject i as well as any between-subject variables. Upon conditioning on subject, α_i becomes a nuisance parameter which cancels out of the conditional logistic likelihood and is not estimated. From this model, the relative odds of a transmural infarction and its 95% confidence interval associated with an $10.3 \mu\text{g}/\text{m}^3$ (IQR) increase in Tier 1 $\text{PM}_{2.5}$ concentration was estimated within each tertile of AER. This was done for the cool (November to April) and warm (May to October) seasons separately because $\text{PM}_{2.5}$ concentrations and composition are distinctly different over these two periods (Appendix B8).³⁶

In this alternative analysis using interaction terms to estimate the relative odds of a transmural MI associated with increased $\text{PM}_{2.5}$ concentration within the low, middle, and high AER groups, the case-crossover analysis described above was essentially

stratified by modeled residential AER and estimates of the relative odds of transmural MI across AER tertiles were compared. In contrast with the “Modeled Exposure Tier Analysis”, which was strictly a within-subject, temporal analysis, this “AER Effect Modification Analysis” is a spatio-temporal comparison of exposure effects across AER tertiles and, thus, across subjects. This analysis also focused on a single parameter that influences the indoor transport of ambient PM_{2.5} in order to reduce the number of assumptions and associated uncertainty in comparison to the more complicated refined exposure surrogates explored above.

To evaluate whether spatially varying factors in addition to AER (e.g. PM_{2.5} chemical composition, study population characteristics) could contribute to variability in relative risk of MI across AER tertiles, a case-crossover analysis stratified by monitoring-site community was also conducted and study population characteristics were compared across AER tertiles. All data sets were constructed using SAS software (version 9.1.3; SAS Institute Inc., Cary, NC), and all analyses were conducted using R (version 2.6.1; R Foundation for Statistical Computing, Vienna, Austria).

3.4 Results

3.4.1 Modeled Exposure Tier Analyses

The summary statistics for Tiers 1, 2a, 2b, and 3 ambient PM_{2.5} exposure estimates have been described previously^{2,21} and are provided in Appendix B6. While the refined ambient PM_{2.5} exposure concentrations (Tiers 2a, 2b, 3) for each case and control period were approximately half of Tier 1 (central-site) values on average, they were all highly correlated with the Tier 1 concentrations ($r = 0.98, 0.98, \text{ and } 0.98$ for Tiers 2a, 2b, and 3, respectively). All relative odds estimates reported below were scaled to the IQR

increase of each Tier: Tier 1 ($10.3 \mu\text{g}/\text{m}^3$), Tier 2a ($5.4 \mu\text{g}/\text{m}^3$), Tier 2b ($5.4 \mu\text{g}/\text{m}^3$), and Tier 3 ($5.4 \mu\text{g}/\text{m}^3$).

Each $10.3 \mu\text{g}/\text{m}^3$ increase in the Tier 1 $\text{PM}_{2.5}$ concentration was associated with a significant increase in the odds of transmural MI (OR = 1.10, 95% CI = 1.01, 1.19). Each IQR increase ($5.4 \mu\text{g}/\text{m}^3$) in the Tier 2a, Tier 2b, and Tier 3 $\text{PM}_{2.5}$ concentrations was associated with the same size increase in the relative odds of a transmural MI with similar 95% confidence intervals across exposure tiers (Tier 2a: OR = 1.10, 95% CI = 1.01, 1.20; Tier 2B: 1.10, 95% CI = 1.01, 1.20; Tier 3: 1.11, 95% CI = 1.02, 1.20; Table 3-1). Similarly, model fits, as measured by the AIC value, were not substantially different across exposure tiers (Table 3-1). Further, using the z-score method, no additional significant relative increase in odds of transmural MI associated with the refined exposure estimates in addition to that associated with Tier 1 $\text{PM}_{2.5}$ concentrations was found (Table 3-2). For example, each IQR ($1.22 \mu\text{g}/\text{m}^3$) increase in the z-score for Tier 1 $\text{PM}_{2.5}$ concentration was associated with a significant increase in the relative odds of a transmural infarction (OR = 1.11, 95% CI = 1.00, 1.23), but an IQR ($0.21 \mu\text{g}/\text{m}^3$) increase in the Tier 2a z-score difference was associated with only a small, non-significant increase in the relative odds (OR = 1.03, 95% CI = 0.90, 1.18). Similarly, increases in the relative odds of MI associated with IQR increases in Tier 2b and Tier 3 z-score differences (0.21 and $0.28 \mu\text{g}/\text{m}^3$, respectively) were small and not statistically significant (Table 3-2) and, thus, added no explanatory power over the Tier 1 estimate.

3.4.2 AER Effect Modification Analyses

As an alternative analysis, whether modeled residential AERs in the 24 hour period immediately before emergency department arrival modified our estimate of the

relative odds of a transmural MI associated with each $10.3 \mu\text{g}/\text{m}^3$ (IQR) increase in the Tier 1 $\text{PM}_{2.5}$ concentration was evaluated. MIs were evenly distributed between the warm (May to October) and cool (November to April) seasons. Summary statistics of the AER distributions for the warm and cool seasons are shown in Appendix B7.

Heterogeneity in the relative odds of transmural MI across AER tertiles was observed, with homes in higher AER tertiles having larger ORs than homes in the low AER tertile. In the warm season, each $10.3 \mu\text{g}/\text{m}^3$ increase in the Tier 1 $\text{PM}_{2.5}$ concentration was associated with increased relative odds of a transmural MI in the middle AER tertile (OR = 1.16, 95% CI = 0.96, 1.39) and high AER tertile (OR = 1.15, 95% CI = 0.98, 1.35), but not the low AER tertile (OR = 0.96, 95% CI = 0.74, 1.25) (Figure 3-1). When the model was run again with the middle and high AER tertiles combined, each $10.3 \mu\text{g}/\text{m}^3$ increase in Tier 1 $\text{PM}_{2.5}$ was associated with a significant increase in relative odds of a transmural MI for the middle and high AER tertiles, combined (OR = 1.15, 95% CI = 1.02, 1.31). Similarly, in the cool season, an increase in the relative odds of a transmural MI associated with each $10.3 \mu\text{g}/\text{m}^3$ increase in Tier 1 concentrations was observed for the middle and high AER tertiles (both individually and combined), but not for the low AER tertile (Figure 3-1).

To further explore the apparent effect-modification by AER, whether AER was actually a surrogate for another spatially-varying factor that might explain the observed variability in relative odds of transmural MI across AER tertiles was assessed. First, the distribution of monitoring sites to which MI patients were assigned within each AER tertile was evaluated and a case-crossover analysis stratified by monitoring-site community was conducted. In the low AER tertile, the majority of study subjects were

residents of the New Brunswick monitoring-site community. In the high AER tertile, the majority of subjects were residents of the Elizabeth monitoring-site community (Appendix B9). Each monitor-specific IQR increase in ambient $PM_{2.5}$ concentration was associated with a non-statistically-significant increase in the relative odds of transmural MI in both New Brunswick (OR = 1.15, 95% CI = 0.95, 1.39) and Elizabeth (OR = 1.11, 95% CI = 0.97, 1.27; Table 3-3). For the other 5 monitors, ORs ranged from 0.78 in Millville to 1.23 in Rahway. However, given the sample sizes, ORs, and 95% confidence intervals within each monitoring location (Table 3-3), there is no clear difference in the relative odds of transmural MI associated with each IQR increase in $PM_{2.5}$ concentration across monitors.

3.5 Discussion

In this case-crossover study of transmural myocardial infarction, use of refined surrogates of personal exposure to $PM_{2.5}$ of outdoor origin that account for the exposure-modifying effects of human activity patterns and/or the indoor transport of ambient $PM_{2.5}$, did not result in larger estimates of the relative odds of a transmural infarction associated with each IQR increase in $PM_{2.5}$ concentration in the previous 24 hours, smaller confidence intervals, nor better model fits compared to analyses that used $PM_{2.5}$ concentrations measured at central-site monitors. However, effect modification of this relative odds estimate by estimated residential AER was observed. This may be attributable to a greater degree of exposure error and resulting bias towards the null in the low AER tertile (less penetration of ambient PM indoors, and therefore more error in estimating one's personal exposure to $PM_{2.5}$ of outdoor origin) compared to the higher

AER tertiles (more penetration of ambient PM indoors and therefore less exposure error), or residual confounding by some unmeasured factor.

Spatial variability, time activity, and losses with outdoor-to-indoor transport are all sources of exposure error in epidemiologic analyses that use central site monitor concentrations as surrogates for exposure to ambient (outdoor-generated) PM_{2.5}. Several recent studies have reported larger effect estimates and/or smaller confidence intervals when exposures were estimated using models that account for spatial variability in outdoor air pollutant concentrations on local scales (e.g. interpolation methods, land use regression) in place of concentrations measured at a single monitor or averaged over all monitors in a region.^{e.g. 3-6} However, at the time the work herein was published, none had directly compared central-site PM_{2.5} with models accounting for human activity patterns and the indoor transport of ambient PM_{2.5} in a large epidemiologic study. Ebelt et al.³⁷ estimated individual-level ambient PM_{2.5} exposure in a panel study of 16 subjects using individual-level time-activity diaries (to estimate time spent indoors) and indoor PM_{2.5} concentrations estimated using a mass balance model. Associations between cardiopulmonary outcomes (e.g., heart rate variability, forced expiration volume) and ambient PM_{2.5} exposure were calculated with this exposure metric, as well as ambient PM_{2.5} concentrations measured at central-site monitors. Contrary to the findings presented in this study, the Ebelt et al.³⁷ analyses that used individual-level information to model ambient PM_{2.5} exposures resulted in larger health effect estimates and smaller confidence intervals compared to the analyses that used central-site ambient PM_{2.5} concentrations.

Multiple factors could have contributed to the differences between the findings presented here and those of Ebelt et al.³⁷ One possibility is that many of the factors that

are accounted for in the refined exposure estimates would not be expected to contribute to exposure error (or bias) in the case-crossover design. Because cases serve as their own controls in this design, factors that differ across subjects, but are largely constant within subjects (e.g., proximity to local $\text{PM}_{2.5}$ sources and differences in AERs or particle losses that stem from differences in housing stock, air conditioning prevalence or human activity patterns) would be expected to have a minimal impact on effect estimates. Similarly, with the case-control period confined to one calendar month, any factors that vary on time-scales longer than a month (e.g., seasonal variability in AER driven by indoor-outdoor temperature differences, natural ventilation, or air conditioning use) would be expected to have little or no effect on the relative odds estimates. Control periods are also matched to case periods by weekday, calendar month, and hour of the day, likely reducing the influence of much of the within-subject variability in human activity patterns occurring on these time-scales. Although not directly evaluated in this study, time-series analyses, which are also temporal contrasts of daily pollutant concentrations and daily counts of health outcomes, may also be only minimally impacted by these factors.

In addition, the refined exposure estimates used in Ebelt et al.³⁷ were based on subject-level time-activity diaries and home-specific penetration and persistence of ambient $\text{PM}_{2.5}$, while here, human activities and the indoor transport of ambient $\text{PM}_{2.5}$ were modeled using census-tract level data and were then averaged over the area within a 10 km radius of each central-site monitor. For example, human activity patterns simulated with SHEDS for Tier 2a exposure estimates were estimated based on census-tract level demographic data. Similarly, modeled AER distributions for each census tract

were used in the calculation of Tier 2b APP and Tier 3 exposure surrogates, rather than individual-level AERs. Further, species mass fractions were not available in every monitoring area and were estimated as the mass fractions measured at the New Brunswick monitor, which was most highly correlated with the other monitors across the state. These were used with local mass concentrations. The spatial resolution of data used to calculate the refined exposure estimates is a limitation of this study. Uncertainty resulting from these limitations could have contributed to exposure error in the refined exposure surrogates and, thus, the potential benefits of the refined exposure surrogates may not have been fully realized.³⁸ However, when exposures were estimated at the zip-code level, rather than averaging over 10 km (Appendix B10), no increase in ORs, reduction in 95% CIs, nor improved model fits were observed. The potential for uncertainty due to averaging and the associated exposure error and bias was likely reduced in the "AER Effect Modification" analyses because a single parameter was focused on, requiring fewer assumptions and, thus, reduced possibility of compounding of exposure prediction errors. It is possible that simpler methods to account for variability in exposure to PM_{2.5} of outdoor origin resulting mostly from variability in the indoor transport of ambient PM_{2.5} (e.g., including AER as an interaction term in the conditional logistic regression model) may more accurately capture variability in effect than these more complicated exposure models, which could be subject to greater uncertainty.

The differences in the results of the "Modeled Tiered Exposure" and "AER Effect Modification" analyses may also be explained, in part, by differences in study design. In the tiered exposure analysis, the relative odds of transmural MI within different time periods during each subject's person-time were essentially compared. Therefore,

non-time varying confounders such as subject characteristics (age, health history, etc.), residential location (and any potential differences in the pollutant mixture due to different pollution sources, source proximity), and housing characteristics (leakage) were controlled by design. In the effect modification analysis, each relative odds estimate within each AER tertile also has this feature. However, when these AER-tertile-specific relative odds estimates were then contrasted, different subjects with their inherent differences in these characteristics were compared. Thus, these characteristics may now act as confounders in this analysis. As a result, differences in these AER-tertile-specific relative odds estimates could be due, in part, to differences in AER, as well as differences in subject characteristics (e.g. age, co-morbidity, proximity to sources, housing stock, access to healthcare, smoking status, etc.) if those characteristics are covariant with AER. For example, low socio-economic status (SES) has been identified as a predictor of susceptibility to negative health outcomes associated with PM exposure.³⁹ Further, low income residents tend to live in homes with higher AERs and, therefore, are exposed to a larger fraction of ambient PM_{2.5} (and smaller fraction of indoor emissions) than residents with higher SES.^{40,41} In fact, because SES is a predictor of AER, poverty status is included in the residential AER model (Appendix B).³¹ Thus, it is conceivable that the results showing effect modification of the PM_{2.5}-MI association by AER could actually reflect effect modification by SES or a combination of AER and SES. It is also possible that higher AERs, in addition to access to health care and other factors, help to explain the associations between low SES and adverse health outcomes observed in previous studies. Notably, differences in age, gender, race/ethnicity, and co-morbidities by AER tertile were not observed (Appendices B8, B9). Further, if location-specific factors other

than AER were contributing to these findings, an increased relative odds of transmural MI in the monitoring-site community in which the majority of subjects were assigned to the high-AER tertile (i.e., Elizabeth) and a smaller effect estimate in communities in which the majority subjects were assigned to the low-AER tertile (i.e. New Brunswick) would be expected. Instead, larger relative odds of transmural MI were observed in New Brunswick compared to Elizabeth (Table 3-3), which suggests that the observed effect modification is related to variability in AER.

The modification of MI risk by community-average AER is consistent with the results of concurrently conducted studies that found that percent increases in short-term mortality associated with given increases in outdoor ozone and PM₁₀ concentrations were larger for cities with higher annual average AERs compared to those with smaller AERs.^{42,43} Previous studies have also shown that home-ventilation conditions (e.g., infiltration through cracks in the building shell, air flow through open windows) and activities that affect particle losses indoors (e.g., AC use) impact ambient PM_{2.5} exposures.^{e.g. 9,44} Sarnat et al.⁴⁴ concluded that ambient monitors were good surrogates for exposure in well-ventilated homes, but were poor exposure surrogates in homes with windows and doors closed. The results presented here are also consistent with studies that have demonstrated a reduced risk of morbidity and mortality with increased prevalence of central AC.¹²⁻¹⁶ As noted above, *F* tends to be lower for homes with central AC in use due to increased particle losses in AC filters.¹⁷⁻¹⁹ Further, AERs tend to be lower for homes with AC in use compared to those with open windows,⁴⁵ which also contributes to lower *F* values.

3.6 Conclusions

Use of refined exposure surrogates that account for human activity patterns and/or the indoor transport of ambient PM_{2.5} in this case-crossover study did not result in larger health effect estimates, narrower confidence intervals, or better model fits compared to the analyses that used central-site PM_{2.5} concentrations to estimate PM_{2.5} exposure. For the level of exposure-estimate-refinement considered here, these findings add support to the use of central-site PM_{2.5} concentrations for epidemiological studies that employ similar case-crossover study designs and other similar temporal analytic methods. These findings also illustrate that variability in factors that influence the fraction of ambient PM_{2.5} in indoor air (e.g., AER) can bias health effects estimates in study designs for which a spatio-temporal comparison of exposure effects across subjects is conducted.

3.7 References

1. Mustafic, H.; Jabre, P.; Caussin, C.; Murad, M. H.; Escolano, S.; Tafflet, M.; Perier, M.C.; Marijon, E.; Vernerey, D.; Empana, J. P.; Jouven, X., Main air pollutants and myocardial infarction: a systematic review and meta-analysis. *JAMA*.
2. Rich, D. Q.; Kipen, H. M.; Zhang, J.; Kamat, L.; Wilson, A. C.; Kostis, J. B., Triggering of transmural infarctions, but not non-transmural infarctions, by ambient fine particles. *Environ. Health Perspect.* **2010**, 118, 1229 - 1235.
3. Poulou, T.; Kanaroglou, P. S.; Elliott, S. J.; Penegelly, D., Assessing the health impacts of air pollution: a re-analysis of the Hamilton children's cohort data using a spatial analytic approach. *Int. J. Environ. Health Res.* **2008**, 18, 17 - 35.
4. Kim, S. Y.; Sheppard, L.; Kim, H., Health effects of long-term air pollution: influence of exposure prediction methods. *Epidemiology* **2009**, 20, 442 - 450.
5. Sahsuvaroglu, T.; Jerrett, M.; Sears, M. R.; McConnell, R.; Finkelstein, N.; Arain, A.; Newbold, B.; Burnett, R., Spatial analysis of air pollution and childhood asthma in Hamilton, Canada: comparing exposure methods in sensitive subgroups. *Environ. Health* **2009**, 8, doi:10.1186/1476-069X-8-14.
6. Son, J. Y.; Bell, M. L.; Lee, J. T., Individual exposure to air pollution and lung function in Korea: spatial analysis using multiple exposure approaches. *Environ. Res.* **2010**, 110, 739 - 749.

7. Sarnat, S. E.; Coull, B. A.; Ruiz, P. A.; Koutrakis, P.; Suh, H. H., The influences of ambient particle composition and size on particle infiltration in Los Angeles, CA, Residences. *J. Air Waste Manage. Assoc. Association* **2006**, 56, 186 - 196.
8. Hering, S. V.; Lunden, M. M.; Thatcher, T. L.; Kirchstetter, T. W.; Brown, N. J., Using regional data and building leakage to assess indoor concentrations of particles of outdoor origin. *Aerosol Sci. Technol.* **2007**, 41, 639 - 654.
9. Meng, Q. Y.; Spector, D.; Colome, S.; Turpin, B. J., Determinants of indoor exposure to PM_{2.5} of indoor and outdoor origin during the RIOPA study. *Atmos. Environ.* **2009**, 43, 5750 - 5758.
10. Zeger, S. L.; Thomas, D.; Dominici, F.; Samet, J. M.; Schwartz, J.; Dockery, D.; Cohen, A., Exposure measurement error in time-series studies of air pollution: Concepts and consequences. *Environ. Health Perspect.* **2000**, 108, 419 - 426.
11. Bateson, T. F.; Coull, B. A.; Hubbell, B.; Kazuhiko, I.; Jerrett, M.; Lumley, T.; Thomas, D.; Vedal, S.; Ross, M., Panel discussion review: session 3 – issue involved in interpretation of epidemiological analyses – statistical modeling. *J. Exposure Anal. Environ. Epidemiol.* **2007**, 17, S90 - S96.
12. Janssen, N. A.; Schwartz, H. J.; Zanobetti, A.; Suh, H. H., Air conditioning and source-specific particles as modifiers of the effect of PM₁₀ on hospital admissions for heart and lung disease. *Environ. Health Perspect.* **2002**, 110, 43 - 49..
13. Zeka, A.; Zanobetti, A.; Schwartz, J., Short-term effects of particulate matter on cause specific mortality: effects of lags and modification by city-specific characteristics. *Occup. Environ. Med.* **2005** 62, 718 - 725.
14. Franklin, M.; Zeka, A.; Schwartz, J., Association between PM_{2.5} and all-cause and specific cause mortality in 27 US communities. *J. Exposure Sci. Environ. Epidemiol.* **2007**, 17, 279 - 287.
15. Bell, M. L.; Ebisu, K.; Peng, R. D.; Dominici, F., Adverse health effects of particulate air pollution: modification by air conditioning. *Epidemiology* **2009**, 19, 680 - 689.
16. Baxter, L. K.; Özkaynak, H.; Franklin, M.; Schultz, B. D.; Neas, L. M., The use of improved exposure factors in the interpretation of fine particulate matter epidemiological results. *Air Quality, Atmosphere, and Health* **2013**, 6, 195 - 204.
17. Thornburg, J.W.; Ensor, D. S.; Rodes, C. E.; Lawless, P. A.; Sparks, L. E.; Mosely, R. B., Penetration of particles into buildings and associated physical factors, Part I: model development and computer simulations. *Aerosol Sci. Technol.* **2001**, 34, 284 - 296.

18. Waring, M. S.; Siegel, J. A., Particle loading rates for HVAC filters, heat exchangers, and ducts. *Indoor Air* **2008**, 18, 209 - 224.
19. Stephens, B.; Siegel, J. A., Comparison of test methods for determining the particle removal efficiency of filters in residential and light-commercial central HVAC systems. *Aerosol Sci. Technol.* **2012**, 46, 504 - 513.
20. Hodas, N.; Meng, Q. Y.; Lunden, M.; Rich, D. Q.; Özkaynak, H.; Baxter, L.; Zhang, Q.; Turpin, B. J., Variability in the fraction of ambient fine particulate matter found indoors and heterogeneity in health effect estimates. *J. Exposure Sci. Environ. Epidemiol.* **2012**, 22, 448 - 454.
21. Baxter, L.; Burke, J.; Lunden, M.; Turpin, B. J.; Rich, D. Q.; Thevenet-Morrison, K.; Hodas, N.; Özkaynak, H., Influence of human activity patterns and residential air exchange rates on modeled distributions of PM_{2.5} exposure compared to central-site monitoring data. *J. Exposure Sci. Environ. Epidemiol.* **2013**, 23, 241 - 247
22. Kostis, J. B.; Wilson, A. C.; O'Dowd, K.; Gregory, P.; Chelton, S.; Cosgrove, N. M.; Chirala, A.; Cui, T., Sex differences in the management and long-term outcome of acute myocardial infarction. A statewide study. MIDAS Study Group. Myocardial Infarction Data Acquisition System. *Circulation* **1994**, 90, 1715 - 1730.
23. Kostis, J. B.; Wilson, A. C.; Lacy, C. R.; Cosgrove, N. M.; Ranjan, R.; Lawrence-Nelson, J., Time trends in the occurrence and outcome of acute myocardial infarction and coronary heart disease death between 1986 and 1996 (a New Jersey statewide study). *Am. J. Cardiol.* **2001**, 88, 837 - 841.
24. US EPA. Technology Transfer Network (TTN)—AQS Datamart. **2008**, Available: <http://www.epa.gov/ttn/airs/aqsdatamart/index.htm> [accessed 19 July 2010].
25. Burke, J. M.; Zufall, M. J.; Özkaynak, H., A population exposure model for particulate matter: Case study results for PM_{2.5} in Philadelphia, PA. *J. Exposure Anal. Environ. Epidemiol.* **2001**, 11, 470 - 489.
26. McCurdy, T.; Glen, G.; Smith, L.; Lakkadi, Y., The National Exposure Research Laboratory's Consolidated Human Activity Database. *J. Exposure Anal. Environ. Epidemiol.* **2000**, 10, 1 - 13.
27. Lunden, M. M.; Thatcher, T. L.; Hering, S. V.; Brown, N. J., The use of time- and chemically-resolved particulate data to characterize the infiltration of outdoor PM_{2.5} into a residence in the San Joaquin Valley. *Environ. Sci. Technol.* **2003**, 37, 4724 -4732.
28. Lunden, M. M.; Revzan, K. L.; Fischer, M. L.; Thatcher, T. L.; Littlejohn, D.; Hering, S. C.; Brown, N. J., The transformation of outdoor ammonium nitrate aerosols in the indoor environment. *Atmos. Environ.* **2003**, 37, 5633 - 5644.

29. Sherman, M. H.; Grimsrud, D. T., Measurement of infiltration using fan pressurization and weather data. Lawrence Berkeley National Laboratory Report. LBNL-10852, Berkeley, CA, **1980**.
30. Sherman, M. H.; Dickerhoff, D. J., Airtightness of US dwellings. *ASHRAE Transactions* **1998**, 104, 1359 - 1367.
31. Chan, W. R.; Nazaroff, W. W.; Price, P. N.; Sohn, M. D.; Gadgil, A. J., Analyzing a database of residential air leakage in the United States. *Atmos. Environ.* **2005**, 39, 3445 - 3455.
32. Maclure, M., The case-crossover design: a method for studying transient effects on the risk of acute events. *Am. J. Epidemiol.* **1991**, 133, 144 - 153.
33. Levy, D.; Lumley, T.; Sheppard, L.; Kaufman, J.; Checkoway, H., Referent selection in case-crossover analyses of acute health effects of air pollution. *Epidemiology* **2001**, 12, 186 - 192.
34. Steadman, R. G., The assessment of sultriness. Part II: effects of wind, extra radiation and barometric pressure on apparent temperature. *J. Appl. Meteorol.* **1979**, 18, 874 - 885.
35. Zanobetti, A.; Schwartz, J., The effect of particulate air pollution on emergency admissions for myocardial infarction: a multicity case-crossover analysis. *Environ. Health Perspect.* **2005**, 113, 978 - 982.
36. Chuersuwan, N.; Turpin, B. J., Evaluation of time resolved PM_{2.5} data in urban/suburban areas of New Jersey. *J. Air Waste Manage. Assoc.* **2000**, 50, 1780 - 1789.
37. Ebelt, S. T.; Wilson, W. E.; Brauer, M., Exposure to ambient and nonambient components of particulate matter: a comparison of health effects. *Epidemiology* **2005**, 16, 296 - 405.
38. Szpiro, A. A.; Paciorek, C. J.; Sheppard, L., Does more accurate exposure prediction necessarily improve health effect estimates? *Epidemiology* **2011**, 22, 680 - 685.
39. Sacks, J. D.; Stanek, L. W.; Luben, T. J.; Johns, D. O.; Buckley, B. J.; Brown, J. S.; Ross, M., Particulate matter-induced health effects: who is susceptible? *Environ. Health Perspect.* **2011**, 119, 446 - 454.
40. Sarnat, J. A.; Sarnat, S. E.; Chang, H.; Mulholland, J.; Özkaynak, H.; Isakov, V., Spatiotemporally-resolved air exchange rate as a modifier of acute air pollution related morbidity. *J. Exposure Sci. Environ. Epidemiol.* (in press).

41. Sarnat, S. E.; Sarnat, J. A.; Mulholland, J.; Isakov, V.; Özkaynak, H.; Chang, H., Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. *J. Exposure Sci. Environ. Epidemiol.* **2013**, 23, 593 - 605.
42. Chen, C.; Zhao, B.; Weschler, C. J., Assessing the influence of indoor exposure to “outdoor ozone” on the relationship between ozone and short-term mortality in U.S. communities. *Environ. Health Perspect.* **2012**, 120, 235 - 240.
43. Chen, C.; Zhao, B.; Weschler, C. J., Indoor exposure to "outdoor PM10": assessing its influence on the relationship between PM10 and short-term mortality in U.S. cities. *Epidemiology* **2012**, 23, 870 - 878.
44. Sarnat, J. A.; Koutrakis, P.; Suh, H. H., Assessing the relationship between personal particulate and gaseous exposures of senior citizens living in Baltimore, MD. *J. Air Waste Manage. Assoc.* 2000, 50, 1184 - 1198.
45. Breen, M. S.; Breen, M.; Williams, R. W.; Schultz, B. D., Predicting residential air exchange rates from questionnaires and meteorology: model evaluation in central North Carolina. *Environ. Sci. Technol.* **2010**, 44, 9349 - 9356.

Table 3-1. Relative increase in odds of a transmural infarction associated with an IQR increase in PM_{2.5} concentration, by exposure Tier

Tier	IQR	N	AIC	OR	95% CI	p-value
Tier 1	10.3	1561	4397.4	1.10	1.01, 1.19	0.03
Tier 2A SHEDS	5.4		4397.2	1.10	1.01, 1.20	0.03
Tier 1	10.3	1552*	4367.7	1.09	1.01, 1.19	0.04
Tier 2B APP	5.4		4366.8	1.10	1.01, 1.20	0.02
Tier 1	10.3	1561	4396.4	1.10	1.01, 1.19	0.03
Tier 3 HYBRID	5.4		4396.1	1.11	1.02, 1.20	0.01

* Subjects were excluded if there was a period of more than nine days between STN PM_{2.5} species concentration measurements for the case period or all control periods

Table 3-2. Relative increase in odds of a transmural infarction associated with each IQR increase in PM_{2.5} concentration, by exposure Tier. Z-score method.

Tier	IQR	N	OR	95% CI	p-value
Tier 1	1.22	1561	1.11	1.00, 1.23	0.04
Tier 2a SHEDS	0.25		1.03	0.90, 1.18	0.65
Tier 1	1.22	1552*	1.12	1.02, 1.23	0.02
Tier 2b APP	0.21		1.05	0.97, 1.14	0.21
Tier 1	1.22	1561	1.12	1.03, 1.22	0.01
Tier 3 HYBRID	0.18		1.05	0.99, 1.11	0.12

* Subjects were excluded if there was a period of more than nine days between STN PM_{2.5} species concentration measurements for the case period or all control periods

For Tier 1, IQR refers to the interquartile range of z-scores, while for the refined exposure models (Tiers 2a, 2b, and 3), it refers to the interquartile range of the z-score difference (e.g., the difference between the Tier 1 and Tier 2a z-scores).

Table 3-3. Relative odds of transmural infarction associated with each interquartile range increase in PM_{2.5} concentration, stratified by monitoring-site, in order of increasing median air exchange rate

Monitor location	Median Air Exchange Rate (h⁻¹)	IQR (µg/m³)	OR	95% CI	p-value
Flemington	0.32	8.9	0.98	0.40, 2.39	0.96
New Brunswick	0.41	8.4	1.15	0.95, 1.39	0.15
Camden	0.50	10.3	1.04	0.86, 1.25	0.68
Millville	0.50	9.5	0.78	0.47, 1.30	0.34
Rahway	0.52	9.3	1.23	0.87, 1.74	0.24
Elizabeth	0.60	11.7	1.11	0.97, 1.27	0.13
Jersey City	0.66	12.2	1.17	0.86, 1.59	0.32

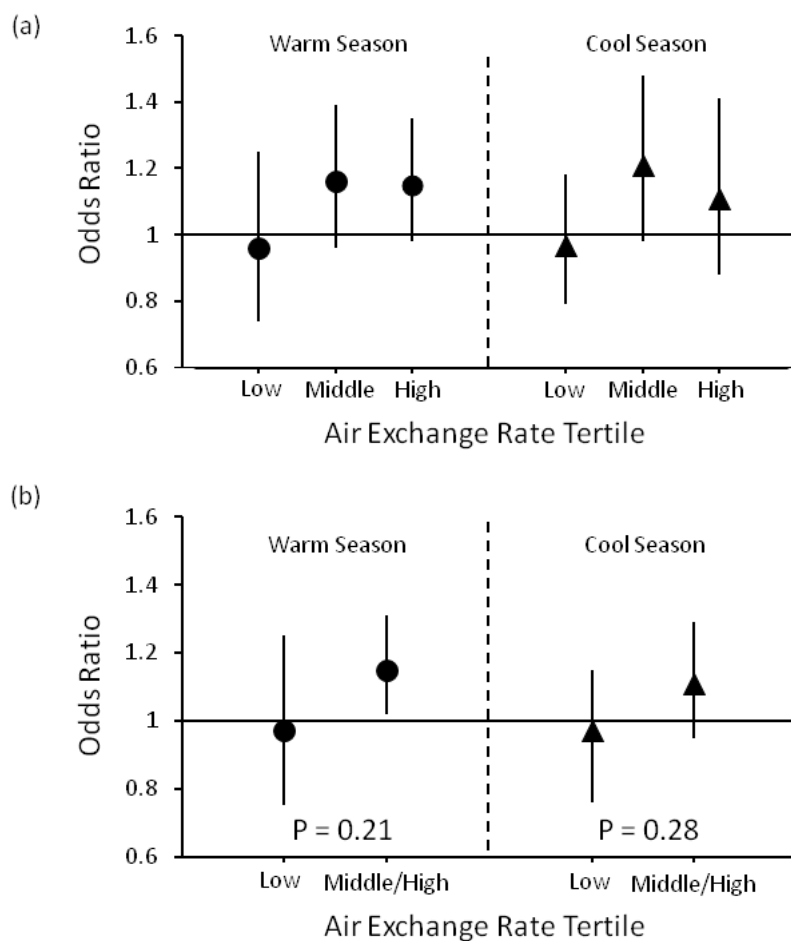


Figure 3-1. Relative odds of transmural infarction associated with each interquartile range increase in Tier 1 (central-site) PM_{2.5} concentration, stratified by air exchange rate tertile. (a) low, middle, and high AER tertiles and (b) for low and middle/high AER tertiles combined

Chapter 4. Toward Refined Estimates of Ambient PM_{2.5} Exposure: Evaluation of a Physical Outdoor-to-Indoor Transport Model

Material in this chapter has been published previously as:

Hodas, N.; Meng, Q. Y.; Lunden, M. M.; Turpin, B. J., Toward Refined Estimates of Ambient PM_{2.5} Exposure: Evaluation of a Physical Outdoor-to-Indoor Transport Model. *Atmos. Environ.* **2014**, 83, 229 - 236.

4.1. Abstract

Because people spend the majority of their time indoors, the variable efficiency with which ambient PM_{2.5} penetrates and persists indoors is a source of error in epidemiologic studies that use PM_{2.5} concentrations measured at central-site monitors as surrogates for ambient PM_{2.5} exposure. To reduce this error, practical methods to model indoor concentrations of ambient PM_{2.5} are needed. Toward this goal, we evaluated and refined an outdoor-to-indoor transport model using measured indoor and outdoor PM_{2.5} species concentrations and air exchange rates from the Relationships of Indoor, Outdoor, and Personal Air Study. Herein, we present model evaluation results, discuss what data are most critical to prediction of residential exposures at the individual-subject and populations levels, and make recommendations for the application of the model in epidemiologic studies. This paper demonstrates that not accounting for certain human activities (air conditioning and heating use, opening windows) leads to bias in predicted residential PM_{2.5} exposures at the individual-subject level, but not the population level. The analyses presented also provide quantitative evidence that shifts in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport contribute significantly to variability in indoor ambient organic carbon concentrations and suggest that methods

to account for these shifts will further improve the accuracy of outdoor-to-indoor transport models.

4.2. Introduction

While people spend the majority of time indoors,¹ fine particulate matter (PM_{2.5}) concentrations measured at outdoor central-site monitors are commonly used as surrogates for exposure to PM_{2.5} of outdoor (ambient) origin in epidemiological studies. The use of central-site PM_{2.5} concentrations as ambient PM_{2.5} exposure surrogates inherently assumes that indoor and outdoor ambient PM_{2.5} concentrations are highly correlated. However, the fraction of ambient PM_{2.5} that penetrates and persists indoors (F) varies with multiple factors including meteorological conditions, the physical and chemical properties of ambient PM_{2.5}, housing characteristics, and home ventilation conditions.²⁻⁹ Exposure error associated with not accounting for variability in F is expected to contribute to an underestimation of health effects associated with ambient PM_{2.5} exposures.¹⁰

In order to reduce this exposure error, practical methods to predict indoor concentrations of ambient PM_{2.5} are needed. Toward this goal, we evaluated and refined a physical mass-balance model using measurements from the Relationships of Indoor, Outdoor, and Personal Air (RIOPA) study.^{11,12} An earlier version of the model was applied in two epidemiologic studies: one that explored associations between ambient PM_{2.5} exposures and myocardial infarction (MI) and the other, associations with birth outcomes.¹³⁻¹⁵ The work herein provides a partial validation of the exposure estimates used in those studies, while also providing new insights that are used to refine the model. This paper highlights the measurements and data most critically needed to facilitate the prediction of residential ambient PM_{2.5} exposures in epidemiological studies.

4.3. Methods

4.3.1 Modeled Indoor PM_{2.5} Concentrations

Indoor concentrations of ambient particulate sulfate, elemental carbon (EC), and organic carbon (OC) were calculated for RIOPA homes (Appendix C1) with a mass-balance model. The model describes the concentration of chemically non-reactive PM_{2.5} species j in indoor air ($C_{in,j}$) as a function of its outdoor concentration ($C_{out,j}$), residential air exchange rate (AER), particle penetration efficiency (P_j), and the depositional loss rate of species j in indoor air ($k_{dep,j}$)⁵:

$$\frac{dC_{in,j}}{dt} = C_{out,j}(P_j \times AER) - C_{in,j}(k_{dep,j} + AER) \quad (1)$$

Forty-eight hour average outdoor sulfate, EC, and OC concentrations and AERs measured at each RIOPA home (Appendix C2) were used as model inputs (nitrate was not measured during RIOPA). Details regarding RIOPA study measurements are provided in Appendix C. Briefly, AERs were measured with a perfluorocarbon tracer method.¹⁶ PM_{2.5} filter samples were analyzed for EC and OC ($\mu\text{gC}/\text{m}^3$) with a Sunset carbon analyzer and for sulfur by energy-dispersive XRF spectrometry and expressed as sulfate.^{11,12} OC was corrected for the adsorption artifact by subtracting the organic mass on the backup filter.¹² Due to the long averaging time of RIOPA measurements, indoor concentrations were calculated with the time-averaged solution to equation 1:

$$C_{in,j} = C_{out,j} \left((AER \times P) / (AER + k_{dep,j}) \right) \quad (2).$$

A review of published species size distributions from diverse geographic locations and seasons (Appendix C, Appendix C3) was conducted to identify "typical" size distributions (i.e., number of modes, mass median diameter of each mode, and the

fraction of mass in each mode) for sulfate, EC, and OC. Values of k_{dep} were then selected for the mass median diameter of each size mode of each PM_{2.5} species size distribution (Table 4-1) using the fourth-order polynomial fit to measured particle-size-resolved deposition rates from Nazaroff.³ While this method provides a means to estimate reasonable values of k_{dep} , the reader should be aware that factors in addition to particle size can contribute to variability in k_{dep} (e.g. particle density, room airflow conditions)^{3,17} and there is heterogeneity in measured size-resolved particle deposition rates across studies.³ A constant P of 0.8, the median value reported by Chen and Zhao⁷ for particles in the size range considered here, was used for all species. Like k_{dep} , many factors contribute to variability in P . For example, laboratory studies have demonstrated that the geometry and roughness of cracks in a building shell can contribute to variability in P ;^{3,7,18} however, these cracks have not been well characterized for individual homes and are likely to be highly variable.³ As a result, this variability is not accounted for in our calculations. In subsequent sections, we explore other contributors to variability in P such as particle size and home ventilation conditions.

4.3.2. Model Evaluation

We compared modeled indoor concentrations of *ambient* sulfate, EC, and OC with the measured indoor concentrations of these PM_{2.5} species (Appendix C, Appendix C2) for each (occupied) RIOPA home. In epidemiologic analyses, the extent to which a model is successful in predicting exposures at the individual-subject level is described by the covariance between actual and estimated exposures. As a result, we examined correlations between measured and modeled indoor concentrations. Paired t-tests were also conducted to evaluate whether pairs of measured and modeled indoor PM_{2.5} species

concentrations were significantly different at the 95% confidence level. To assess model performance at the population level, chi-square tests were used to examine whether cumulative distributions of measured and modeled indoor concentrations had the same underlying distribution at a 95% confidence level. All analyses were conducted with SAS software (version 9.3; SAS Institute Inc., Cary, NC).

Using the same methods, we also evaluated whether measured residential *outdoor* PM_{2.5} concentrations were good predictors of indoor ambient PM_{2.5} concentrations. Much of the recent work aimed at refining ambient PM_{2.5} exposure surrogates has focused on accounting for spatial variability in *outdoor* PM_{2.5} concentrations (e.g. land use regression, interpolation methods).^{19,20} Ambient PM_{2.5} concentrations measured outside of RIOPA homes provide spatially-resolved measures of outdoor PM_{2.5} concentrations. A comparison between measured residential *outdoor* PM_{2.5} concentrations and modeled indoor ambient PM_{2.5} concentrations evaluates whether exposure metrics that account for outdoor-to-indoor transport offer improvement over exposure metrics that account only for spatial variability in outdoor concentrations.

4.3.3. Attributing model-measurement differences: Human activities

To focus our efforts to refine the outdoor-to-indoor transport model, we explored the contributions of several factors to differences between modeled and measured indoor PM_{2.5} species concentrations. First, we evaluated the extent to which model-measurement differences could be attributed to the fact that the model does not account for the effects of human activities likely to influence F . Human-activity variables that were likely to influence the efficiency with which ambient PM_{2.5} penetrated and/or persisted in RIOPA homes were selected from questionnaires administered to RIOPA participants to

characterize home-occupant activities during sample collection (Appendix C).^{6,11} The activities were: (1) time with windows open, (2) time with central air conditioning (AC) in use, and (3) time with central heating in use. Differences between modeled and measured indoor $PM_{2.5}$ concentrations were regressed on these activity variables using multiple linear regression (MLR) with stepwise selection ($\alpha = 0.15$ for variable entrance and removal threshold; SAS version 9.3). Variance inflation factors indicated that the human activities were not significantly correlated with each other. Outliers were detected based on the student's t and a value was considered an outlier if t was greater than 2. It should be noted that outliers are likely indicators of strong indoor sources. Outlier homes were excluded from all following analyses to avoid influence of strong indoor sources of $PM_{2.5}$ on model evaluation results, as the model predicts only the contribution of *ambient* $PM_{2.5}$ to indoor concentrations and not the contribution of indoor sources.

We refined the model to account for the human-activity variables selected as significant predictors of model-measurement differences based on assumptions about the ways in which each activity variable would influence F . For homes with open windows, there is little to no removal of the particles entering the home and, thus, we assumed a penetration efficiency (P_{window}) of 1.0 (Table 4-1).⁷ For homes with central AC or heating in use, a filter penetration efficiency term (P_{filter}) was multiplied by the right side of equation (2) to account for losses in the filters of central heating and cooling systems. Values of P_{filter} were selected from particle-size-resolved filtration efficiencies for residential furnace filters (assuming a pressure drop of 125 Pa across the filter to account for particle loading)²¹ using the same assumptions about species size distributions as were used to select k_{dep} values (Table 4-1). If more than one activity occurred within a home, we accounted only for the dominant activity (i.e. the activity carried out for the longer

period of time). The performance of this refined version of the model was evaluated using the same methods as described above.

4.3.4. Attributing model-measurement differences: Indoor sources of OC

Because sulfate and EC are non-volatile and have minimal indoor sources, indoor concentrations of these species are likely driven by outdoor-to-indoor transport.⁴ However, organics comprised the majority of PM_{2.5} emitted or formed inside RIOPA homes (on average, 41 – 76%).²² Because the aim of the current modeling is to predict indoor concentrations of *ambient* PM_{2.5}, there is a need for an estimate of the measured indoor OC that can be attributed to outdoor sources. We estimated this by regressing measured indoor OC concentrations on measured outdoor OC concentrations using robust regression (SAS version 9.3). Robust regression down-weights outliers and, thus, reduces the influence of strong indoor sources on the regression equation.^{23,24} The intercept of the resulting regression equation provides an average indoor-source strength and the slope is a population-average estimate of F . We multiplied this population-average F by each measured outdoor OC concentration to calculate the distribution of measured indoor OC of ambient origin.^{23,24} When this approach was used for sulfate, which is dominated by outdoor sources, F estimated by robust regression was in good agreement with F calculated as the ratio of measured indoor to measured outdoor sulfate (Appendix C, Appendix C4).

4.3.5. Attributing model-measurement differences: Uncertainty in OC size distributions

Ambient OC size distributions are more variable across sampling locations and seasons than sulfate and EC. We conducted a sensitivity analysis to explore whether

uncertainty in ambient outdoor OC size distributions (and, thus, k_{dep}) was a source of error in modeled indoor ambient OC. Indoor ambient OC concentrations were calculated assuming two alternative size distributions: (1) a bimodal distribution with an ultrafine peak at 0.08 μm and a broad accumulation mode peaking at 0.4 μm with 20% and 80% of OC mass comprising each mode, respectively ($k_{dep} = 0.07 \text{ h}^{-1}$)^{25,26} and (2) a trimodal distribution with an ultrafine peak (comprising 20% of OC mass) and an accumulation mode comprised of a condensation mode (0.2 μm) and droplet mode (0.7 μm) of equal mass proportions. The k_{dep} values for alternative size distribution (2) are the same as those shown in Table 4-1, but the mass fractions comprising the condensation and droplet modes are different from those explored in the main analysis. Indoor OC concentrations calculated assuming each of the three size distributions were compared to evaluate the sensitivity of the model to uncertainty in k_{dep} associated with variability in OC size distributions.

4.3.6. Attributing model-measurement differences: Phase changes of ambient organics

Predicting the outdoor-to-indoor transport of particulate OC is further complicated by the fact that organics can undergo phase changes due to indoor-outdoor differences in temperature, surface area, and the availability of particulate matter for sorption.^{22,27-29} Because ambient OC is comprised of thousands of compounds with largely unknown identities,³⁰ it is not possible to calculate the change in gas-particle partitioning with outdoor-to-indoor transport from first principles. In order to explore the influence of phase changes on F , a surrogate is needed. We used 5 - 7 ring polycyclic aromatic hydrocarbons (PAHs), which were measured in the gas and particle phases

inside and outside of 76 RIOPA homes,³¹ for this purpose. The PAHs included were benzo[b+k]fluoranthene, benzo[e]pyrene, benzo[a]pyrene, perylene, indeno[1,2,3-c,d]pyrene, dibenzo[a,c+a,h]anthracene, benzo[g,h,i]perylene, and coronene. These PAHs are predominantly of outdoor origin and were mainly in the particle phase under the ambient conditions measured outside RIOPA homes,³¹ making them a useful surrogate for ambient particulate OC.

Using MLR, we explored the extent to which variability in measured indoor OC concentrations could be explained by (1) physical losses associated with outdoor-to-indoor transport (i.e., those already accounted for in the model) and (2) shifts in the gas-particle partitioning of ambient organics with indoor transport (using changes in partitioning of 5-7 ring PAHs). We regressed measured indoor OC on modeled indoor OC and on the indoor-outdoor difference in the pooled gas-particle partitioning coefficient (K_p) of the 5 - 7 ring PAHs ($\alpha = 0.15$ for variable entrance and removal threshold; SAS version 9.3). K_p was calculated as the ratio of the pooled concentration of PAHs in the particle phase to their concentration in the gas phase, normalized by the total $PM_{2.5}$ concentration.³² No collinearity between variables was found and outliers were removed using the same criteria as described above.

4.4. Results and Discussion

4.4.1. Initial Model

Agreement between modeled indoor ambient $PM_{2.5}$ concentrations and measured indoor concentrations varied by species. While modeled indoor ambient EC concentrations were well correlated with measured indoor EC ($R^2 = 0.70$), the model generally underestimated indoor EC (Figure 4-1a, 4-2a). In fact, pairs of measured and

modeled indoor EC were significantly different and modeled and measured values did not have the same underlying distribution at a 95% confidence level, suggesting that model refinements are needed to predict residential EC exposures.

For sulfate, the initial model performed reasonably well at the population level, but the model under-predicted indoor sulfate for many high concentration homes (Figure 4-2b). Cumulative distributions of measured and modeled indoor sulfate (Figure 4-1b) had the same underlying distribution ($P = 0.87$). While modeled and measured values were well correlated ($R^2 = 0.86$; Figure 4-2b), modeled indoor sulfate concentrations were significantly lower than measured concentrations at a 95% confidence level according to a paired t-test. Thus, the initial model could be applied to estimate residential sulfate exposures at the population level, but refinements are needed to improve exposure estimates at the individual-subject level, particularly for high-end exposures.

Measured indoor OC concentrations were not well captured by the initial model (Figure 4-1c, 4-2c), which accounts for physical losses of ambient OC during outdoor-to-indoor transport into closed homes without air conditioning, but does not account for phase changes or indoor sources. Modeled indoor particulate OC of ambient origin explained only 4% of the variability in total particulate OC measured indoors ($R^2 = 0.04$). Further, measured and modeled indoor OC concentrations were significantly different at a 95% confidence level at both the individual- (paired) and population (distribution) levels. Contributors to this poor agreement are explored below.

4.4.2. Model refinement: accounting for human activities

Human activities that were not accounted for in the initial model helped to explain differences between modeled and measured indoor sulfate (Table 4-2), but not EC and OC. For sulfate, all activity variables included in the MLR analysis were selected as significant predictors of model-measurement differences ($\alpha < 0.15$) and together explained 31% of the variance in these differences (Table 4-2). We refined the model to account for these activities based on our assumptions regarding the ways in which each activity would influence F (Table 4-1).

Improved agreement between measured and modeled indoor sulfate at the individual-subject level was substantial (Figure 4-3a). Pairs of measured and modeled indoor sulfate concentrations were no longer significantly different ($P = 0.60$). Indoor sulfate concentrations modeled with the refined model explained 90% of the variance in measured indoor sulfate, compared to 86% for the initial model. While use of the refined model also improved agreement between measured and modeled indoor sulfate distributions ($P = 0.996$; Figure 4-3c), the initial model distribution was not significantly different from the measured distribution to begin with. Thus, while the initial model is adequate for predicting sulfate distributions, we recommend the use of the refined model when estimating residential sulfate exposures at the individual-subject level. Notably, most epidemiologic studies do not focus on sulfate exposures, but rather on exposure to total ambient $PM_{2.5}$. When using a mass-balance model like the one explored here, exposure to total ambient $PM_{2.5}$ would be calculated by summing predicted indoor concentrations of the individual species. The fact that a refined version of the model is needed to predict residential sulfate exposures has implications for the design of epidemiologic studies focused on both $PM_{2.5}$ species and total ambient $PM_{2.5}$ exposures,

as it requires the collection of human activity data (e.g., pertaining to windows, air conditioning, and heating) using questionnaires or activity diaries or a method to estimate human activity patterns (e.g. sampling from a distribution of published time-activity patterns³³) over the length of the study.

While accounting for human activities in the model improved model-measurement agreement for sulfate, the examined human activities had a minimal impact on F for ambient EC. The small impact of human activities can likely be explained by the EC size distribution. Values of P_{filter} for 80 nm particles are ~90%²¹ and, thus, use of central AC or heating is expected to result in only small losses of EC. This also suggests that for the ultrafine-mode, overall penetration efficiencies may be greater than the 0.8 used in the initial calculations. We re-calculated indoor ambient EC assuming a P value of 0.9. With this refinement, the model captured indoor EC concentrations at both the population and individual-subject levels (Figure 4-3b, 4-3d). Measured and modeled EC had the same underlying distribution ($P = 0.65$), they were well correlated ($R^2 = 0.70$), and pairs of measured and modeled indoor EC were not significantly different at a 95% confidence level ($P = 0.16$). These results suggest that human activities might not need to be accounted for when calculating residential EC exposures, but that P values can vary across $PM_{2.5}$ species due to differences in size distributions.

Like EC, human activities were not selected as significant predictors of differences between measured and modeled indoor OC. Based on the assumed size distribution for OC, we would expect that the influence of human activities on F would be minimal for OC in the ultrafine mode, but similar to that for sulfate for the accumulation mode fraction. However, the effect of human activities was likely

overshadowed by the substantial contribution of indoor emissions²² to measured indoor OC.

4.4.3. Accounting for indoor sources of OC

The robust-regression estimated F value for OC was 0.53, suggesting that, on average, 53% of the ambient OC penetrated and persisted indoors (Appendix C5). This value is higher than the F value for OC reported in Polidori et al.,²² which estimated contributions of outdoor-generated OC to total OC measured in RIOPA homes using a Random Component Superposition (RCS) statistical model ($F = 0.32$). Figure 4-4 compares cumulative distributions of ambient indoor OC estimated with the mass-balance model with our robust-regression estimate of F (a comparison with the RCS-estimated F value is available in Appendix C6). Agreement between the distributions of indoor ambient OC estimated with the model and with robust regression (Figure 4-4; mean \pm standard deviation = $2.54 \pm 1.61 \mu\text{gC}/\text{m}^3$ and $1.91 \pm 1.26 \mu\text{gC}/\text{m}^3$, respectively) is improved compared to agreement between modeled indoor ambient OC and measured (total) indoor OC ($6.08 \pm 3.77 \mu\text{gC}/\text{m}^3$). However, the two distributions are still significantly different ($P = 0.0004$). Accounting for human activities and the higher penetration efficiency of ultrafine-mode particles ($P = 0.9$) in calculations of indoor ambient OC did not reduce this bias (mean \pm standard deviation $2.74 \pm 1.81 \mu\text{gC}/\text{m}^3$; Appendix C6).

4.4.4. Variability and uncertainty in OC size distributions

The model showed little sensitivity to the uncertainty in k_{dep} associated with variability in OC size distributions. Estimated indoor OC concentrations were highly correlated across the size-distributions scenarios ($R^2 > 0.99$) and distributions of modeled

indoor OC were in good agreement (mean \pm standard deviation = 2.54 ± 1.61 , 2.62 ± 1.66 , and 2.58 ± 1.64 $\mu\text{gC}/\text{m}^3$ for the initial and two alternative size-distribution scenarios, respectively). The OC size distributions considered here are based on measurements conducted in urban regions in which OC is comprised of a mix of locally- and regionally-generated PM (Appendix C). It is possible that these size distributions are not representative of the OC measured outside of some of the RIOPA study homes. For example, for homes in close proximity to primary $\text{PM}_{2.5}$ sources (as is the case for many RIOPA homes), the majority of OC might be in the ultrafine mode, which would result in smaller depositional losses (and possibly greater P values) than those calculated assuming that accumulation-mode OC comprised a substantial fraction of the OC. It should be noted, however, that this would result in increased calculated indoor ambient OC concentrations and the model already has an upward bias (Figure 4-4).

4.4.5. Shifts in the gas-particle partitioning of ambient OC

We did find evidence that shifts in the gas-particle partitioning of ambient OC with outdoor-to-indoor transport contributed to variability in particulate OC measured inside RIOPA homes. The model-estimated indoor OC, which was included in the MLR analysis to represent physical particle losses associated with outdoor-to-indoor transport, was the most significant predictor of variability in measured indoor OC concentrations ($P = 0.0003$), explaining 20% of this variability ($R^2 = 0.20$). The indoor-outdoor difference in K_p for the 5 - 7 ring PAHs, which we used as a surrogate for changes in the gas-particle partitioning of ambient OC, was also selected as significant predictor of variability in indoor OC ($P = 0.05$), explaining 5% of this variability ($R^2 = 0.05$). While PAHs account for only a small fraction of total OC, this new finding for OC is consistent with previous

work which demonstrated that shifts in the gas-particle partitioning of outdoor-generated PAHs with outdoor-to-indoor transport contributes substantially to variability in residential PAH exposures.^{29,33} Much of the remaining variability can likely be attributed to indoor OC sources, which contributed to 3 - 99% of the OC in these RIOPA homes (calculated by subtracting the robust-regression estimate of indoor ambient OC from total measured indoor OC).

While physical loss was the dominant contributor to variability in indoor ambient OC, our results suggest that refining the model to account for phase changes of OC with outdoor-to-indoor transport would improve the predictive abilities of the model. This is an important area of future work that requires further characterization of the thermodynamic properties of ambient OC and a better understanding of the chemistry that occurs in indoor air, including interactions between indoor- and outdoor-emitted organics.³⁴

4.4.6. Further recommendations for epidemiologic studies

Exposure research has focused on accounting for spatial variability in outdoor air pollution concentrations (e.g., use of residential outdoor concentrations rather than central-site concentrations through land use regression, interpolation between sites, etc).^{19,20} One objective of this study was to explore whether a model that brings the residential outdoor air pollution indoors, offers additional improvement. The mass-balance model did offer improvement over the use of measured outdoor concentrations as residential ambient PM_{2.5} exposure surrogates. As noted above, EC and sulfate have minimal indoor sources⁴ and, thus, the vast majority of the sulfate and EC measured inside RIOPA homes can be attributed to PM_{2.5} of outdoor origin. As expected, measured outdoor sulfate and EC concentrations were well-correlated with measured indoor

concentrations ($R^2 = 0.77$ and 0.69 , respectively), but correlations between modeled and measured indoor concentrations are even stronger ($R^2 = 0.86$ and 0.90 for the initial and refined sulfate models and $R^2 = 0.70$ for EC for both models). Measured outdoor and indoor OC were weakly correlated ($R^2 = 0.03$), undoubtedly because of the substantial contributions of indoor sources to indoor OC concentrations. The mass-balance model offered only a small improvement over measured outdoor OC concentrations ($R^2 = 0.004$). However, when the influence of indoor sources was reduced using the robust-regression estimate of F , indoor OC concentrations calculated with the mass-balance model performed better than measured outdoor OC concentrations (mean \pm standard deviation = 1.91 ± 1.26 , 2.54 ± 1.61 , and 3.61 ± 2.38 $\mu\text{gC}/\text{m}^3$ for the robust-regression estimate of indoor ambient OC, modeled indoor ambient OC, and measured outdoor OC, respectively). Notably, in the two epidemiologic studies discussed above (in which a version of this mass-balance model was used to estimate the fraction of *central-site* $\text{PM}_{2.5}$ found in study-subject homes) the spatial resolution of residential ambient $\text{PM}_{2.5}$ exposure estimates was identified as a possible source of error.¹⁵ A two-step approach involving a method to account for local-scale variability in outdoor $\text{PM}_{2.5}$ followed by the use of an outdoor-to-indoor transport model might offer the best results when predicting residential $\text{PM}_{2.5}$ exposures.

It should also be noted that AERs measured at each individual home were used as model inputs in our calculations; however, these data are not generally available for an epidemiologic study population. The Lawrence Berkeley National Laboratory Infiltration model, which accounts for air exchange due to air flow through cracks in a residence³⁵ has recently been refined to include natural ventilation through open windows^{13,36} and can

be used to calculate AER distributions for a study population using readily-available housing data (from the United States Census and the American Housing Survey) and routinely measured meteorological parameters. A refined version of the LBNL Infiltration model was used for this purpose in the two epidemiologic studies mentioned above.^{14,15}

4.5. Conclusions

The evaluation and refinement of an outdoor-to-indoor transport model using measured indoor and outdoor $PM_{2.5}$ species concentrations and AERs from the RIOPA study illustrates that the modeling tools presented here offer improvement over the use of outdoor $PM_{2.5}$ concentrations to estimate residential ambient $PM_{2.5}$ exposure. The level of model refinement and data required to facilitate the use of this model in large epidemiologic studies varies across $PM_{2.5}$ species. Accounting for AC and heating use and open windows led to reduced bias in predicted F values for sulfate at the individual-subject level, but this refinement was not needed for EC nor at the population level for sulfate. This refinement did not resolve the large model-measurement differences for OC. We did, however, find quantitative evidence that shifts in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport contribute significantly to variability in F . Our results suggest that the collection of human activity data or a method to predict these human activity patterns can lead to substantial improvements in individual-subject level residential ambient $PM_{2.5}$ exposure estimates. This work also highlights the need for a method to account for shifts in the gas-particle partitioning of ambient OC in outdoor-to-indoor transport models. While further refinements are recommended, this mass-balance model is a practical method that can be applied in large epidemiologic studies to predict residential ambient $PM_{2.5}$ exposures. The input parameters (i.e. k_{dep} , P_{filter})

provided here are based on a comprehensive assessment of PM_{2.5} species size distributions and their evaluation using RIOPA data provides confidence in this version of the mass-balance model as a robust tool for reducing exposure misclassification in epidemiologic studies.

4.6. References

1. Klepeis, N. E.; Nelson, W. C.; Ott, W. R.; Robinson, J. P.; Tsang, A. M.; Switzer, P.; Behar, J. V.; Stephen, C. H.; Engelman, W. H., The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *J. Exposure Sci. Environ. Epidemiol.* **2001** 11, 231 - 252.
2. Riley, W. J.; McKone, T. E.; Lai, A. C. K.; Nazaroff, W. W., Indoor particulate matter of outdoor origin: importance of size-dependent removal mechanisms. *Environ. Sci. Technol.* **2002**, 36, 200 - 207.
3. Nazaroff, W. W., Indoor particle dynamics. *Indoor Air* **2004**, 14, 175-183.
4. Sarnat, S. E.; Coull, B. A.; Ruiz, P. A.; Koutrakis, P.; Suh, H. H., The influences of ambient particle composition and size on particle infiltration in Los Angeles, CA, Residences. *J. Air Waste Manage. Assoc. Association* **2006**, 56, 186 - 196.
5. Hering, S. V.; Lunden, M. M.; Thatcher, T. L.; Kirchstetter, T. W.; Brown, N. J., Using regional data and building leakage to assess indoor concentrations of particles of outdoor origin. *Aerosol Sci. Technol.* **2007**, 41, 639 - 654.
6. Meng, Q. Y.; Spector, D.; Colome, S.; Turpin, B. J., Determinants of indoor exposure to PM_{2.5} of indoor and outdoor origin during the RIOPA study. *Atmos. Environ.* **2009**, 43, 5750 - 5758.
7. Chen, C.; Zhao, B., Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. *Atmos. Environ.* **2011**, 45, 275-288.
8. Chen, C.; Zhao, B.; Weschler, C. J., Indoor exposure to "outdoor PM₁₀": assessing its influence on the relationship between PM₁₀ and short-term mortality in U.S. cities. *Epidemiology* **2012**, 23, 870 - 878.

9. Meng, Q.; Williams, R.; Pinto, J. P., Determinants of the associations between ambient concentrations and personal exposures to ambient PM_{2.5}, NO₂, and O₃ during DEARs. *Atmos. Environ.* **2012**, 63, 109 - 116.
10. Zeger, S. L.; Thomas, D.; Dominici, F.; Samet, J. M.; Schwartz, J.; Dockery, D.; Cohen, A., Exposure measurement error in time-series studies of air pollution: Concepts and consequences. *Environ. Health Perspect.* **2000**, 108, 419 - 426.
11. Weisel, C. P.; Zhang, J.; Turpin, B. J.; Morandi, M. T.; Colome, S.; Stock, T. H.; Spektor, D. M.; Korn, L.; Winer, A.; Alimokhtari, S.; Kwon, J.; Mohan, K.; Harrington, R.; Giovanetti, R.; Cui, W.; Afshar, M.; Maberti, S.; Shendell, D., Relationship of Indoor, Outdoor and Personal Air (RIOPA) study: study design, methods and quality assurance/control results. *J. Exposure Sci. Environ. Epidemiol.* **2005**, 15, 123-137.
12. Turpin, B. J.; Weisel, C. P.; Morandi, M.; Colome, S.; Eisenreich, S.; Buckley, B., Relationships of Outdoor Indoor and Personal Air (RIOPA): part II. Analysis of concentrations of particulate matter species. Research Report (Health Effects Institute) **2007**, 130, 79-92.
13. Turpin, B.; Rich, D. Q.; Lunden, M. M.; Hodas, N.; Thevenet-Morrison, K.; Ohman-Strickland, P., Refined Exposure Surrogates for Ambient PM in Epidemiologic Studies: Accounting for Temporal/Spatial Variations in Infiltration, **2012**, Final Report for EPA Cooperative Agreement CR-83407201-1.
14. Baxter, L.; Burke, J.; Lunden, M.; Turpin, B. J.; Rich, D. Q.; Thevenet-Morrison, K.; Hodas, N.; Ozkaynak, H., Influence of human activity patterns and residential air exchange rates on modeled distributions of PM_{2.5} exposure compared to central-site monitoring data. *J. Exposure Sci. Environ. Epidemiol.* **2013**, 23, 241 - 247.
15. Hodas, N.; Turpin, B. J.; Lunden, M. M.; Baxter, L.; Özkaynak, H.; Burke J.; Ohman-Strickland, P.; Thevenet-Morrison, K.; Rich, D. Q., Refined ambient PM_{2.5} exposure surrogates and the risk of myocardial infarction. *J. Exposure Sci. Environ. Epidemiol.* **2013**, advance online publication, doi:10.1038/jes.2013.24.
16. Yamamoto, N.; Shendell, D. G.; Winer, A. M.; Zhang, J., Residential air exchange rates in three major US metropolitan areas: results from the Relationship Among Indoor, Outdoor, and Personal Air Study 1999–2001. *Indoor Air* **2010**, 20, 85-90.
17. Lai, A.C.K; Nazaroff, W.W., Modeling indoor particle deposition from turbulent flow onto smooth surfaces. *Journal of Aerosol Science* **2000**, 31, 463-476.
18. Liu, D.L., Nazaroff, W.W., Modeling pollutant penetration across building envelopes. *Atmospheric Environment* **2001** 35, 4451-4462.

19. Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahuvaroglu, T.; Morrison, J.; Giovis, C., A review and evaluation of intraurban air pollution exposure models. *J. Exposure Anal. Environ. Epidemiol.* **2005**, 15, 185 - 204.
20. Hoek, G.; Brunekreef, B.; Goldbohm, S.; Fischer, P.; van den Brandt, P. A., Association between mortality and indicators of traffic-related air pollution in the Netherlands. *Lancet* **2002**, 360, 1203 - 1209.
21. Hanley, J. T.; Ensor, D. S.; Smith, D. D.; Sparks, L. E., Fractional aerosol filtration efficiency of in-duct ventilation air cleaners. *Indoor Air* **1994**, 3, 169 - 178.
22. Polidori, A.; Turpin, B.; Meng, Q. Y.; Lee, J. H.; Weisel, C.; Morandi, M.; Colome, S.; Stock, T.; Winer, A.; Zhang, J.; Kwon, J.; Alimokhtari, S.; Shendell, D.; Jones, J.; Farrar, C.; Maberti, S., Fine organic particulate matter dominates indoor-generated PM_{2.5} in RIOPA homes. *J. Exposure Sci. Environ. Epidemiol.* **2006**, 16, 321 - 331.
23. Ott, W.; Wallace, L.; Mage, D., Predicting particulate (PM₁₀) personal exposure distributions using a random component superposition statistical model. *J. Air Waste Manage. Assoc.* **2000**, 50, 1390 - 1406.
24. Meng, Q. Y.; Turpin, B. J.; Polidori, A.; Lee, J. H.; Weisel, C. P.; Morandi, M.; Winer, A.; Zhang, J., PM_{2.5} of ambient origin: estimates and exposure errors relevant to PM epidemiology. *Environ. Sci. Technol.* **2005**, 39, 5105 - 5112.
25. Turpin, B. J.; Saxena, P.; Allen, G.; Koutrakis, P.; McMurry, P.; Hildemann, L., Characterization of the southwestern desert aerosol, Meadview, AZ. *J. Air Waste Manage. Assoc.* **1997**, 47, 344 - 356.
26. Miguel, M. H.; Eiguren-Fernandez, A.; Jaques, P. A.; Froines, J. R.; Grant, B. L.; Mayo, P. R.; Sioutas, C., Seasonal variation of the particle size distribution of polycyclic aromatic hydrocarbons and of major aerosol species in Claremont, California. *Atmos. Environ.* **2004**, 38, 3241-3251.
27. Naumova, Y. Y.; Offenberg, J. H.; Eisenreich, S. J.; Meng, Q. Y.; Polidori, A.; Turpin B. J.; Weisel, C. P.; Morandi, M. T.; Colome, S. D.; Stock, T. H.; Winer, A. M.; Alimokhtari, S.; Kwon, J.; Maberti, S.; Shendell, D.; Jones, J.; Farrar, C., Gas/particle distribution of polycyclic aromatic hydrocarbons in coupled indoor/outdoor atmospheres. *Atmos. Environ.* **2003**, 37, 703 - 719.
28. Lunden, M. M.; Kirchstetter, T. W.; Thatcher, T. L.; Hering, S. V.; Brown, N. J., Factors affecting the indoor concentration of carbonaceous aerosols of outdoor origin. *Atmos. Environ.* **2008**, 42, 5660 - 5671.
29. Shi, S., Zhao, B., Comparison of the predicted concentration of outdoor originated indoor Polycyclic Aromatic Hydrocarbons between a kinetic partition model and a linear

instantaneous model for gas-particle partition. *Atmospheric Environment* **2012**, 59, 93-101.

30. Goldstein, A. H.; Galbally, I. E., Known and unexplored organic constituents in the Earth's atmosphere. *Environ. Sci. Technol.* **2007**, 41, 1514-1521.

31. Naumova, Y. Y.; Eisenreich, S. J.; Turpin, B. J.; Weisel, C. P.; Morandi, M. T.; Colome S. D.; Totten, L. A.; Stock, T. H.; Winer, A. M.; Alimokhtari, S.; Kwon, J.; Shendell, D.; Jones, J.; Maberti, S.; Wall, S. J., Polycyclic aromatic hydrocarbons in the indoor and outdoor air of three cities in the U.S. *Environ. Sci. Technol.* **2002**, 36, 2552-2559.

32. Pankow, J. F., An absorption model of gas/particle partitioning of organic compounds in the atmosphere. *Atmos. Environ.* **1994**, 28, 185-188.

33. Zhou, B.; Zhao, B., Population inhalation exposure to Polycyclic Aromatic Hydrocarbons and associated lung cancer risk in Beijing region: contributions of indoor and outdoor sources and exposure. *Atmospheric Environment* **2012**, 62, 472-480.

34. Weschler, C.J., Chemistry in indoor environments: 20 years of research. *Indoor Air* **2011**, 21, 205-218.

35. Sherman, M. H.; Grimsrud, D. T., Measurement of infiltration using fan pressurization and weather data. Lawrence Berkeley National Laboratory Report. LBNL-10852, Berkeley, CA, **1980**.

36. Breen, M. S.; Breen, M.; Williams, R. W.; Schultz, B. D., Predicting residential air exchange rates from questionnaires and meteorology: model evaluation in central North Carolina. *Environ. Sci. Technol.* **2010**, 44, 9349 - 9356.

	EC	Sulfate		OC		
		Mode 1	Mode 2	Mode 1	Mode 2	Mode 3
Mass Fraction in Mode	1.0	0.2	0.8	0.4	0.12	0.48
Particle Diameter (μm)	0.08	0.2	0.7	0.08	0.2	0.7
k_{dep} (h^{-1})	0.05	0.05	0.13	0.05	0.05	0.13
P	0.80/0.90 ^b	0.80	0.80	0.80/0.90 ^b	0.80	0.80
P_{filter}^a	0.90	0.90	0.65	0.90	0.90	0.65
P_{window}^a	1.0	1.0	1.0	1.0	1.0	1.0

^aRefined model: activities selected as significant predictors of variability in model-measurement differences were included in the refined model.

^bRefined model: greater penetration efficiency of ultrafine-mode particles was accounted for in the refined model

Table 4-1. Ambient PM_{2.5} species particle diameters and associated particle deposition loss rate coefficients (k_{dep}), penetration efficiencies (P), central heating and air conditioning filter penetration efficiencies (P_{filter}), and penetration efficiencies for homes with open windows (P_{window}) for elemental carbon (EC), sulfate, and organic carbon (OC).

Sulfate (n = 203)				
Selection Step	Activity	Partial R ²	Model R ²	P
1	Central Air Conditioning	0.20	0.20	< 0.0001
2	Open Windows	0.09	0.29	< 0.0001
3	Central Heating	0.02	0.31	0.027

Table 4-2. Multiple linear regression (MLR) analysis investigating the contribution of human activities to variability in model-measurement differences for sulfate. Partial R² describes the variance in model-measurement differences explained by each human-activity variable individually. Model R² describes the total variance in these differences described by the full MLR model at each selection step. Indoor sulfate of outdoor origin is modeled. Measurements are of total indoor sulfate. Previous work suggests indoor sulfate is predominately of outdoor origin (Sarnat et al., 2006).

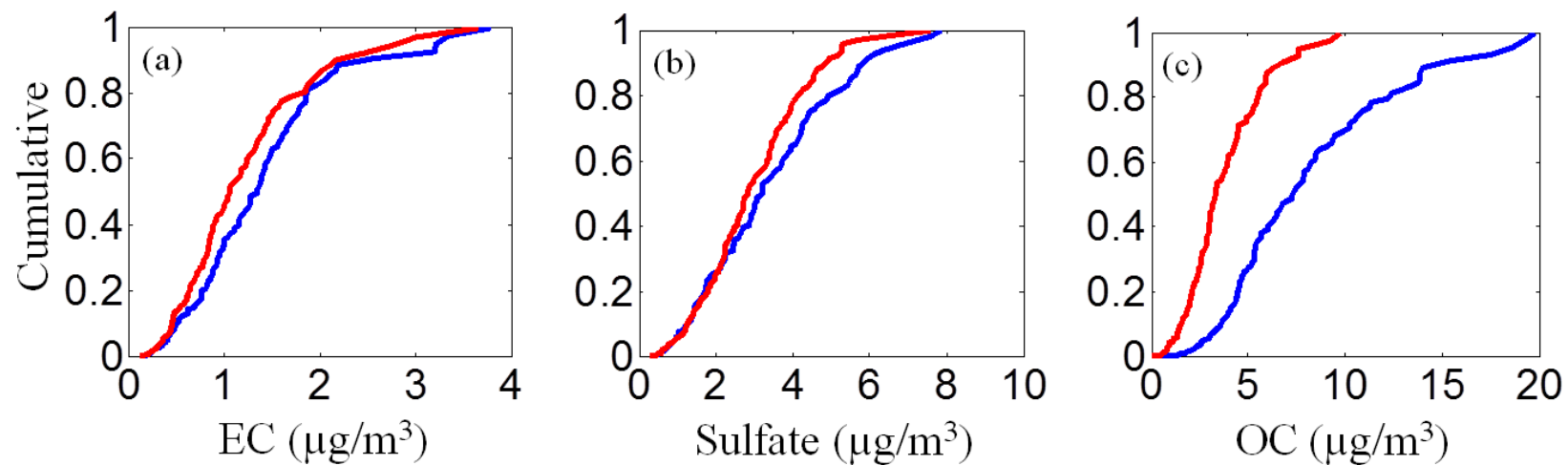


Figure 4-1. Cumulative distributions: measured indoor species (blue) and indoor species of ambient origin modeled with the initial model (red): (a) elemental carbon (EC), (b) sulfate, and (c) organic carbon (OC).

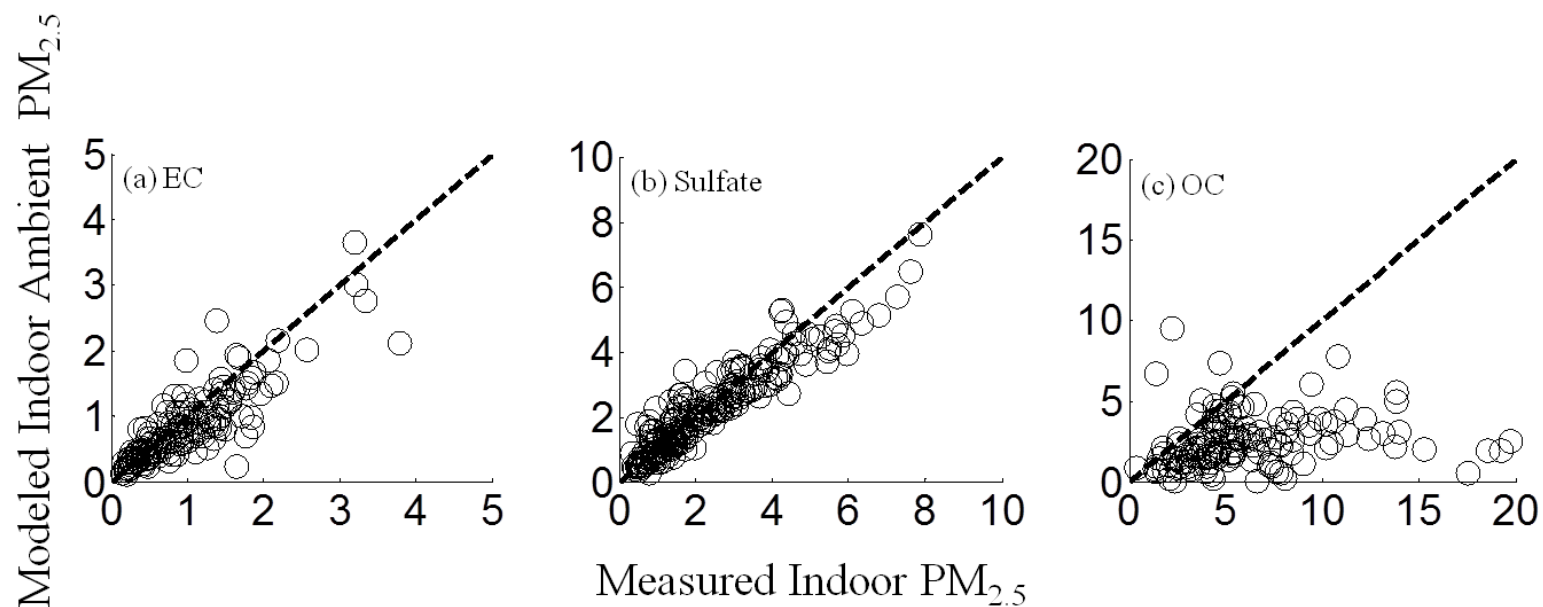


Figure 4-2. Indoor PM_{2.5} species concentrations ($\mu\text{g}/\text{m}^3$): modeled with the initial model and measured (a) elemental carbon (EC), (b) sulfate, and (c) organic carbon (OC). The dashed line is the 1:1 line. Note the model predicts indoor concentrations of ambient origin, whereas measurements also include the contribution from indoor sources.

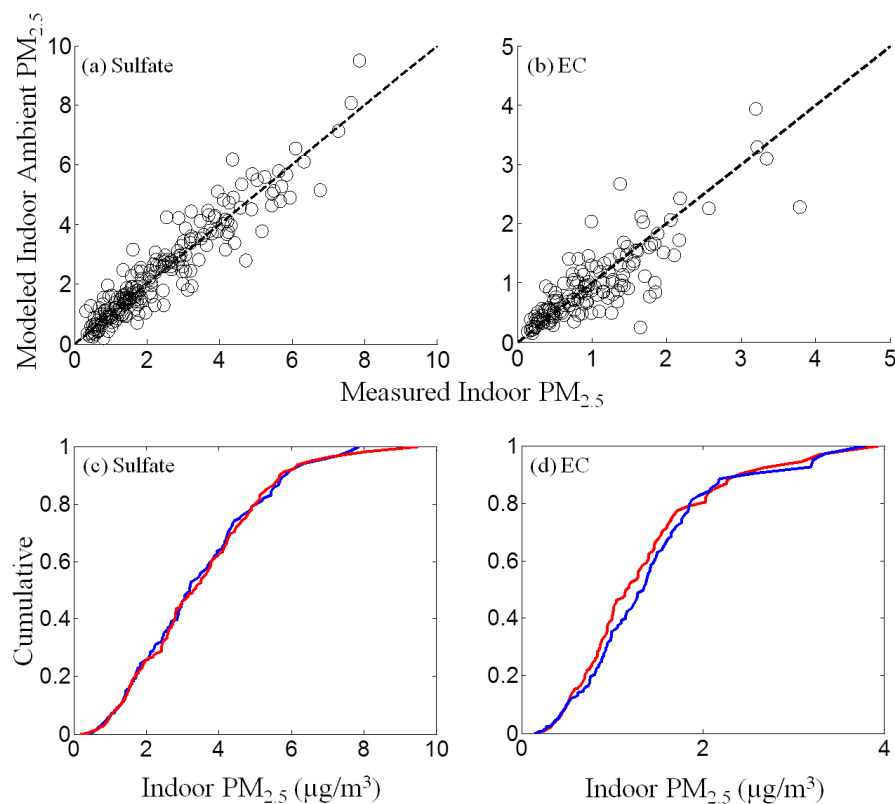


Figure 4-3. Indoor PM_{2.5} species concentrations ($\mu\text{g}/\text{m}^3$): modeled with the refined model and measured (a) sulfate and (b) elemental carbon (EC). The dashed line is the 1:1 line. Cumulative distributions: measured indoor species (blue) and indoor species of ambient origin modeled with the refined model (red): (c) sulfate and (d) elemental carbon (EC). Note that for sulfate the refined model accounts for human activities, while for EC it accounts for the greater penetration efficiency of ultrafine-mode particles.

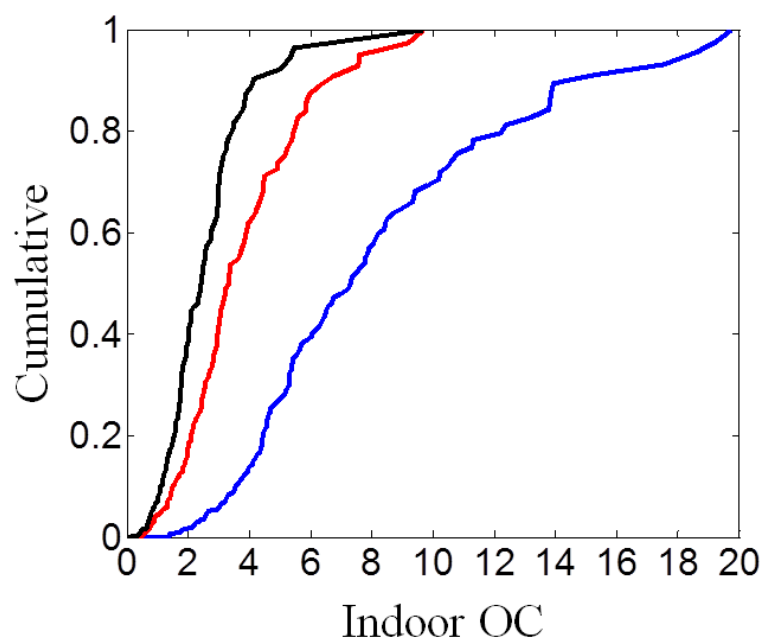


Figure 4-4. Cumulative distributions of measured indoor organic carbon (OC) concentrations (blue), indoor OC of ambient origin estimated with the mass-balance model (red), and indoor OC of ambient origin estimated with the population average F value calculated using robust regression (black).

Chapter 5. Shifts in the gas-particle partitioning of ambient organics with transport into the indoor environment

Material in this chapter has been published previously as:

Hodas, N.; Turpin, B. J., Shifts in the gas-particle partitioning of ambient organics with transport into the indoor environment, *Aerosol Sci. Technol.* **2014**, 48, 271 - 281.

5.1. Abstract

Predicting indoor exposures to ambient organic aerosol (OA) is complicated by shifts in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport. This analysis aims to quantify the effect of changes in temperature and OA loading on the gas-particle partitioning of ambient organics transported indoors and explores whether accounting for shifts in partitioning closes the gap between measured indoor ambient OA concentrations and indoor concentrations calculated in a previous analysis using a model that accounts for only the physical processes that influence outdoor-to-indoor transport. Changes in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport were calculated for 167 homes using measured temperatures and OA concentrations and published OA volatility distributions. Initially, it was assumed that ambient OA could be represented with a single volatility distribution. The analysis was then repeated treating ambient OA as the sum of distinct components derived from factor analysis of aerosol mass spectra (e.g. hydrocarbon-like OA, oxygenated OA), each with a distinct volatility distribution. The sensitivity of these calculations to uncertainty in the thermodynamic properties of ambient OA was also evaluated by varying the enthalpy of vaporization. Partitioning shifts were sensitive to enthalpy-of-vaporization assumptions and resulted in changes in indoor ambient OA concentrations of 13 - 27%. The

calculations indicate that phase changes are important determinants of residential exposure to ambient OA and are of sufficient magnitude to close the gap between measured and modeled indoor concentrations of ambient OA.

5.2. Introduction

Ambient PM_{2.5} exposure mitigation strategies, risk assessment and health studies all benefit from accurate exposure prediction. Because people spend about 70% of time in their homes,¹ the residence is an important setting for exposure to PM_{2.5} of ambient (outdoor) origin. The efficiency with which ambient PM_{2.5} penetrates into and persists in indoor air, and thus the fraction of ambient PM_{2.5} that people are exposed to in their homes, varies across PM_{2.5} species due to differences in particle size distributions and thermodynamic properties.²⁻⁵ For chemically non-reactive species (e.g. sulfate, elemental carbon), outdoor-to-indoor transport is governed by the physical losses associated with penetration across the building envelope and deposition indoors.⁶⁻⁹ It has been demonstrated for nitrate and polycyclic aromatic hydrocarbons (PAHs), however, that semi-volatile species can also undergo phase changes with outdoor-to-indoor transport due to changes in temperature, surface area, and the availability of particulate matter for sorption.¹⁰⁻¹² Organics are a major component of outdoor and indoor-generated PM_{2.5}.¹³ While previous work provides evidence that organics also undergo phase changes with outdoor-to-indoor transport,^{9,12} to my knowledge this process has not been incorporated into models used to predict ambient PM_{2.5} exposure.

The thermodynamic principles governing gas-particle partitioning of ambient organics are well established.¹⁴ Increasing temperature increases organic vapor pressures, shifting organic mass from the particle phase to the gas phase. On the other hand, organic

aerosol (OA) provides a medium for sorption of semi-volatile organics. As a result, increasing concentrations of OA (for example with the introduction of indoor sources) shift organic mass from the gas to the particle phase. While the thermodynamics are understood, modeling these processes for atmospheric organics is hindered by their complexity. Atmospheric organic matter is comprised of thousands of compounds with largely unknown identities and a broad range of thermodynamic properties.^{15,16} Thus, explicitly modeling the gas-particle partitioning of all individual organic compounds comprising ambient OA is not practical. The partitioning behavior of atmospheric organics is often parameterized with a volatility basis set (VBS), which treats organics as a distribution of compounds binned by their volatilities.¹⁷ More specifically, the VBS is a distribution of saturation vapor pressures expressed in concentration units (C^*) with \log_{10} spacing that span the range of atmospherically-relevant organic saturation concentrations.¹⁷ A volatility distribution with C^* on the x-axis and total (gas + particle) organic mass (OM) on the y-axis can be used to describe the gas-particle partitioning of ambient organics as a function of temperature and organic aerosol loading.¹⁷

In this chapter, the first study to utilize this volatility distribution to model shifts in the gas-particle partitioning of ambient organics with transport into the indoor environment is presented. Changes in temperature, and that indoor sources add considerable OA for sorptive partitioning are considered. Recently, volatility distributions were generated for ambient OA and OA components measured in regions dominated by anthropogenic $\text{PM}_{2.5}$ sources.¹⁸ These volatility distributions were utilized to explore the thermodynamic behavior of ambient OA with outdoor-to-indoor transport

for multiple homes located in three urban regions of the United States: Elizabeth, NJ, Los Angeles County, CA, and Houston, TX.

Notably, Chapter 4 demonstrated that shifts in gas-particle partitioning were significant contributors to variability in measured indoor OA for these homes and that not accounting for these shifts significantly contributed to error in predicted indoor concentrations of ambient OA.⁹ In that study, indoor concentrations of ambient OA were calculated with a single compartment mass balance model that accounted for the physical processes that govern outdoor-to-indoor transport (i.e. the efficiency with which particles penetrate across the building envelope, depositional losses indoors, and losses in the filters of HVAC systems), but not shifts in gas-particle partitioning. A comparison between these modeled indoor concentrations of ambient OA and a statistical estimate of the *measured* indoor OA that could be attributed to outdoor sources demonstrated a lack of closure. As noted in Chapter 4, an estimate of the measured ambient OA indoors was needed because the indoor concentrations measured in these occupied homes include OA of both outdoor and indoor origin.⁹ This chapter uses volatility distributions to illustrate and quantify the effect of changes in temperature and OA loading on the gas-particle partitioning of ambient organics found indoors and explores whether these changes in OA concentrations are of adequate magnitude to achieve model-measurement closure.

5.3. Methods

5.3.1. Overview

As is described in detail below, the change in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport was calculated for homes sampled during the Relationships of Indoor, Outdoor, and Personal Air (RIOPA) Study. Changes

in partitioning due to indoor-outdoor temperature differences were considered, as was the fact that indoor sources emit OA to which ambient organics can sorb. The extent to which positive (net sorption) and negative (net volatilization) shifts were driven by changes in OA loading and changes in temperature was examined. Finally, the magnitude of shifts in partitioning were compared with the size of the measurement-model gap discussed above. This work was accomplished using measured temperatures and OA concentrations from RIOPA^{19,20} and published volatility distributions for ambient OA.¹⁸ It was assumed that volatility distributions generated for ambient OA measured during the MILAGRO campaign in Mexico City are representative of the volatility distributions of the ambient OA at each of the three RIOPA sites. This assumption is supported by the fact that mass thermograms (measurements of the fraction of mass remaining in the particle phase as a function of temperature) for the ambient OA in Riverside, CA are similar to those for Mexico City, suggesting similarities in the thermodynamic properties of ambient OA for regions dominated by anthropogenic OA sources.^{18,21} Initially, the volatility of ambient OA was represented with a single volatility distribution using an enthalpy of vaporization (ΔH_{vap}) of 100 kJ/mol, a value considered reasonable when ambient OA is treated as a mixture of compounds with a range of volatilities, as is the case for the VBS.^{17,18} However, a preferred approach would be one in which differences in the sources and formation mechanisms (and thus differences in thermodynamic properties) of ambient OA with season and geographic region are accounted for. Thus, this analysis was repeated treating ambient OA as the sum of several distinct components, each with its own distinct volatility distribution, derived from factor analysis of aerosol mass spectra (assuming an ΔH_{vap} of 100 kJ/mol for all components). In this alternative analysis,

volatility distributions in different cities and seasons were represented by an average of the of the OA component volatility distributions weighted by the mass fractions of the OA components in each of those seasons and locations. The sensitivity of these calculations to uncertainty in the thermodynamic properties of ambient OA was also evaluated by repeating these analyses assuming an ΔH_{vap} of 50 kJ/mol.

5.3.2. RIOPA Study Measurements

RIOPA study measurements are described in detail by Weisel et al.¹⁹ and Turpin et al.²⁰ Measurements included indoor and outdoor particulate organic carbon (OC) samples collected for 173 homes. Briefly, OC concentrations were measured by thermal-optical transmittance with a Sunset Carbon Analyzer using the NIOSH temperature protocol and were corrected for the adsorption of gas-phase semi-volatile organic compounds on the quartz fiber filters used for collection. OC concentrations ($\mu\text{gC}/\text{m}^3$) were converted to OA concentrations ($\mu\text{g}/\text{m}^3$) assuming an OA:OC ratio of 1.4. A total of 167 homes had all data required for the calculations conducted in this chapter (indoor and outdoor temperature and OC concentrations). Summary statistics for the RIOPA data used in our calculations are shown in Table 5-1.

5.3.3. Volatility Distributions

The calculation of the volatility distributions for ambient OA is described in detail elsewhere.¹⁸ Briefly, during the MILAGRO campaign in Mexico City, aerosol was pulled through a thermodenuder that heated the aerosol stepwise between ambient temperature and 230°C. Organic mass fragments remaining in the particle phase at each temperature were measured in an Aerosol Mass Spectrometer (TD-AMS).²¹ These measurements and a detailed model of aerosol evaporation in the TD-AMS system were used to generate

volatility distributions for ambient OA at 25°C and the campaign-average OA loading (17 $\mu\text{g}/\text{m}^3$).¹⁸ The fitting approach involved fixing two sets of free parameters: (1) the range of C^* bins that described the volatility distributions of the measured aerosol and (2) the total OM (gas-phase + particle-phase) in each of those C^* bins ($C_{i,tot}$). The authors used an iterative approach to determine the range of C^* bins that maximized agreement between modeled aerosol evaporation and the TD-AMS measurements. $C_{i,tot}$ in each of the bins was calculated assuming an exponential relationship between C^* and $C_{i,tot}$ and, again, adjusting this relationship iteratively until model-measurement agreement was maximized. Volatility distributions were calculated for a range of ΔH_{vap} assumptions, as well as for OA components derived from factor analysis of aerosol mass spectra: hydrocarbon-like OA (HOA), biomass burning OA (BBOA), oxygenated OA (OOA), semi-volatile oxygenated OA (SV-OOA), and low-volatility oxygenated OA (LV-OOA).¹⁸ Factor analysis of AMS spectra categorizes OA based on the temporal variability of measured compound mass fragments. As a result, any given OA component is comprised of compounds with similar sources, formation mechanisms, and physiochemical.²²⁻²⁷ Thus, while the contribution of any given OA component will vary temporally and spatially,²⁴ the physiochemical properties of this OA component are likely to be similar across seasons and geographic regions.

5.3.4. Shifts in Gas-Particle Partitioning with Outdoor-to-Indoor Transport

In the initial analysis in which ambient OA volatility was represented with a single distribution, volatility distributions were first calculated for the ambient conditions (i.e. temperature and OA loading) measured outside of each RIOPA home using the parameters presented in Cappa and Jimenez¹⁸ (Appendix D1). Volatility distributions for

the average outdoor conditions for the homes considered here (18.2°C, 4.87 µg/m³) are shown in Figure 5-1. The full bar for each saturation vapor pressure bin (C_i^*) indicates the total OM (gas + particle phase) in that volatility (saturation vapor pressure) bin. The shaded region indicates the fraction of that OM that is in the particle phase (ξ_i) assuming absorptive partitioning into a single, well-mixed condensed phase:

$$\xi_i = \left(1 + \frac{C_i^*}{C_{OA}}\right)^{-1} \quad (1)$$

where C_{OA} is the OA mass concentration. Organic aerosol concentrations measured outside each RIOPA home ($C_{OA,out}$) were used as inputs for C_{OA} in equation (1) in these calculations.

As noted above, the published volatility-distribution parameters (Appendix D1)¹⁸ are for an ambient temperature of 25°C. The distribution of C^* bins for the temperature measured outside of each home was calculated assuming that the temperature-dependence of C^* can be described by the Clausius-Clapeyron equation:

$$C^*(T) = C^*(T_{ref}) \left(\frac{T_{ref}}{T}\right) \exp \left[-\frac{\Delta H_{vap}}{R} \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right] \quad (2)$$

where T_{ref} is 25°C, T is the temperature measured outside each RIOPA home, and R is the ideal gas constant.¹⁷ The same ΔH_{vap} assumptions as were used to construct the volatility distributions were also used to calculate changes in C^* with temperature.

It should be noted that these volatility distributions were used only to simulate the change in partitioning behavior of the ambient organics. In other words, the volatility distributions were used to calculate the *fraction* of total ambient OM (gas and particle phase) that was in the condensed phase (ξ_{tot}) outside of each RIOPA home:

$$\xi_{tot} = \frac{\sum_i (\xi_i \times C_{i,tot})}{\sum_i (C_{i,tot})} \quad (3)$$

This calculated fraction of OM in the particle phase outside of each RIOPA home ($\xi_{tot,out}$) and the measured outdoor OA concentration were then used to determine the total OM in the gas plus particle phase outside of each home:

$$C_{tot} = \frac{C_{OA,out}}{\xi_{tot,out}} \quad (4)$$

Using the same methods, the fraction of that ambient C_{tot} that would be found in the particle phase for the temperature and OA loading measured *inside* each RIOPA home ($\xi_{tot,in}$) was then calculated. In other words, the new equilibrium partitioning of the ambient OM after it was transported into the indoor environment was simulated. Note that measured indoor OA concentrations, which were used to estimate the total OA loading indoors (i.e. C_{OA} in equation 1), include emissions from indoor sources into which ambient organics can partition, as well as the fraction of ambient OA that has penetrated into and persisted in indoor air. It was assumed that there was no change in the thermodynamic properties of the ambient organics with outdoor-to-indoor transport (i.e. the same volatility distribution could be used to represent ambient OA before and after transport indoors). Possible limitations of this assumption are discussed in the Results and Discussion section (Section 5.4). Indoor concentrations of *ambient* OA ($C_{OA,in,amb}$) were calculated by multiplying $\xi_{tot,in}$ by C_{tot} . Changes in OA mass due to re-partitioning with outdoor-to-indoor transport were calculated by subtracting measured outdoor OA concentrations from those calculated after repartitioning to indoor conditions (i.e., temperature and OA loading):

$$\Delta C_{OA} = C_{OA,in,amb} - C_{OA,out} \quad (5)$$

In a more sophisticated analysis, ambient OA was treated as a mixture of components derived from factor-analysis of aerosol mass spectra. In these chemically-

resolved calculations, ΔH_{vap} is assumed to be uniform across the OA components (100 kJ/mol); however, the volatility distribution for total OA varies across locations and seasons because the mass fractions of the OA components differ and each component has a unique volatility distribution (Figures 5-1c - 5-1g). While measurements of OA outside of each home were available, measurements of AMS OA components were not. Thus, measured OA component mass fractions for Riverside, New York City, and Houston²⁸⁻³⁵ were used to apportion measured OA between the components for LA County, Elizabeth, and Houston RIOPA homes, respectively (Appendix D2). The measured mass fractions are season-specific and, thus, only homes that were sampled during the seasons for which OA component mass fractions were available were included in the calculations that utilized the volatility distributions for OA components. For Houston and Riverside, only warm season (May - October) mass fractions were available, while for New York, component mass fraction measurements were available for both the warm and cool (November - April) seasons (Appendix D2). Aggregate volatility distributions for each RIOPA region (i.e. those calculated as an average of these component-specific distributions weighted by the mass fraction of each component) are provided in Appendix D4. Volatility distributions were calculated for the ambient conditions (temperature and OA loading) measured outside each RIOPA home for each of the OA components using the OA-component-specific parameters presented in Cappa and Jimenez¹⁸ (Appendix D1). Note that for these calculations, outdoor concentrations of each OA component (rather than the total outdoor OA concentrations) are used as inputs for C_{OA} in equation 1. These were calculated by multiplying the component mass fractions in Appendix D2 by the OA measured outside of each RIOPA home. Shifts in the gas-particle partitioning of

each of these OA components with outdoor-to-indoor transport were then calculated using the same methods as described above. The total ΔC_{OA} due to shifts in partitioning for each RIOPA home was calculated by summing ΔC_{OA} across all of the OA components.

In order to determine the data most important for predicting shifts in gas-particle partitioning with outdoor-to-indoor transport, whether changes in temperature or changes in OA loading were the dominant drivers of variability in calculated gas-particle partitioning shifts was then explored. This will clarify what measurements and data are most needed in order to incorporate this process into predictive ambient PM exposure models. For example, can shifts in partitioning largely be predicted knowing only the indoor temperature or is it necessary to also characterize indoor OA emission rates? Using multiple linear regression (MLR) with stepwise selection ($\alpha = 0.15$ for variable entrance and removal threshold; SAS version 9.3, SAS Inc., Cary, NC), ΔC_{OA} was regressed on the indoor-outdoor temperature difference and the indoor-outdoor difference in OA mass loading, which was calculated by subtracting measured outdoor OA mass concentrations from measured indoor OA mass concentrations. This analysis systematically evaluates the predictive capability afforded by including either one or both variables. It should again be kept in mind that the total OA loading measured inside each home includes OA emitted or formed inside the home and is a different value than the OA of *ambient* origin calculated above. Variance inflation factors indicated that there was no correlation between indoor-outdoor differences in temperature and OA loading. For unoccupied homes, we would expect these two quantities to be correlated because temperature influences the fraction of OM in the particle phase;¹⁴ however, for these

occupied homes, indoor sources of organics are the driving force for variability in measured indoor OA mass concentrations.¹³ Outliers were identified with a student's t-test and were removed if $t > 2$. The MLR analysis was conducted for each ΔH_{vap} assumption, as well as for each urban region in order to explore variability in partitioning shifts across climatic regions.

5.3.5. Model-Measurement Closure

In order to explore whether changes in OA concentrations due to shifts in partitioning with outdoor-to-indoor transport can explain, at least in part, the gap between measured and modeled indoor ambient OA concentrations observed in the previous study discussed above, the previously calculated distribution of model error (Chapter 4) was compared with the distributions of ΔC_{OA} calculated here. In Chapter 4, (1) indoor concentrations of ambient OA calculated for RIOPA homes using an outdoor-to-indoor transport model that accounted only for home ventilation and physical loss processes (i.e. air exchange rates, particle penetration efficiencies, and depositional losses) and (2) a statistical estimate of the OA measured indoors that could be attributed to outdoor sources were compared. An estimate of the measured indoor OA of ambient origin was required because indoor OA measurements included ambient OA that had penetrated and persisted indoors, as well as OA emitted by indoor sources. It was concluded that remaining differences between modeled values and the statistical estimate of measured indoor OA of ambient origin could likely be attributed to phase changes with outdoor-to-indoor transport. If shifts in partitioning explain this closure gap, it is expected that the magnitude of ΔC_{OA} is equal in magnitude, but opposite in sign of the model-measurement disagreement (i.e. ΔC_{OA} would offset model error). In order to compare these quantities

directly, values of model error (modeled - measured) were multiplied by -1. Thus, if the distributions of "model error" and ΔC_{OA} are similar, shifts in partitioning with outdoor-to-indoor transport could plausibly close this model-measurement disagreement. The distribution of model error was compared to all calculations of ΔC_{OA} (i.e. those assuming $\Delta H_{vap} = 100$ kJ/mol, those assuming $\Delta H_{vap} = 50$ kJ/mol, and those treating ambient OA as a mixture of factor-analysis components).

5.4. Results and Discussion

In the main analysis (i.e. $\Delta H_{vap} = 100$ kJ/mol), partitioning shifts resulted in a loss of OA mass for 56% of homes (i.e., net volatilization; shifts from the particle phase towards the gas phase), and changes in ambient organic aerosol concentrations due to these shifts in partitioning (ΔC_{OA}) ranged from -4.6 to 2.4 $\mu\text{g}/\text{m}^3$ for individual homes (Figure 5-2). Negative values indicate net volatilization and positive indicate net absorption with outdoor-to-indoor transport. Calculated shifts in partitioning with outdoor-to-indoor transport were highly sensitive to ΔH_{vap} assumption (Figure 5-2). Both the magnitudes and the direction of partitioning shifts varied across these analyses. While the main analysis resulted in a loss of OA mass for 56% of homes, partitioning shifts resulted in an increase in OA mass for 61% of homes in the sensitivity analysis ($\Delta H_{vap} = 50$ kJ/mol). Values of ΔC_{OA} for individual homes spanned a wider range in the sensitivity analysis (-10.5 to 5.7 $\mu\text{g}/\text{m}^3$) compared to the main analysis. Outdoor-to-indoor transport resulted in an absolute change in C_{OA} (i.e. net absorption or volatilization) of only 13%, on average, in the main analysis, while a 27% change, on average, was observed in the sensitivity analysis.

Differences with ΔH_{vap} assumptions can partially be attributed to the fact that shifts in C^* with temperature are dependent on ΔH_{vap} (equation 2). However, this effect is relatively small.¹⁸ The differences in gas-particle partitioning shifts are mostly the result of differences in the volatility distributions for these two ΔH_{vap} assumptions. The volatility distribution generated for $\Delta H_{vap} = 100$ kJ/mol has more mass in lower C^* bins.¹⁸ In other words, when an ΔH_{vap} of 100 kJ/mol is assumed, lower volatility material comprises a larger fraction of the OA and, thus, this OA is less sensitive to indoor-outdoor differences in OA loading than that with an ΔH_{vap} of 50 kJ/mol. As is evident from Figure 5-3, which shows shifts in gas-particle partitioning that would result if only indoor-outdoor temperature differences or only indoor-outdoor differences in OA loading were considered, the change in OA associated with temperature is relatively similar across ΔH_{vap} values, but the OA with an ΔH_{vap} of 50 kJ/mol shows a much greater sensitivity to changes in OA loading with outdoor-to-indoor transport.

This effect is also evident in our MLR results (Table 5-2). In the main analysis ($\Delta H_{vap} = 100$ kJ/mol), the change in temperature with outdoor-to-indoor transport was the dominant predictor of variability in ΔC_{OA} , explaining 45% of this variability ($R^2 = 0.45$). The indoor-outdoor difference in OA loading explained 24% ($R^2 = 0.24$) of the variability in ΔC_{OA} . When an ΔH_{vap} of 50 kJ/mol was assumed, however, the indoor-outdoor difference in OA loading was the dominant driver of variability in shifts in partitioning, explaining 51% of the variability in ΔC_{OA} ($R^2 = 0.51$). The indoor-outdoor temperature difference explained 23% of this variability ($R^2 = 0.23$). The MLR results indicate that both temperature and OA emissions inside homes are important for predicting shifts in partitioning of ambient organics with outdoor-to-indoor transport. Note that indoor-

outdoor temperature differences also affect ambient exposure prediction by affecting residential air exchange rates.

Geographic differences in the results (Figure 5-4, Table 5-2) demonstrate that heterogeneity across climatic regions in indoor-outdoor temperature differences and in the human activities that influence these differences (e.g. air conditioning (AC) and heating use) impact gas-particle partitioning shifts with outdoor-to-indoor transport. In the relatively moderate climate of Los Angeles County, where both heating and AC use were low,^{19,36} indoor temperatures were greater than outdoor temperatures for 98% of homes and the average indoor-outdoor temperature difference was 4.8°C. In Elizabeth, where some of the homes sampled in the winter had heating in use,^{19,36} indoor temperatures were again greater than outdoor temperatures for the vast majority of homes (96%), but the average indoor-outdoor temperature difference was about twice that observed in Los Angeles County (9.9°C). Finally in Houston, where central AC use was greater than for the other two urban regions,^{19,36} there was a decrease in temperature with indoor transport for 30% of homes and indoor-outdoor temperature differences ranged from -7.2°C to 12.8°C. Interestingly, Houston is the only region for which there was a net increase in OA mass (shifts from the gas phase to the particle phase) for the majority of homes (64% compared to 31 and 30% for Los Angeles County and Elizabeth; Figure 5-4), indicating that human activities in addition to those associated with indoor emissions of OA (i.e., heating and AC use) can influence gas-particle partitioning. The greater range of indoor-outdoor temperature differences in Elizabeth and Houston likely explains why partitioning shifts were dominantly driven by changes in temperature for those two regions, while the change in OA loading was the dominant predictor of shifts in

partitioning for Los Angeles County homes (Table 5-2). While not explored here directly because of sample-size restrictions, these results also indicate that there are likely seasonal differences in shifts in gas-particle partitioning (and OA exposure) in regions for which there is seasonal heterogeneity in indoor-outdoor temperature differences and in heating and AC use.

Shifts in gas-particle partitioning calculated for RIOPA homes using the volatility distribution generated for OA components were qualitatively similar to those calculated in the main analysis ($\Delta H_{vap} = 100$ kJ/mol); however, this aerosol demonstrated slightly lower sensitivity to temperature and OA-loading changes with outdoor-to-indoor transport (Figure 5-5). These shifts resulted in an absolute change in C_{OA} (i.e. net absorption or volatilization) of 11%, on average, compared to 13% for the main analysis. HOA was most sensitive to changes in temperature and OA loading with outdoor-to-indoor transport (20% change in OA concentrations, on average), followed by "other" OA (13% change on average), and OOA (12%, on average; Figure 5-5). Indoor transport resulted in a 7% change in C_{OA} , on average, for SV-OOA, whereas outdoor-to-indoor transport did not induce shifts in partitioning for LV-OOA (Figure 5-5).

While there is uncertainty in the thermodynamic properties of ambient OA, we can speculate about which calculations of ΔC_{OA} are most realistic. Partitioning models in which total ambient OA is represented by one or two compounds require ΔH_{vap} values ≤ 50 kJ/mol to reproduce partitioning behavior observed in chamber studies and in the field.^{17,18,37,38} However, when a wider range of thermodynamic properties is considered in the representation of ambient OA (e.g. use of a volatility basis set as is the case here versus use of a two-product model) a more realistic ΔH_{vap} of 100 kJ/mol reproduces

observed partitioning behavior.^{17,18} As a result, the calculations for which ΔH_{vap} was assumed to be 100 kJ/mol are likely to be more representative of the partitioning behavior of ambient OA than those calculated assuming an ΔH_{vap} of 50 kJ/mol. Treating ambient OA as a mixture of components derived from factor analysis of aerosol mass spectra likely provides the most realistic and robust estimate of gas-particle partitioning shifts with indoor transport. As noted above, the physiochemical properties of any particular component (e.g., HOA) are likely similar across seasons and geographic regions because each component is comprised of species that are covariant due to having common sources or formation mechanisms.²²⁻²⁷ Notably, mass thermograms for OA components generated from TD-AMS measurements in Riverside, CA were similar to those for Mexico City, suggesting similarities in the thermodynamic properties of these OA components for regions dominated by anthropogenic OA sources.^{18,21} This supports the use of these volatility distributions for a wide range of locations. However, AMS data describing the mass fractions of each OA component are also required before these volatility distributions can be applied to predict indoor concentrations of ambient OA. Greater use of the AMS to measure outdoor aerosol in exposure studies will facilitate such exposure modeling efforts.

Distributions of measurement-model disagreement and calculated values of ΔC_{OA} are compared in Figure 5-6. Median values of residual model error and of ΔC_{OA} differ (note that median values of ΔC_{OA} are near zero); however, the distributions of model error and ΔC_{OA} are not significantly different. Thus, shifts in gas-particle partitioning provide a plausible explanation for the previously observed lack of closure (Chapter 4). As expected based on the discussion above regarding the most realistic and robust volatility-

distribution assumptions, distributions of model-measurement disagreement and ΔC_{OA} are closest in magnitude when an ΔH_{vap} of 100 kJ/mol is assumed and when ambient OA is treated as a mixture of AMS factors (e.g., HOA, OOA).

One limitation of this work is that the increased surface area in the indoor environment, compared to the outdoor environment, is not accounted for. Sorption to indoor surfaces such as carpets, wallboard, furniture, HVAC surfaces and even home occupants can be an important sink for gaseous SVOCs³⁹ and could result in a shift of OM of outdoor origin from the particle phase towards the gas phase in order to reach a new equilibrium.^{12,39} Thus, it can be concluded that interactions with indoor surfaces would result in greater evaporative losses of OA than were calculated for the RIOPA homes. As noted in the Methods section, we assumed that there was no change in the thermodynamic properties of the ambient organics with outdoor-to-indoor transport (i.e. the same volatility distribution could be used to represent ambient organics before and after transport indoors). Many factors such as particle size and composition, characteristics of cracks in the building shell, human activities (e.g. opening/closing windows), and compound volatility and reactivity contribute to variability in the efficiency with which pollutants penetrate and persist indoors.^{9,40} Reactive gases are expected to encounter larger losses and non-reactive gases smaller losses than particulate organics. Such differential losses may alter the ambient volatility distribution of ambient organics. For example, depletion of the particle phase and not the gas phase would result in smaller evaporative losses or larger increases in ambient OA due sorption to indoor-generated OA than the values calculated here. It should also be noted that the timescale

required for organic compounds to reach equilibrium partitioning between the gas and particle phases can vary over several orders of magnitude depending on vapor pressure.⁴¹

This work also does not consider interactions of ambient OA with gas-phase organics of indoor origin (e.g. terpenoids from cleaning products, PAHs from combustion sources)⁴²⁻⁴⁴. Weschler and Nazaroff³⁹ note that the partitioning of gas-phase organics emitted or formed indoors to OA of outdoor origin could both increase the mass concentrations of the particles transported indoors and alter their chemical composition. These interactions illustrate the difficulties that could arise in separating the ambient and non-ambient contributions to OA exposure. Further, the importance of liquid water as a partitioning and reaction medium has been demonstrated for atmospheric aerosols,⁴⁵ but its role in the indoor environment has not been explored. In the case of PAHs, Naumova et al.¹¹ found that the majority of variability (84.5%) in gas-particle partitioning indoors could be explained by PAH vapor pressure, indoor temperature, and the characteristics of the PM_{2.5} measured indoors (mass fractions of elemental and organic carbon).

5.5. Conclusions

Accounting for shifts in gas-particle partitioning of ambient organics with transport into the indoor environment improved model-measurement closure. Calculated shifts resulted in changes in OA mass of between 11 and 27%, on average, depending on the assumed ΔH_{vap} and whether OA was represented with a single volatility distribution or with a distribution generated assuming a mixture of AMS factor-analysis components. While uncertainties in the thermodynamic properties of ambient OA contribute to uncertainty in the magnitude and direction of partitioning shifts with outdoor-to-indoor transport, all calculations indicate that phase changes are important determinants of

residential OA exposure. Of the calculations presented here, treatment of ambient OA as a mixture of components with distinct sources, formation mechanisms, and physiochemical properties (AMS factors, e.g. HOA, OOA) is likely to offer the most robust estimates of shifts in partitioning with outdoor-to-indoor transport across seasons and geographic regions. Expanding the spatial and temporal coverage of TD-AMS measurements, volatility distribution calculations like those presented in Cappa and Jimenez,¹⁸ and OA component mass fraction measurements will help to identify seasonal and geographic variations in OA volatility distributions and will help facilitate estimates of shifts in partitioning with outdoor-to-indoor transport.

5.6. References

1. Klepeis, N. E.; Nelson, W. C.; Ott, W. R.; Robinson, J. P.; Tsang, A. M.; Switzer, P.; Behar, J. V.; Stephen, C. H.; Engelman, W. H., The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *J. Exposure Sci. Environ. Epidemiol.* **2001** 11, 231 - 252.
2. Lunden, M. M.; Thatcher, T. L.; Hering, S. V.; Brown, N. J., The use of time- and chemically-resolved particulate data to characterize the infiltration of outdoor PM_{2.5} into a residence in the San Joaquin Valley. *Environ. Sci. Technol.* **2003**, 37, 4724 -4732.
3. Sarnat, S. E.; Coull, B. A.; Ruiz, P. A.; Koutrakis, P.; Suh, H. H., The influences of ambient particle composition and size on particle infiltration in Los Angeles, CA, Residences. *J. Air Waste Manage. Assoc. Association* **2006**, 56, 186 - 196.
4. Meng, Q. Y.; Turpin, B. J.; Lee, J. H.; Polidori, A.; Weisel, C. P.; Morandi, M.; Colome, S.; Zhang, J.; Stock, A.; Winer, A., How does infiltration behavior modify the composition of ambient PM_{2.5} in indoor spaces? An analysis of RIOPA data. *Environ. Sci. Technol.* **2007**, 41: 7315-7321.
5. Hodas, N.; Meng, Q. Y.; Lunden, M.; Rich, D. Q.; Özkaynak, H.; Baxter, L.; Zhang, Q.; Turpin, B. J., Variability in the fraction of ambient fine particulate matter found indoors and heterogeneity in health effect estimates. *J. Exposure Sci. Environ. Epidemiol.* **2012**, 22, 448 - 454.

6. Riley, W. J.; McKone, T. E.; Lai, A. C. K.; Nazaroff, W. W., Indoor particulate matter of outdoor origin: importance of size-dependent removal mechanisms. *Environ. Sci. Technol.* **2002**, 36, 200 - 207.
7. Nazaroff, W. W., Indoor particle dynamics. *Indoor Air* **2004**, 14, 175-183.
8. Hering, S. V.; Lunden, M. M.; Thatcher, T. L.; Kirchstetter, T. W.; Brown, N. J., Using regional data and building leakage to assess indoor concentrations of particles of outdoor origin. *Aerosol Sci. Technol.* **2007**, 41, 639 - 654.
9. Hodas, N., Meng, Q.Y., Lunden, M.M., and Turpin, B.J., Toward refined estimates of ambient PM_{2.5} exposure: Evaluation of a physical outdoor-to-indoor transport model, **2013**, *Atmos. Environ.*, accepted.
10. Lunden, M. M.; Revzan, K. L.; Fischer, M. L.; Thatcher, T. L.; Littlejohn, D.; Hering, S. C.; Brown, N. J., The transformation of outdoor ammonium nitrate aerosols in the indoor environment. *Atmos. Environ.* **2003**, 37, 5633 - 5644.
11. Naumova, Y. Y.; Offenberg, J. H.; Eisenreich, S. J.; Meng, Q. Y.; Polidori, A.; Turpin B. J.; Weisel, C. P.; Morandi, M. T.; Colome, S. D.; Stock, T. H.; Winer, A. M.; Alimokhtari, S.; Kwon, J.; Maberti, S.; Shendell, D.; Jones, J.; Farrar, C., Gas/particle distribution of polycyclic aromatic hydrocarbons in coupled indoor/outdoor atmospheres. *Atmos. Environ.* **2003**, 37, 703 - 719.
12. Lunden, M. M.; Kirchstetter, T. W.; Thatcher, T. L.; Hering, S. V.; Brown, N. J., Factors affecting the indoor concentration of carbonaceous aerosols of outdoor origin. *Atmos. Environ.* **2008**, 42, 5660 - 5671.
13. Polidori, A.; Turpin, B.; Meng, Q. Y.; Lee, J. H.; Weisel, C.; Morandi, M.; Colome, S.; Stock, T.; Winer, A.; Zhang, J.; Kwon, J.; Alimokhtari, S.; Shendell, D.; Jones, J.; Farrar, C.; Maberti, S., Fine organic particulate matter dominates indoor-generated PM_{2.5} in RIOPA homes. *J. Exposure Sci. Environ. Epidemiol.* **2006**, 16, 321 - 331.
14. Pankow, J. F., An absorption model of gas/particle partitioning of organic compounds in the atmosphere. *Atmos. Environ.* **1994**, 28, 185-188.
15. Goldstein, A. H.; Galbally, I. E., Known and unexplored organic constituents in the Earth's atmosphere. *Environ. Sci. Technol.* **2007**, 41, 1514-1521.
16. Hallquist, M.; Wenger, J. C.; Baltensperger, U.; Rudich, Y.; Simpson, D.; Claeys, M.; Dommen, J.; Donahue, N. M.; George, C.; Goldstein, A. H.; Hamilton, J. F.;

Hermann, H.; Hoffman, T.; Iinuma, Y.; Jang, M.; Jenkin, M. E.; Jimenez, J. L.; Kiendler-Sharr, A.; Maenhaut, W.; McFiggans, G.; Mentel, Th. F.; Monod, A.; Prevot, A. S. H.; Seinfeld, J. H.; Surratt, J. D.; Szmigielski, R.; Wildt, J., The formation, properties and impact of secondary organic aerosol: current and emerging issues. *Atmos. Chem. Phys.* **2009**, 9, 5155–5236.

17. Donahue, N. M.; Robinson, A. L.; Stanier, C. O.; Pandis, S. N., Coupled partitioning, dilution, and chemical aging of semivolatile organics. *Environ. Sci. Technol.*, **2006**, 40: 2635-2643.

18. Cappa C. D.; Jimenez J., Quantitative estimates of the volatility of ambient organic aerosol. *Atmos. Chem. Phys.* **2010**, 10, 5409–5424.

19. Weisel, C. P.; Zhang, J.; Turpin, B. J.; Morandi, M. T.; Colome, S.; Stock, T. H.; Spektor, D. M.; Korn, L.; Winer, A.; Alimokhtari, S.; Kwon, J.; Mohan, K.; Harrington, R.; Giovanetti, R.; Cui, W.; Afshar, M., Maberti, S., Shendell, D., Relationship of Indoor, Outdoor and Personal Air (RIOPA) study: study design, methods and quality assurance/control results. *J. Exposure Sci. Environ. Epidemiol.* **2005**, 15, 123-137.

20. Turpin, B. J.; Weisel, C. P.; Morandi, M.; Colome, S.; Eisenreich, S.; Buckley, B., Relationships of Outdoor Indoor and Personal Air (RIOPA): part II. Analysis of concentrations of particulate matter species. Research Report (Health Effects Institute) **2007**, 130, 79-92.

21. Huffman, J. A.; Docherty, K. S.; Aiken, A. C.; Cubison, M. J.; Ulbrich, I. M.; DeCarlo, P. F.; Sueper, D.; Jayne, J. T.; Worsnop, D. R.; Ziemann, P.; Jimenez, J. L., Chemically-resolved volatility measurements from two megacity field studies. *Atmos. Chem. Phys.* **2009**, 9, 7161-7182.

22. Zhang, Q.; Alfarra, M. R.; Worsnop, D. R.; Allan, J. D.; Coe, H.; Canagaratna, M. R.; Jimenez, J., Deconvolution and quantification of hydrocarbon-like and oxygenated organic aerosols based on aerosol mass spectrometry. *Environ. Sci. Technol.* **2005**, 39, 4938-4952.

23. Zhang, Q.; Worsnop, D. R.; Canagaratna, M. R.; Jimenez, J. L., Hydrocarbon-like and oxygenated organic aerosols in Pittsburgh: insights into sources and processes of organic aerosols. *Atmos. Chem. Phys.* **2005**, 5, 3289-3311.

24. Zhang Q., Jimenez J.L., Canagaratna M.R., Allan J.D., Coe H., Ulbrich I., Alfarra, M. R.; Takami, A.; Middlebrook, A. M.; Sun, Y. L.; Dzpeina, K.; Dunlea, E.; Docherty, K.; DeCarlo, P. F.; Salcedo, D.; Onasch, T.; Jayne, J. T.; Miyoshi, T.; Shimono, A.; Hatakeyama, S.; Takegawa, N.; Kondo, Y.; Schneider, J.; Drewnick, F.; Borrmann, S.; Weimer, S.; Demerjian, K.; Williams, P.; Bower, K.; Bahreini, R.; Cottrell, L.; Griffin, R. J.; Rautiainen, J.; Sun, J. Y.; Zhang, Y. M.; Worsnop, D. R., Ubiquity and dominance of oxygenated species in organic aerosol in anthropogenically-

influenced Northern Hemisphere midlatitudes. *Geophys. Res. Lett.* **2007**, 34, doi:10.1029/2007GL029979.

25. Zhang, Q.; Jimenez, J. L.; Canagaratna, M. R.; Ulbrich, I. M.; Ng, N. L.; Worsnop, D. R.; Sun, Y., Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review. *Anal. Bioanal. Chem.* **2011**, 401, 3045-3067.

26. Lanz, V. A.; Alfarra, M. R.; Baltensperger, U.; Buchmann, B.; Hueglin, C.; Prevot, A. S. H., Source apportionment of submicron organic aerosols at an urban site by factor analytical modeling of aerosol mass spectra. *Atmos. Chem. Phys.*, **2007**, 7, 1503-1522.

27. Ulbrich, I. M.; Canagaratna, M. R.; Zhang, Q.; Worsnop, D. R.; Jimenez, J. L., Integration of organic components from Positive Matrix Factorization of aerosol mass spectrometric data. *Atmos. Chem. Phys.* **2009**, 9, 2891-2918.

28. Tanaka, P.; Riemer, D. D.; Chang, S.; Yarwood, G.; McDonald-Buller, E. C.; Apel, E. C.; Orlando, J. J.; Silva, P. J.; Jimenez, J. L.; Canagaratna, M. R.; Neece, J. D.; Mullins, C. B.; Allen, D. T., Direct evidence for chlorine-enhanced urban ozone formation in Houston, Texas. *Atmos. Environ.* **2003**, 37, 1393-1400.

29. Drewnick, F.; Schwab, J. J.; Jayne, J. T.; Canagaratna, M.; Worsnop, D. R.; Demerjian, K. L., Measurement of ambient aerosol composition during the PMTACS-NY using an aerosol mass spectrometer. Part I: mass concentrations. *Aerosol Sci. Technol.* **2004**, 38, 92-103

30. Drewnick, F.; Jayne, J. T.; Canagaratna, M.; Worsnop, D. R.; Demerjian, K. L., Measurement of ambient aerosol composition during the PMTACS-NY 2001 using an aerosol mass spectrometer. Part II: chemically speciated mass distributions. *Aerosol Sci. Technol.* **2004**, 38, 104 - 117.

31. DeCarlo, P. F.; Kimmel, J. R.; Trimborn, A.; Northway, M. J.; Jayne, J. T.; Aiken, A. C.; Gonin, M.; Fuhrer, K.; Horvath, T.; Docherty, K. S.; Worsnop, D. R.; Jimenez, J. L., Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. *Anal. Chem.* **2006**, 78, 8281-8289.

32. Weimer, S.; Drewnick, F.; Högrefe, O.; Schwab, J. J.; Rhoads, K.; Orsini, D.; Canagaratna, M. R.; Worsnop, D. R.; Demerjian, K. L.; Size-selective nonrefractory ambient aerosol measurements during the Particulate Matter Technology Assessment and Characterization Study–New York 2004 Winter Intensive in New York City. *J. Geophys. Res.* **2006**, 111, D18305, doi:10.1029/2006JD007215.

33. Cubison, M. J.; Alfarra, M. R.; Allan, J.; Bower, K. N.; Coe, H.; McFiggans, G. B.; Whitehead, J. D.; Williams, P. I.; Zhang, Q.; Jimenez, J. L.; Hopkins, J.; Lee, J., The characterisation of pollution aerosol in a changing photochemical environment. *Atmos. Chem. Phys.* **2008**, 6, 5573-5588.

34. Docherty, K. S.; Stone, E. A.; Ulbrich, I. M.; DeCarlo, P. F.; Snyder, D. C.; Schauer, J. J.; Peltier, R. E.; Weber, R. J.; Murphy, S. M.; Seinfeld, J. H.; Grover, B. D.; Eatough, D. J.; Jimenez, J. L., Apportionment of primary and secondary aerosols in Southern California during the 2005 Study of Organic Aerosols in Riverside (SOAR-1). *Environ. Sci. Technol.* **2008**, 42, 7655-7662.
35. Jimenez, J. L.; Canagaratna, M. R.; Donahue, N. M.; Prevot, A. S. H.; Zhang, Q.; Kroll, J. H.; DeCarlo, P. F.; Allan, J. D.; Coe, H.; Ng, N. L.; Aiken, A. C.; Docherty, K. S.; Ulbrich, I. M.; Grieshop, A. P.; Robinson, A. L.; Duplissy, J.; Smith, J. D.; Wilson, K. R.; Lanz, V. A.; Hueglin, C.; Sun, Y. L.; Tian, J.; Laaksonen, A.; Raatikainen, T.; Rautiainen, J.; Vaattovaara, P.; Ehn, M.; Kulmala, M.; Tomlinson, J. M.; Collins, D. R.; Cubison, M. J.; Dunlea, E. J.; Huffman, J. A.; Onasch, T. B.; Alfarra, M. R.; Williams, P. I.; Bower, K.; Kondo, Y.; Schneider, J.; Drewnick, F.; Borrmann, S.; Weimer, S.; Demerjian, K.; Salcedo, D.; Cottrell, L.; Griffin, R.; Takami, A.; Miyoshi, T.; Hatakeyama, S.; Shimojo, A.; Sun, J. Y.; Zhang, Y. M.; Dzepina, K.; Kimmel, J. R.; Sueper, D.; Jayne, J. T.; Herndon, S. C.; Trimborn, A. M.; Williams, L. R.; Wood, E. C.; Middlebrook, A. M.; Kolb, C. E.; Baltensperger, U.; Worsnop, D. R., Evolution of organic aerosol in the atmosphere. *Science* **2009**, 326, 1525 - 1529.
36. Meng, Q. Y.; Spector, D.; Colome, S.; Turpin, B. J., Determinants of indoor exposure to PM_{2.5} of indoor and outdoor origin during the RIOPA study. *Atmos. Environ.* **2009**, 43, 5750 - 5758.
37. Offenberg, J. H.; Kleindienst, T. E.; Jaoui, M.; Lewandowski, M.; Edney, E. O., Thermal properties of secondary organic aerosols. *Geophys. Res. Lett.* 2006, 33, L03816, doi:10.1029/2005GL024623.
38. Stanier, C. O.; Pathak, R. K.; Pandis, S. N., Measurements of the Volatility of Aerosols from α -Pinene Ozonolysis. *Environ. Sci. Technol.* **2007**, 41, 2756–2763.
39. Weschler, C.J, and Nazaroff, W.W., Semivolatile organic compounds in indoor environments. *Atmos. Environ.*, **2008**, 42, 9018-9040.
40. Liu, D. L. and Nazaroff, W. W., Modeling particle penetration across building envelopes. *Atmos. Environ.*, **2001**, 35, 4451-4462.
41. Liu, C.; Shiu, S.; Weschler, C.; Zhao, B.; Zhang, Y., Analysis of the dynamic interaction between SVOCs and airborne particles. *Aerosol Sci. Technol.*, **2013**, 47, 125-136.
42. Weschler, C. J; Nazaroff, W. W.; Cleaning products and air fresheners: exposure to primary and secondary air pollutants. *Atmos. Environ.* **2004**, 38, 2841-2865.

43. Waring, M. S.; Wells, J. R.; Siegel, J. A., Secondary organic aerosol formation from ozone reactions with single terpenoids and terpenoid mixtures. *Atmos. Environ.*, **2011**, 45, 4235-4242.
44. Weschler, C. J., Chemistry in indoor environments: 20 years of research. *Indoor Air*, **2011**, 21, 205-218.
45. Ervens, B.; Turpin, B. J.; Weber, R. J., Secondary organic aerosol formation in cloud droplets and aqueous particles (aqSOA): A review of laboratory, field and model studies, *Atmos. Chem. Phys.* **2011**, 11, 11069-11102.

	5 th	25 th	Median	75 th	95 th
	Percentile	Percentile		Percentile	Percentile
Outdoor OA ^a ($\mu\text{g}/\text{m}^3$)	1.07	2.66	4.25	6.47	10.90
Indoor OA ^a ($\mu\text{g}/\text{m}^3$)	2.45	4.95	7.42	11.04	27.43
Outdoor Temperature ($^{\circ}\text{C}$)	3.8	13.6	19.5	24.1	28.6
Indoor Temperature ($^{\circ}\text{C}$)	19.5	22.2	23.9	25.6	27.8

^aOA concentrations ($\mu\text{g}/\text{m}^3$) were estimated from measurements of organic carbon concentrations ($\mu\text{gC}/\text{m}^3$) assuming an OM:OC ratio of 1.4.

Table 5-1. Summary statistics of measured indoor and outdoor temperatures and organic aerosol (OA) concentrations for the RIOPA study homes included in this paper. Only homes for which all measurements shown above were available were included in this analysis (n = 167).

	Selection Step	Variable	Partial R ²	Model R ²	Coefficient Estimate	P
$\Delta H_{vap} = 100 \text{ kJ/mol}$						
All Homes (n = 167)	1	Temperature	0.45	0.45	-0.07	<0.0001
	2	OA Loading	0.24	0.69	0.05	<0.0001
Los Angeles County, CA (n = 44)	1	OA Loading	0.58	0.58	0.12	<0.0001
	2	Temperature	0.20	0.78	-0.10	<0.0001
Elizabeth, NJ (n = 54)	1	Temperature	0.54	0.54	-0.08	<0.0001
	2	OA Loading	0.24	0.78	0.03	<0.0001
Houston, TX (n = 69)	1	Temperature	0.30	0.30	-0.07	<0.0001
	2	OA Loading	0.32	0.62	0.05	<0.0001
$\Delta H_{vap} = 50 \text{ kJ/mol}$						
All Homes (n = 167)	1	OA Loading	0.51	0.51	0.11	<0.0001
	2	Temperature	0.23	0.74	-0.08	<0.0001
Los Angeles County, CA (n = 44)	1	OA Loading	0.78	0.78	0.25	<0.0001
	2	Temperature	0.10	0.88	-0.10	<0.0001
Elizabeth, NJ (n = 54)	1	OA Loading	0.56	0.56	0.08	<0.0001
	2	Temperature	0.28	0.84	-0.09	<0.0001
Houston, TX (n = 69)	1	OA Loading	0.58	0.58	0.11	<0.0001
	2	Temperature	0.12	0.70	-0.06	<0.0001

Table 5-2. Multiple linear regression (MLR) analyses investigating the contribution of indoor-outdoor differences in temperature and organic aerosol (OA) loading to changes in ambient OA concentrations (ΔC_{OA}) due to shifts in gas-particle partitioning with outdoor-to-indoor transport. Partial R² describes the variance in ΔC_{OA} explained by each variable individually. Model R² describes the total variance in ΔC_{OA} described by the full MLR model at each selection step. Coefficient estimates describe the change in partitioning (i.e., ΔC_{OA}) per unit difference between indoor and outdoor temperature and OA loading

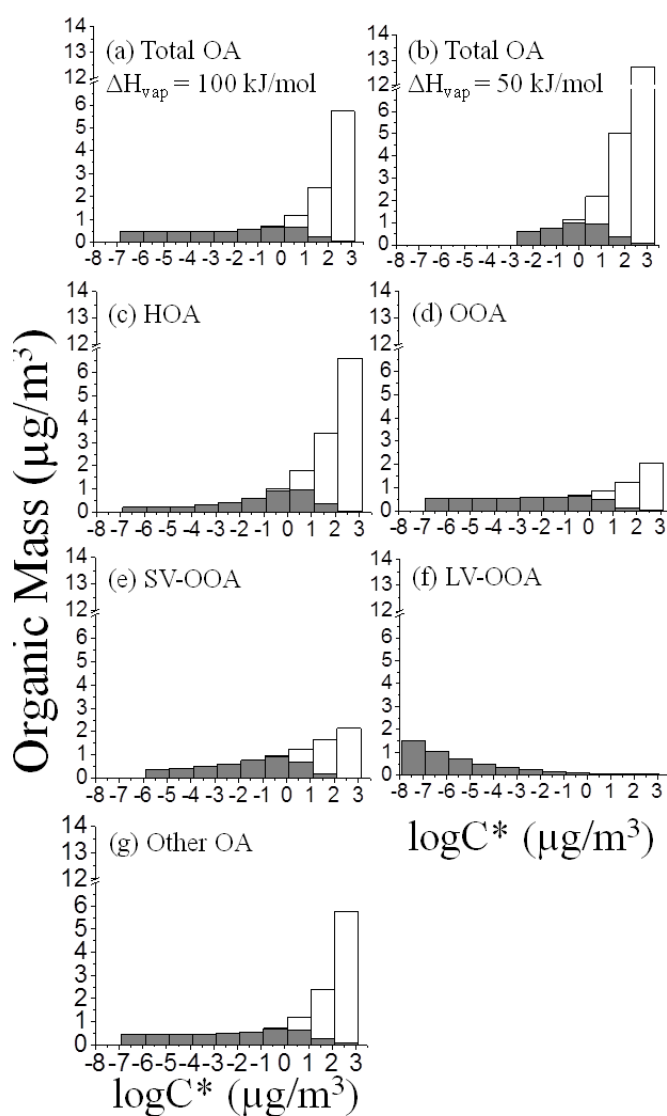


Figure 5-1. Volatility distributions for the average conditions measured outside of RIOPA homes (temperature = 18.2°C and OA loading = $4.87 \mu\text{g}/\text{m}^3$). (a) total OA assuming an enthalpy of vaporization of 100 kJ/mol ; (b) total OA assuming an enthalpy of vaporization of 50 kJ/mol ; (c) - (g) OA components derived from factor analysis assuming an enthalpy of vaporization of 100 kJ/mol for all components: hydrocarbon like OA (HOA), oxygenated OA (OOA), semi-volatile oxygenated OA (SV-OOA), low volatility oxygenated OA (LV-OOA), and "other" OA. The distribution for "other" OA is based on parameters for total OA (Appendix D1). The full bar for each saturation vapor pressure bin (C^*) indicates the total OM (gas + particle phase) in that volatility (saturation vapor pressure) bin. The shaded region indicates the fraction of that OM that is in the particle phase assuming absorptive partitioning into a single, well-mixed condensed phase. Adapted with permission from Cappa and Jimenez (2010).

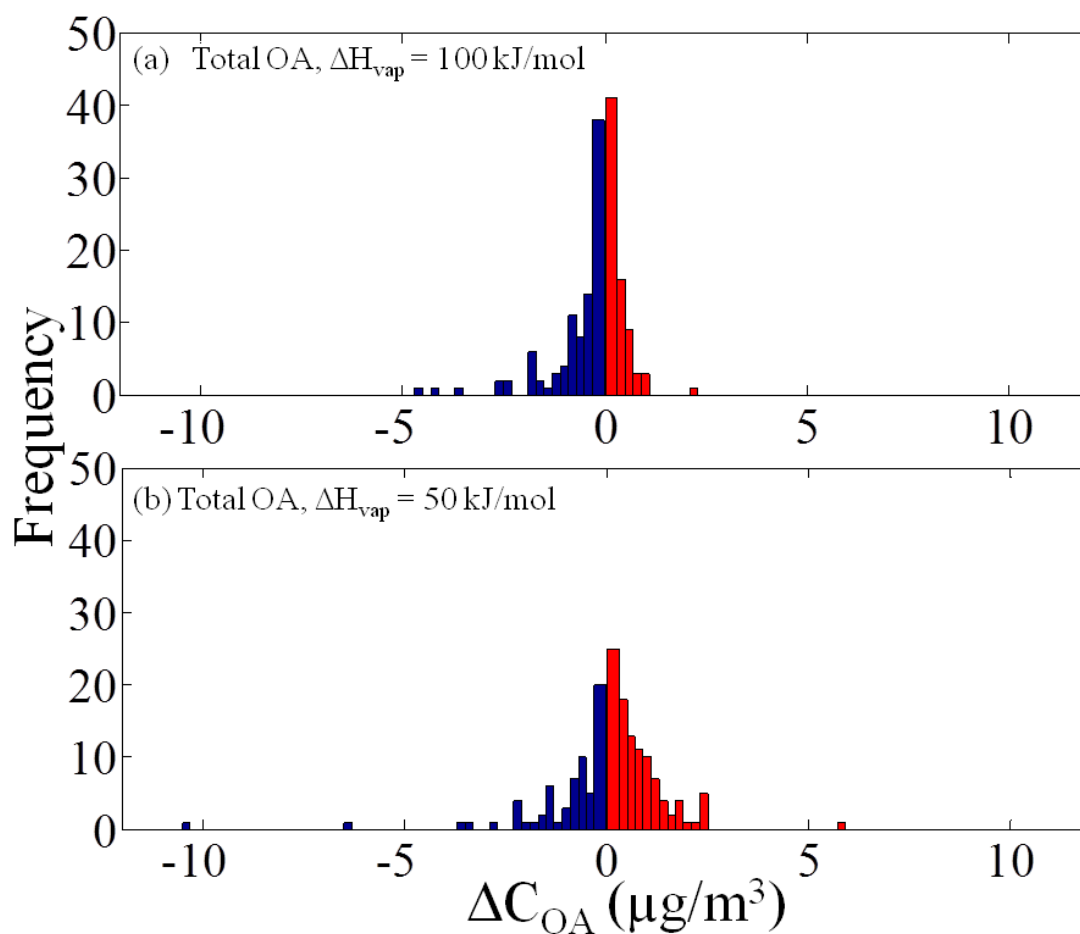


Figure 5-2. Frequency distributions of the change in organic aerosol mass concentrations due to changes in gas-particle partitioning with outdoor-to-indoor transport (ΔC_{OA}) assuming an enthalpy of vaporization (ΔH_{vap}) of (a) 100 kJ/mol and (b) 50 kJ/mol. Negative values indicate net volatilization, whereas positive values indicate net absorption.

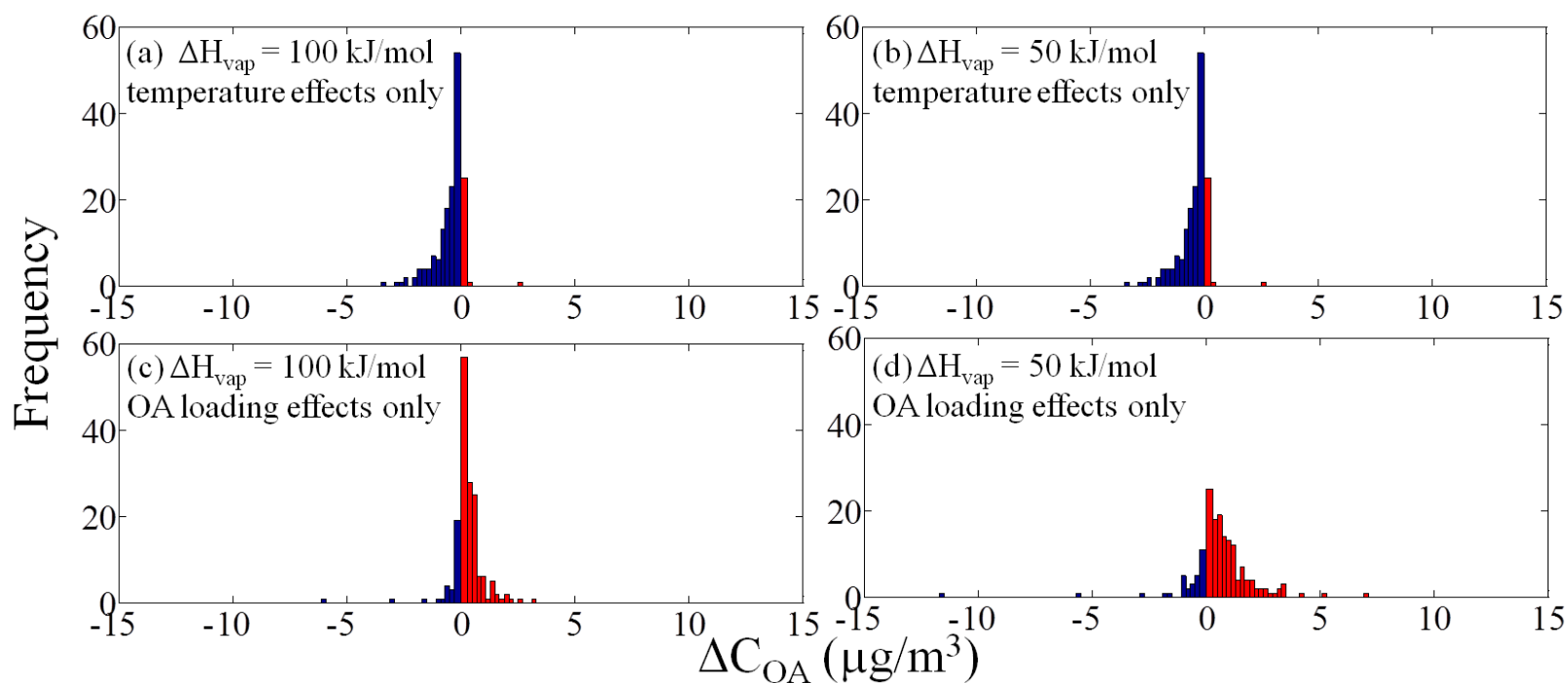


Figure 5-3. Frequency distributions of the change in organic aerosol mass concentrations due to changes in gas-particle partitioning with outdoor-to-indoor transport (ΔC_{OA}) accounting only for (a) - (b) indoor-outdoor temperature differences and (c) - (d) indoor-outdoor differences in OA loading.

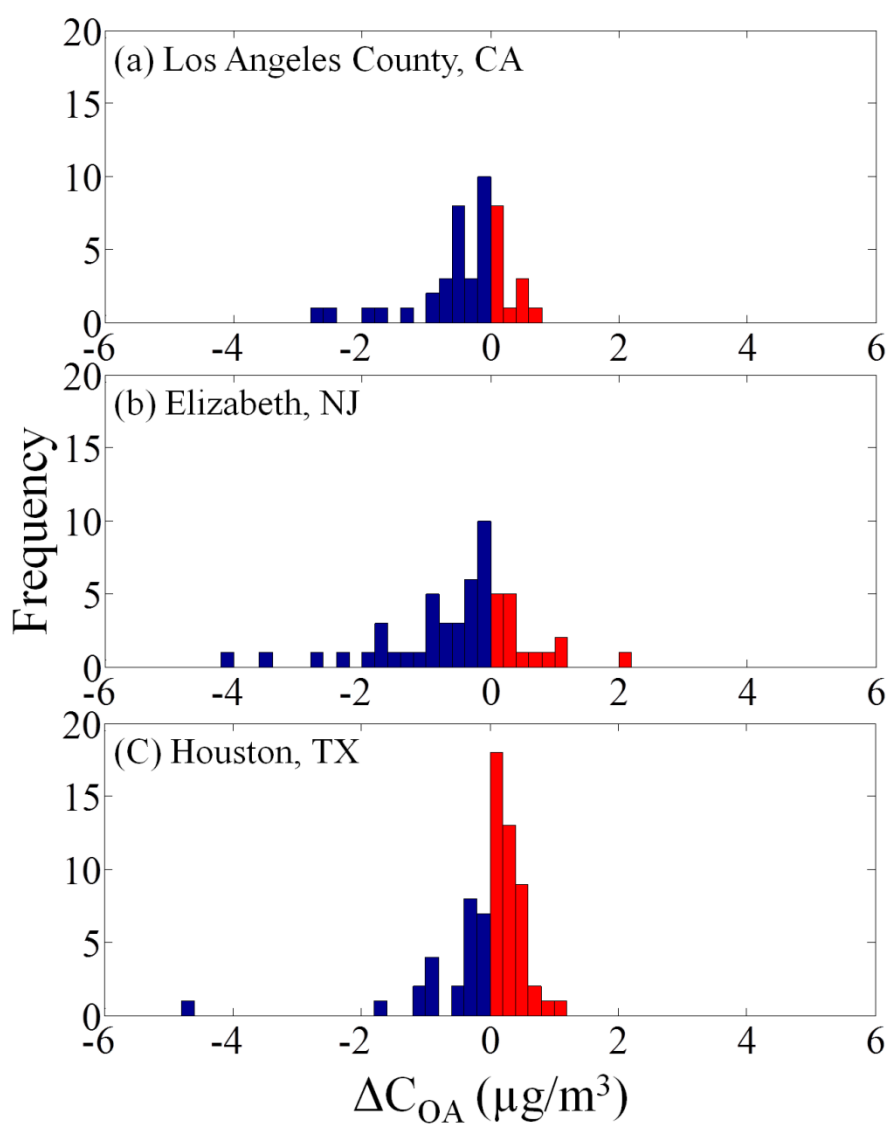


Figure 5-4. Frequency distributions of the change in organic aerosol mass concentrations due to changes in gas-particle partitioning with outdoor-to-indoor transport (ΔC_{OA}) for the three geographically and climatically diverse urban regions studied here: (a) Los Angeles County, CA, (b) Elizabeth, NJ, and (c) Houston, TX.

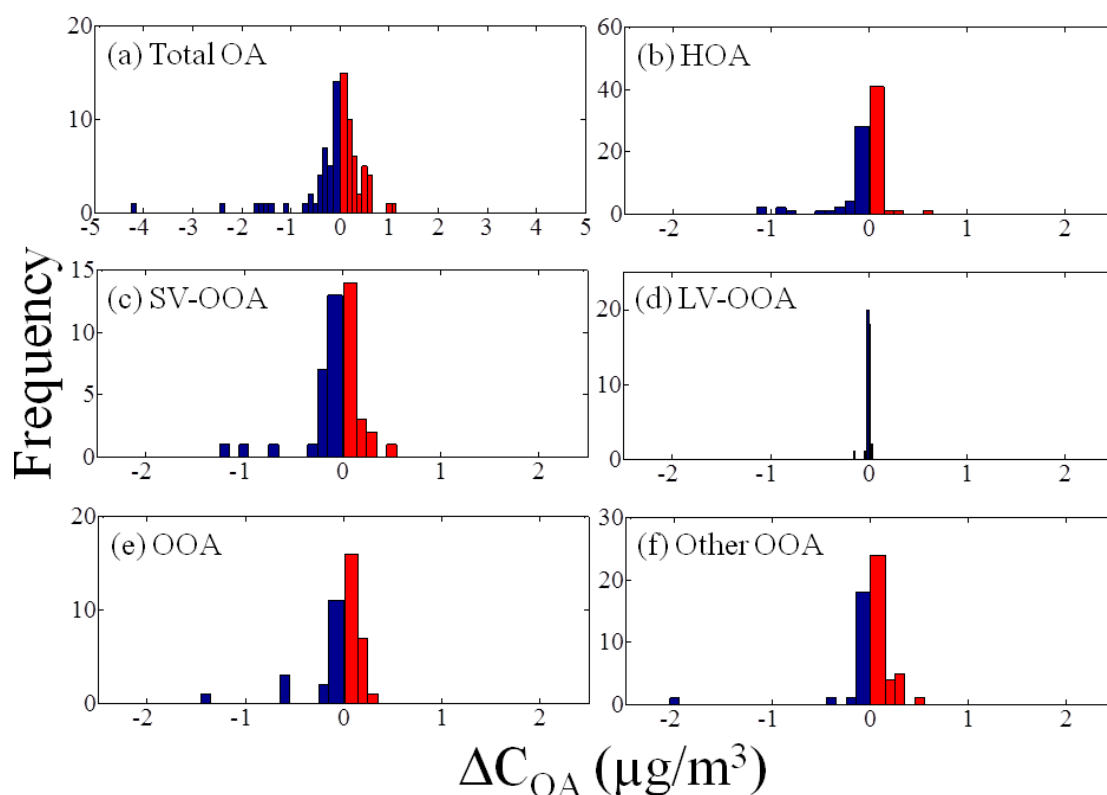


Figure 5-5. Frequency distributions of the change in organic aerosol mass concentrations due to changes in gas-particle partitioning with outdoor-to-indoor transport (ΔC_{OA}) for (a) total OA assuming that OA can be represented as a mixture of factor-analysis components with mass fractions given in Appendix D2 and (b) - (f) for each OA factor analysis component: (b) hydrocarbon-like OA (HOA), (c) semi-volatile oxygenated OA (SV-OOA), (d) low volatility oxygenated OA (LV-OOA), (e) oxygenated OA (OOA), and (f) other OA. The bin widths for each distribution are on the same order of magnitude as the standard deviations of ΔC_{OA} , illustrating differences across OA factor-analysis components. An enthalpy of vaporization of 100 kJ/mol was assumed for all components.

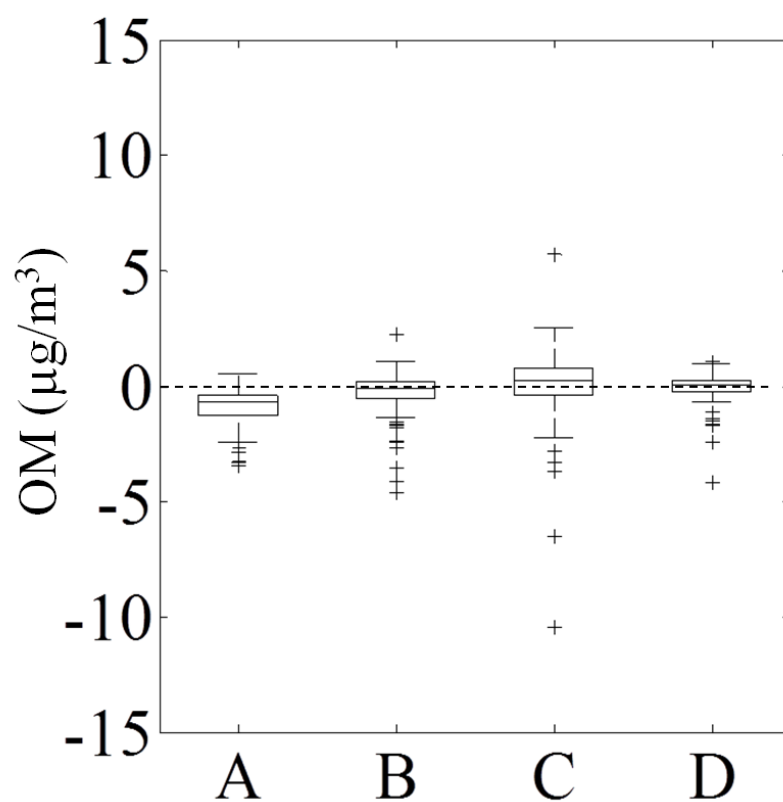


Figure 5-6. Distributions of (A) the difference between a statistical estimate of measured indoor OA of outdoor origin and indoor concentrations of ambient OA modeled accounting for only home ventilation and physical losses, (B) ΔC_{OA} calculated assuming an ΔH_{vap} of 100 kJ/mol, (C) ΔC_{OA} calculated assuming an ΔH_{vap} of 50 kJ/mol, and (D) ΔC_{OA} calculated assuming that OA can be represented as a mixture of factor-analysis components with mass fractions given in Appendix D2.

Chapter 6. Summary, Future Directions, and Implications

6.1 Summary and Concurrent Studies

The research presented in this dissertation develops and implements simple, practical methods and modeling tools that can be applied in large epidemiologic studies to predict residential exposures to ambient $\text{PM}_{2.5}$. Such modeling tools address the exposure error associated with variability in the fraction of ambient $\text{PM}_{2.5}$ that penetrates and persists indoors (F), error that this work demonstrates can bias health effect estimates in air pollution epidemiologic studies.

This research identifies the major drivers of variability in F , as well as circumstances, atmospheric conditions, and epidemiologic study designs for which variability in F is (and is not) a substantial contributor to error in air pollution epidemiology. When F is smaller and more variable, ambient $\text{PM}_{2.5}$ exposures are more different from and less correlated with central-site $\text{PM}_{2.5}$ concentrations. Under these circumstances, the non-Berksonian error and associated bias towards the null (i.e. underestimation of health effect) is greater. Thus, as is demonstrated in Chapter 2, this error is greatest when $\text{PM}_{2.5}$ composition is highly variable and/or enriched in semi-volatile species, when there is temporal and/or spatial variability in home ventilation conditions and the human activities that influence those conditions (e.g. opening windows, using central air conditioning), and when there is variability in residential proximity to roadways and primary $\text{PM}_{2.5}$ sources and in socioeconomic status among epidemiologic study subjects.

Chapter 2 is in agreement with concurrently conducted studies by Chen et al.^{1,2} Those studies used mass balance methods to estimate indoor concentrations of ambient

ozone and PM_{10} and explored the extent to which seasonal and geographic variability in mortality coefficients could be attributed to variability in the efficiency with which those pollutants were transported indoors. More specifically, city- and season-specific values of F were calculated and correlations between these F values and city- and season-specific PM_{10} and ozone mortality coefficients were examined. The authors found that heterogeneity in mortality risk could be attributed, at least in part, to variability in F associated with variability in AER, AC prevalence, and particle-size-dependent penetration and deposition losses. This is in agreement with the results of Chapter 2, which demonstrated that those factors are likely contributors to exposure error and to biases that influence effect estimates derived from epidemiologic studies.

I hypothesized based on the results of Chapter 2 that the case-crossover epidemiologic study design would be largely robust to much of the error associated with variability in F because subjects serve as their own controls, thus avoiding error associated with location of residence, proximity to sources, socioeconomic status, and residential construction. This was tested in Chapter 3. As hypothesized, use of refined exposure surrogates that account for human activity patterns and/or the indoor transport of ambient $PM_{2.5}$ did not result in larger health effect estimates, narrower confidence intervals, nor better model fits compared to the analyses that used central-site $PM_{2.5}$ concentrations to estimate $PM_{2.5}$ exposure for a case-crossover study. These results add support to the use of central-site $PM_{2.5}$ concentrations as ambient $PM_{2.5}$ exposure surrogates for epidemiological analyses that employ methods similar to the case-crossover design (i.e., studies for which a within-subject temporal comparison of exposure effects is conducted). In epidemiologic analyses for which a spatio-temporal

comparison of exposure effects across subjects is conducted, on the other hand, variability in factors that influence F may contribute to bias. This was demonstrated by the heterogeneity in risk of myocardial infarction with residential air exchange rate observed in Chapter 3. For such epidemiologic analyses, the effects of outdoor-to-indoor transport should be accounted for in estimates of ambient $\text{PM}_{2.5}$ exposure.

Contrary to the findings presented in Chapter 3, larger health effect estimates and better model fits were obtained in a concurrently conducted *time-series* study³ when exposure surrogates that accounted for outdoor-to-indoor transport were used compared to analyses that used central-site ambient PM_{10} concentrations. In fact, health effect estimates calculated in this time-series study using mass balance methods to predict indoor ambient PM_{10} concentrations were nearly double those calculated using central-site concentrations. Interestingly, the authors also found that between-city heterogeneity in health effect estimates was reduced when the refined exposure surrogates were used. I speculated (Chapter 3) that time-series analyses, which are temporal contrasts of daily pollutant concentrations and daily counts of health outcomes, may be only minimally impacted by spatially-varying factors that impact F (e.g., proximity to local $\text{PM}_{2.5}$ sources and variability in AERs or particle losses that stem from differences in housing stock or air conditioning prevalence). However, the results of Chen et al.³ suggest that variability in F can contribute to exposure error and bias in such study designs. This is likely because seasonally-varying factors that influence F (e.g., seasonal variability in AER driven by indoor-outdoor temperature differences, natural ventilation, or air conditioning use) can contribute to exposure error and bias in the time-series design. The case-control period in the case-crossover study explored in Chapter 3 was confined to one calendar

month and, thus, factors that vary on time-scales longer than a month would be expected to have little or no effect on the relative odds estimates. Note that both Chen et al.³ and Chapter 3 both found evidence that health effect estimates are larger for subjects with larger residential AERs. The modification of MI risk by community-average AER observed in Chapter 3 is also consistent with the results of Chen et al.^{1,2} which found that mortality associated with given increases in outdoor ozone and PM₁₀ concentrations were larger for cities with higher annual average AERs compared to those with smaller AERs.

It is possible that limitations in the exposure metrics used in Chapter 3, in addition to the use of different epidemiologic study designs, could have contributed to differences between my results and those of Chen et al.³ As is discussed in Chapter 3, uncertainty in the model inputs used to calculate the refined exposure surrogates could have contributed to exposure error and, thus, the potential benefits of the refined exposure surrogates may not have been fully realized. Notably, the refined exposure surrogates used in Chen et al.³ accounted for variability in human activities that can influence F . While the effect of opening windows on AER was accounted for in the exposure surrogates explored in Chapter 3 (using census-tract level AC prevalence data and assumptions regarding the outdoor temperature at which windows are opened), increases in P with window opening and increased indoor losses in the filters of HVAC systems were not considered in the calculations of indoor ambient PM_{2.5}. These factors include air conditioning use and opening windows. Chapter 4 contributes further evidence that it is important to account for these human activities in estimates of ambient PM_{2.5} exposure. That chapter demonstrates that not accounting for certain human activities (i.e., air conditioning and heating use, opening windows) leads to bias in predicted residential PM_{2.5} exposures

at the individual-subject level. Model parameters to account for the influence of human activities such as opening windows and using central air conditioning and heating on F were evaluated, in part, in Chapter 4. Importantly, a comprehensive review of $\text{PM}_{2.5}$ species size distribution measurements was conducted to improve estimates of depositional and filtration loss rates for individual ambient $\text{PM}_{2.5}$ species. To my knowledge, this was the first study to conduct such a comprehensive assessment.

This (Chapter 4) was also the first study to provide quantitative evidence that shifts in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport contribute significantly to variability in indoor ambient organic carbon concentrations. The sum of 5 - 7 ring polycyclic aromatic hydrocarbons (PAHs), which were measured in the gas and particle phases inside and outside of 76 RIOPA homes,⁴ was used as a surrogate for the partitioning behavior of total ambient organic aerosol. It was demonstrated that the shift in partitioning of those PAHs was a significant predictor of variability in the indoor concentration of organic carbon. This is in agreement with two concurrently conducted studies of PAH partitioning,^{5,6} which demonstrated that shifts in the gas-particle partitioning of outdoor-generated PAHs with outdoor-to-indoor transport contribute substantially to variability in residential PAH exposures and can influence health effect estimates derived from epidemiologic studies. It should be noted that those studies were examining shifts in partitioning of PAHs, whereas the work herein examines shifts in partitioning of total organic aerosol.

The findings of Chapter 4 and the concurrent studies discussed above^{5,6} led to the work of Chapter 5. Chapter 5 is the first study to predict shifts in the gas-particle partitioning of ambient organics with transport into the indoor environment (Chapter 5)

and provides valuable information regarding the magnitude of ambient organic aerosol exposure error likely to be associated with such shifts in partitioning.

The methods utilized in Chapter 5 assume that ambient organics achieve instantaneous equilibrium partitioning upon their transport into the indoor environment. The validity of this assumption was explored for PAHs by Shi and Zhao.⁵ Shifts in the gas-particle partitioning of ambient PAHs with transport indoors were modeled using a kinetic partitioning model and a simple, linear instantaneous partitioning model. These modeled values were then compared to partitioning coefficients calculated using measured gas- and particle-phase PAHs. The authors found that for PAHs, model-measurement agreement was improved for the kinetic partitioning model compared to the instantaneous partitioning model. It was noted in Chapter 5 that the timescale required for organic compounds to reach equilibrium partitioning between the gas and particle phases can vary over several orders of magnitude depending on vapor pressure. Research aimed at investigating the validity of the instantaneous partitioning assumption for ambient organics is an important area of ongoing work.⁷

6.2 Future Directions

6.2.1 Epidemiologic Analyses

In the case-crossover epidemiologic study investigating associations between ambient PM_{2.5} exposure and transmural MI (Chapter 3), health effect estimates derived from analyses that used refined exposure estimates (i.e., that accounted for human activity patterns and the indoor transport of ambient PM_{2.5}) did not differ from those that used central-site PM_{2.5} concentrations to estimate exposure. It was concluded that the case-crossover study design is minimally impacted by the error associated with

variability in F because subjects serve as their own controls. On the other hand, this error is likely to contribute to bias in studies for which there is a spatio-temporal comparison of exposure effects across subjects. A systematic evaluation of the extent to which specific epidemiologic study designs are (or are not) robust to the error and bias associated with variability in F is an important area of future work. This can be evaluated by applying the same methods to estimate exposure as presented in Chapter 3 in epidemiologic analyses with varying designs. A comparison of the health effect estimates calculated using the refined exposure estimates and those calculated using central-site $\text{PM}_{2.5}$ concentrations to estimate exposure will illustrate the influence of exposure error associated with variability in F on various study designs. I hypothesize that time series studies and other designs for which a temporal contrast within subjects is conducted will be minimally impacted by error associated with spatially varying factors that influence F (e.g. subject socioeconomic status, proximity to primary $\text{PM}_{2.5}$ sources, home construction), but that temporally varying factors (e.g. seasonal variability in AER, central AC and heating use, human activity patterns) will contribute to exposure error. Note that the time-stratified case-crossover study design used in Chapter 3 is uniquely suited to control for this error because case and control periods are confined to the same day of the week within a single calendar month.^{8,9} For study designs in which there is a spatio-temporal comparison across subjects (e.g. case-control studies, cohort/panel studies), I hypothesize that use of refined exposure surrogates will result in larger health effect estimates and smaller confidence intervals compared to analyses that utilize central-site $\text{PM}_{2.5}$ to estimate exposure.

It was also noted in Chapter 3 that the low spatial resolution of the ambient $PM_{2.5}$ concentrations used as model inputs to calculate indoor ambient $PM_{2.5}$ concentrations was a possible source of error that may have contributed to bias in calculated effect estimates for the refined exposure surrogates. Recall that all subjects living within 10 km of a given central-site monitor were assigned the same ambient $PM_{2.5}$ concentrations measured at that monitor. Many studies have presented methods to account for spatial variability in outdoor $PM_{2.5}$. For example, land use regression models, in which a statistical relationship between land use parameters (e.g. traffic/roadway size, commercial properties) and $PM_{2.5}$ concentrations is developed, interpolation between central-site monitors, and air quality models are commonly used to account for local-scale variability in $PM_{2.5}$ concentrations and composition.^{10,11} It is possible that a two-step approach involving a method to account for local-scale variability in outdoor $PM_{2.5}$ followed by the use of an outdoor-to-indoor transport model will offer the best results when predicting residential $PM_{2.5}$ exposure and will allow for the potential benefits of the refined exposure surrogates presented here to be fully realized. In other words, if both spatial variability in outdoor $PM_{2.5}$ concentrations *and* the effects of outdoor-to-indoor transport are contributing to exposure error, the use of an exposure surrogate that accounts only for one of those factors may not sufficiently reduce that exposure error so as to avoid bias in the health-effect estimates. In order to explore this, health-effect estimates can be recalculated for the case-crossover study discussed in Chapter 3 using a combination of methods to calculate refined estimates of ambient $PM_{2.5}$ exposure: one method to capture spatial variability in $PM_{2.5}$ concentrations and composition and one method to account for the effects of outdoor-to-indoor transport (i.e. the mass balance methods presented in this

dissertation). This will evaluate whether the results of Chapter 3, which found no differences between refined and central-site, can be attributed, at least in part, to the spatial resolution of the outdoor $PM_{2.5}$ concentrations used as outdoor-to-indoor transport model inputs (i.e. a lack of accounting for spatial variability in outdoor $PM_{2.5}$ concentrations and composition).

6.2.2 Data and Measurements

Chapters 4 and 5 make use of exposure measurements from the RIOPA study to evaluate and refine modeling tools and parameters for use in outdoor-to-indoor transport models (e.g. shifts in the gas-particle partitioning of ambient organics, human activities such as central heating and AC use and opening windows). These chapters also illustrate the data and measurements needed to maximize the performance of these methods when applied in future epidemiologic studies. Chapter 4 demonstrates that the collection of human activity data (i.e. opening windows, using central heating or AC) or a method to predict these human activity patterns can lead to substantial improvements in individual-subject level residential ambient $PM_{2.5}$ exposure estimates. The LBNL Infiltration model was recently refined to account for open windows in estimates of air exchange rate by assuming a simple relationship between outdoor temperature and human activities. More specifically, it is assumed that windows are open in homes without central AC when outdoor temperatures exceed 22.5°C (Appendix B). Data regarding central AC prevalence at the census tract level is available from the American Housing Survey. Similar methods could be developed to predict circumstances for which home occupants would be likely to open windows or use central AC or heating. The RIOPA Study provides a unique

opportunity to develop and evaluate such a method because this study includes data on indoor and outdoor temperatures, as well as human activities.

Chapter 5 demonstrated that shifts in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport are important determinants of ambient organic aerosol exposure. In order to account for these phase changes in predictive models, measurements that characterize indoor temperatures or distributions of indoor temperatures, as well as the geographic and seasonal variability in these values are needed. In addition, characterization of the gas- and particle-phase organics emitted indoors is an important area of future work. As noted in Chapter 5, outdoor-generated organic vapors can sorb to organic aerosol emitted indoors. However, indoor-generated organic vapors can also sorb to outdoor-generated organic aerosol, changing the mass, the composition, and, likely, the toxicity of this aerosol.¹² In addition, there are many water-soluble gases emitted indoors and, thus, relative humidity and the hygroscopicity of the aerosol and surfaces found indoors may affect partitioning. Knowledge regarding the compounds, phase, and thermodynamic properties of indoor-generated organics will improve our understanding of interactions between indoor and outdoor generated pollutants. This is an ongoing area of research.¹³

It is also important that the robustness of the volatility basis sets used to represent ambient organic aerosol volatility is evaluated. Similarities between the volatility basis sets for Mexico City and Riverside, CA¹⁴ suggest that these volatility distributions can be generalized to regions influenced by anthropogenic sources; however, this has not been explicitly determined. This can be accomplished by expanding the spatial coverage of thermodenuder-Aerosol Mass Spectrometer measurements and volatility basis set

calculations like those presented in Huffman et al.¹⁵ and Cappa and Jimenez¹⁴. These measurements and calculations are also needed in non-urban areas to characterize the volatility of the organic aerosol in such areas. Similarly, greater temporal and spatial coverage of Aerosol Mass Spectrometer factor-analysis organic aerosol component mass fraction measurements are needed before the methods presented in Chapter 5 can be incorporated into residential ambient PM_{2.5} exposure models.

6.3 Implications and Impacts

6.3.1 Implications for Future Studies

This work provides guidance to those planning future efforts to study the health effects of ambient PM_{2.5}. This work suggests that the case-crossover epidemiologic study design is robust to the exposure error associated with variability in F and supports the use of central-site PM_{2.5} concentrations to estimate exposure in future case-crossover studies. However, for study designs for which there is a comparison of exposure effects across subjects, this work demonstrates that accounting for the effects of outdoor-to-indoor transport in estimates of ambient PM_{2.5} exposure may reduce bias in calculated health-effect estimates. Furthermore, it highlights the measurements and data most critically needed to facilitate the prediction of residential ambient PM_{2.5} exposures in those studies. For studies in which exposures are estimated at the individual-subject level, the collection of human activity data (e.g., pertaining to windows, air conditioning, and heating) using questionnaires or activity diaries or a method to estimate human activity patterns over the length of the study will maximize the performance of the refined exposure estimates presented here. While human activities are not a major source of error for population-level studies, all epidemiologic studies that use these refined exposure surrogates will

benefit from the collection of ambient $\text{PM}_{2.5}$ composition data, as this information is crucial to selecting model parameters that describe particle-size dependent penetration and depositional losses, as well as phase changes of semi-volatile species. Similarly, deploying Aerosol Mass Spectrometers (which can provide particle composition data and organic aerosol component mass fractions for selection) at central-site monitors will provide valuable data that can be used to predict shifts in the gas-particle partitioning of ambient organics with outdoor-to-indoor transport.

6.3.2 Implications for Exposure Mitigation Strategies and Policies

One of the main objectives of air pollution epidemiology is to inform strategies and policies aimed at mitigating exposure to toxic air pollutants. The National Ambient Air Quality Standards (NAAQS) for ambient $\text{PM}_{2.5}$ are determined based on a comprehensive assessment of epidemiological and exposure studies.¹⁶ For the majority of the epidemiologic studies included in this assessment, central-site $\text{PM}_{2.5}$ concentrations were used as ambient $\text{PM}_{2.5}$ exposure surrogates.¹⁶ The results of this dissertation suggest that for some of these epidemiologic studies, variability in F may have contributed to error, resulting in an underestimation of health effects. Thus, it is possible that the current standards are based to some extent on biased health effect estimates and they may not sufficiently protect public health. The use of refined exposure surrogates like those presented here will help to elucidate associations between negative health outcomes and $\text{PM}_{2.5}$ and, as a result, has the potential to inform revisions to these standards.

Similarly, the NAAQS consider only total $\text{PM}_{2.5}$ mass and do not regulate mass concentrations of specific ambient $\text{PM}_{2.5}$ species. Recent epidemiological evidence suggests that some types of $\text{PM}_{2.5}$ (i.e. from certain sources or with a certain

compositions) may be more toxic than others.¹⁷⁻¹⁹ However, there is a great deal of uncertainty regarding which PM_{2.5} species are most toxic. Future applications of the methods presented in this dissertation will help to elucidate relationships between exposure to specific PM_{2.5} species and negative health outcomes. This work demonstrates that exposure misclassification varies across ambient PM_{2.5} species. For example, *F* values are generally lower for ammonium nitrate due to high evaporative losses and they are higher for PM_{2.5} derived from primary sources. The methods presented in this dissertation treat the indoor transport of each species separately and thus constitute an advancement in the quantification of exposure to ambient PM_{2.5} species. Improved knowledge regarding the health effects of ambient PM_{2.5} species will help to determine whether standards for specific species are needed and will inform targeted policies and technological controls aimed at reducing this exposure.

6.3.4 Broader Impacts

In addition to informing exposure mitigation strategies, this work has major implications for environmental justice issues. Because the fraction of outdoor-generated PM_{2.5} that penetrates and persists indoors differs for homes above and below the poverty line, as well as with source proximity, this work is especially important for reducing exposure error for populations in urban areas and with low socioeconomic status. Note that the *F* value calculated for a low-income home in close proximity to a roadway (*F*=0.62) was nearly twice that for a "typical" home located in a region where the PM_{2.5} is enriched in ammonium nitrate (*F*=0.36; Chapter 2). This illustrates the magnitude of the disparity in ambient PM_{2.5} exposures that can exist between low income urban residents and higher income suburban residents in the southwestern U.S., for example. Diseases

with known links to environmental contaminants place huge burdens on healthcare programs, economic programs, and quality of life. Such issues are especially prevalent in inner-city and industrial regions where there are high numbers of PM_{2.5} sources, but economic resources for healthcare programs are limited. The methods presented in this dissertation will help to identify locations and populations for which the potential for ambient PM_{2.5} exposure and the negative health effects associated with this exposure are greatest and, as a result, will inform efforts aimed at environmental justice.

6.4 References

1. Chen, C.; Zhao, B.; Weschler, C. J., Assessing the influence of indoor exposure to “outdoor ozone” on the relationship between ozone and short-term mortality in U.S. communities. *Environmental Health Perspectives* **2012**, 120, 235 - 240.
2. Chen, C.; Zhao, B.; Weschler, C. J., Indoor exposure to "outdoor PM10": assessing its influence on the relationship between PM10 and short-term mortality in U.S. cities. *Epidemiology* **2012**, 23, 870 - 878.
3. Chen, R.; Zhou, B.; Kan, H.; Zhao, B., Associations of particulate air pollution and daily mortality in 16 Chinese cities: An improved effect estimate after accounting for the indoor exposure to particles of outdoor origin. *Environmental Pollution* **2013**, 182, 278 - 282.
4. Naumova, Y.Y., Eisenreich, S.J., Turpin, B.J., Weisel, C.P., Morandi, M.T., Colome S.D., Totten, L.A., Stock, T.H., Winer, A.M., Alimokhtari, S., Kwon, J., Shendell, D., Jones, J., Maberti, S., Wall, S.J., 2002. Polycyclic aromatic hydrocarbons in the indoor and outdoor air of three cities in the U.S. *Environmental Science & Technology* 36, 2552-2559.
5. Shi, S., Zhao, B., Comparison of the predicted concentration of outdoor originated indoor Polycyclic Aromatic Hydrocarbons between a kinetic partition model and a linear instantaneous model for gas-particle partition. *Atmospheric Environment* **2012**, 59, 93-101.
6. Zhou, B.; Zhao, B., Population inhalation exposure to Polycyclic Aromatic Hydrocarbons and associated lung cancer risk in Beijing region: contributions of indoor and outdoor sources and exposure. *Atmospheric Environment* **2012**, 62, 472-480.

7. Liu, C.; Shiu, S.; Weschler, C.; Zhao, B.; Zhang, Y., Analysis of the dynamic interaction between SVOCs and airborne particles. *Aerosol Sci. Technol.*, **2013**, 47, 125-136.
8. Maclure, M., The case-crossover design: a method for studying transient effects on the risk of acute events. *Am. J. Epidemiol.* **1991**, 133, 144 - 153.
9. Levy, D.; Lumley, T.; Sheppard, L.; Kaufman, J.; Checkoway, H., Referent selection in case-crossover analyses of acute health effects of air pollution. *Epidemiology* **2001**, 12, 186 - 192.
10. Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahuvaroglu, T.; Morrison, J.; Giovis, C., A review and evaluation of intraurban air pollution exposure models. *J. Exposure Anal. Environ. Epidemiol.* **2005**, 15, 185 - 204.
11. Ryan, P. H.; LeMasters, G. K., A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhal. Toxicol.* **2013**, 19, 127 - 133.
12. Weschler, C. J. Nazaroff, W. W., Semivolatile organic compounds in indoor environments. *Atmos. Environ.* **2008**, 42, 9018 - 9040.
13. Weschler, C. J., Chemistry in indoor environments: 20 years of research. *Indoor Air* **2011**, 21, 205 - 218.
14. Cappa C. D.; Jimenez, J., Quantitative estimates of the volatility of ambient organic aerosol. *Atmos. Chem. Phys.* **2010**, 10, 5409 - 5424.
15. Huffman J. A.; Docherty, K. S.; Aiken, A. C.; Cubison, M. J.; Ulbrich, I. M.; DeCarlo, P. F.; Sueper, D.; Jayne, J. T.; Worsnop, D. R.; Ziemann, P. J.; Jimenez, J. L., Chemically-resolved volatility measurements from two megacity field studies. *Atmos. Chem. Phys.* **2009**, 9, 7161 - 7182.
16. U.S. EPA. Integrated Science Assessment for Particulate Matter (Final Report); U.S. Environmental Protection Agency (U.S. EPA), Washington, DC, EPA/600/R-08/139F, **2009**, 1–2228.
17. Bell, M. L.; Ebisu, K.; Peng, R. D.; Samet, J. M.; Dominici, F., Hospital admissions and chemical composition of fine particle air pollution. *Am. J. Respir. Crit. Care Med.* **2009**, 179, 1115 - 1120.
18. Krall, J. R.; Anderson, G. B.; Dominici, F.; Bell, M. L.; Peng, R. D., Short-term exposure to particulate matter constituents and mortality in a national study of U.S. urban communities. *Environ. Health Perspect.* **2013**, <http://dx.doi.org/10.1289/ehp.1206185>.
19. Rich, D. Q.; Özkaynak, H.; Crooks, J.; Baxter, L.; Burke, J.; Ohman-Strickland, P.; Thevenet-Morrison, K.; Kipen, H. M.; Zhang, J.; Kostis, J. B.; Lunden, M.; Hodas,

N.; Turpin, B., The triggering of myocardial infarction by fine particles is enhanced when particles are enriched in secondary species. *Environmental Science and Technology* **2013**, 47, 9414 - 9423.

Appendix A: Supporting Information for Chapter 2

Particle Size Distribution Data Analysis

Sampling-campaign-averaged hourly species-resolved size distributions from Queens, NY¹ and Fresno, CA² were analyzed to assign deposition loss rates for each major PM_{2.5} species. It is hypothesized that size distributions for each species are less variable than size distributions for total PM_{2.5} mass and, thus, deposition rates will be more accurate if fit based on *species-resolved* size distributions. Appendices A1 and A2 show the hourly size distributions for each species at the Queens, NY and Fresno, CA sampling locations, respectively. Notably, there is variability across PM_{2.5} species size distributions, but the major characteristics of the size distributions remain relatively constant within PM_{2.5} species within the sampling locations. The effects of variability in species size distributions across sampling locations on F are discussed in Chapter 2.

Accounting for changes in the gas-particle partitioning of nitrate

Due to changes in ambient conditions (i.e. temperature, relative humidity, and surface area for sorption), outdoor-to-indoor transport drives a shift in the equilibrium partitioning of nitrate between the gas and particle phases.³ The driving force for nitrate loss is the equilibrium vapor concentration and subsequent loss of gaseous nitric acid by sorption to indoor surfaces.³ A mass balance model that accounts for physical deposition and the evaporation kinetics of particulate ammonium nitrate was developed and validated against outdoor and indoor particulate nitrate measurements in a closed unoccupied home.³ Hering et al.⁴ fit a simpler expression for k_{evap,NO_3^-} to the more detailed modeling presented by Lunden. We used this simpler expression, which involves

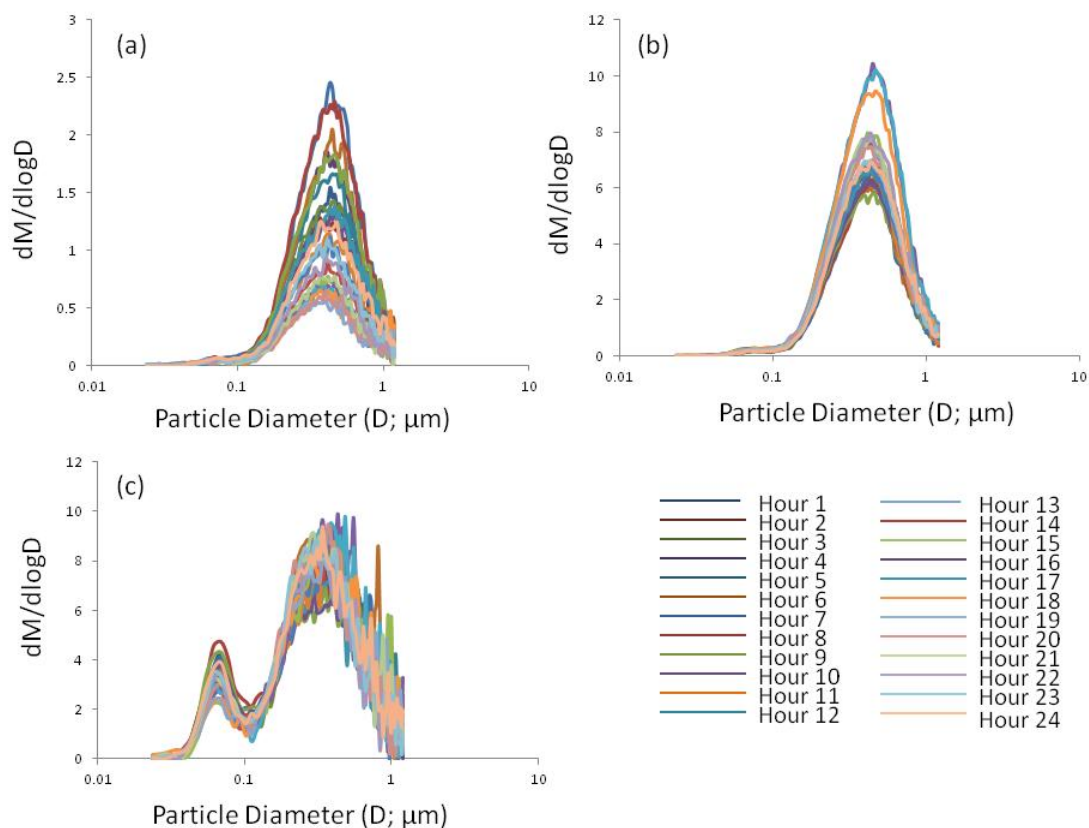
the temperature-dependent equilibrium constant for ammonium nitrate dissociation. We assumed an indoor temperature of 22°C for the purpose of calculating $k_{\text{evap},\text{NO}_3^-}$.

Appendix A References

1. Drewnick, F.; Jayne, J. T.; Canagaratna, M.; Worsnop, D. R.; Demerjian, K. L., Measurement of ambient aerosol composition during the PMTACS-NY 2001 using an aerosol mass spectrometer. Part II: chemically speciated mass distributions. *Aerosol Sci. Technol.* **2004**, 38, 104 - 117.
2. Ge, X. L.; Zhang, Q.; Sun, Y. L.; Ruchi, C. R.; Setyan, A., Impacts of aqueous-phase processing on aerosol chemistry and size distributions in Fresno, California, during wintertime. *Environ. Chem.* **2012**, 221 - 235.
3. Lunden, M. M.; Revzan, K. L.; Fischer, M. L.; Thatcher, T. L.; Littlejohn, D.; Hering, S. C.; Brown, N. J., The transformation of outdoor ammonium nitrate aerosols in the indoor environment. *Atmos. Environ.* **2003**, 37, 5633 - 5644.
4. Hering, S. V.; Lunden, M. M.; Thatcher, T. L.; Kirchstetter, T. W.; Brown, N. J., Using regional data and building leakage to assess indoor concentrations of particles of outdoor origin. *Aerosol Sci. Technol.* **2007**, 41, 639 - 654.

Appendix A1

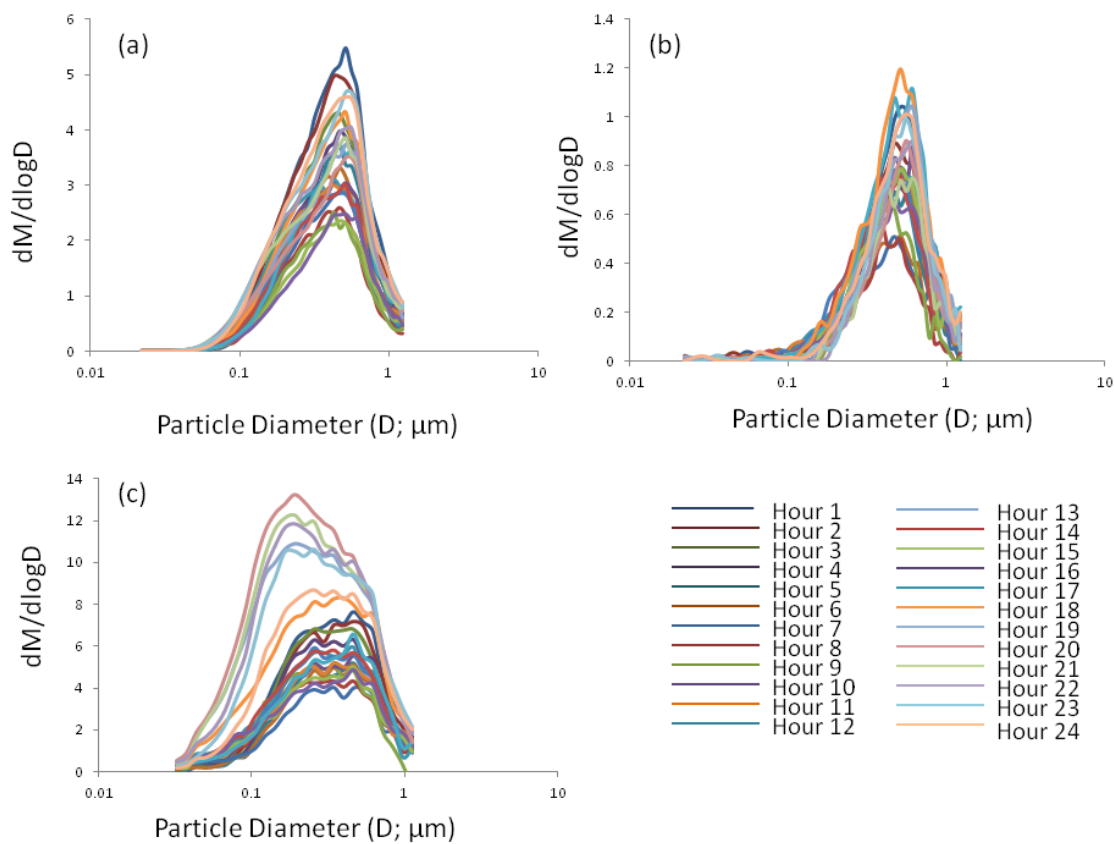
Queens, NY Size Distribution Measurements



Hourly campaign-average size distributions for (a) nitrate, (b) sulfate, (c) organics measured in Queens, NY in August 2001.¹

Appendix A2

Fresno, CA Size Distribution Measurements



Hourly campaign-average size distributions for (a) nitrate, (b) sulfate, (c) organics measured in Fresno, CA in January 2010.²

Appendix B: Supporting Information for Chapter 3

Air Exchange Rate

Calculation Method: Air exchange rates were calculated at the census tract level using the Lawrence Berkeley National Laboratory (LBNL) Infiltration Model¹ and an adjustment to account for open windows. The LBNL infiltration model predicts AER for single-family, closed homes (i.e. windows and doors closed) based on the leakage area, certain house characteristics, and meteorological conditions. Although the mechanisms driving airflow across a crack are well defined, the characteristics of cracks in buildings are not well known and are likely to be highly varied within and across buildings.² As a result, home leakiness is commonly quantified in terms of overall leakiness of the building shell, or effective leakage area (ELA).³ The distributions of normalized leakage (NL; ELA normalized by floor area) in each census tract was calculated using the regression analysis of Chan et al.³ Separate models were used depending on household poverty status because these factors affect home leakiness:³

$$NL_{low\ income} = \exp[11.1 - (5.37 \times 10^{-3} \times \text{year built}) - (4.18 \times 10^{-3} \times \text{floor area})] \quad (1)$$

$$NL_{conventional} = \exp[11.1 - (1.07 \times 10^{-2} \times \text{year built}) - (2.20 \times 10^{-3} \times \text{floor area})] \quad (2)$$

Resident poverty status, home-age and year-built distributions of the housing units in each census tract were retrieved from the Census 2000 Summary File 3 (SF3) available at the American Fact Finder website (<http://factfinder.census.gov>) and American Housing Survey (www.census.gov/hhes/www/ahs.html). In cases where needed parameters were not available separately for single-family and multi-unit residences, housing units that were

listed as “owner-occupied” were assumed to be single-family residences. Floor area, which is not directly available through the Census, was estimated from the distribution of number of rooms in each housing unit (available from SF3) and data relating number of rooms to floor area from the American Housing Survey. AER distributions were then calculated from the NL distributions:

$$AER = \frac{NL}{1000 \times h_f} \left(\frac{2.5}{H} \right)^{0.3} s \quad (3)$$

H is the building height and h_f is the height of the building’s ceiling. The specific infiltration rate (s) is a function of wind velocity (v), the stack parameter (f_s), the wind parameter (f_w), and the indoor-outdoor temperature difference (ΔT):

$$s = \sqrt{f_s^2 \Delta T + f_w^2 v^2} \quad (4)$$

The stack parameter is defined as

$$f_s = \left(\frac{1+R/2}{3} \right) \left(1 - \frac{X^2}{(2-R)^2} \right)^{3/2} \left(\frac{gH}{T_o} \right)^{1/2}. \quad (5)$$

R is the fraction of total leakage area contained within the floor and ceiling areas, X is the difference between floor leakage area and ceiling leakage area. H is the ceiling height, T_o is 298 K (77°F), and g is the acceleration due to gravity. The wind parameter is defined as

$$f_w = C(1 - R)^{1/3} A \left(\frac{H}{10} \right)^B. \quad (6)$$

C is a parameter that describes the magnitude of wind shielding resulting from obstructions surrounding the building. A and B are parameters that depend on the terrain and land usage surrounding the building. Parameter values and full derivation of the model are available from Sherman and Grimsrud.¹

It was assumed that H was constant at 5 m for all homes, the shielding parameter was 4 (obstructions around most of home perimeter), and that half of the total leakage

area in each home was contained within the walls ($R=0.5$). A and B were 0.67 and 0.25, respectively, and correspond to a terrain parameter of 4 (urban, industrial, or forested area). X was held constant at 0.25. It was also assumed that indoor temperature was constant at 20°C. Outdoor temperatures were retrieved from the same airports as were used to generate apparent temperature estimates as described above. Step-by-step examples of these calculations are available elsewhere.⁴

LBNL Infiltration Model Evaluation and Refinement: It was hypothesized that the LBNL infiltration model, which was designed to predict AER for closed, unoccupied housing stock, could be refined to predict AER distributions for occupied homes within a census tract. To focus these efforts, housing characteristics and human activities (which were available from questionnaires administered during the Relationships of Indoor, Outdoor, and Personal Air (RIOPA) Study to collect data regarding subject demographics, housing and neighborhood characteristics, and human activity patterns)⁴ were regressed on the differences between measured and modeled AERs for 50 Elizabeth, NJ RIOPA study homes using multiple linear regression (MLR) with stepwise selection ($\alpha = 0.15$ for variable entrance and removal threshold; SAS version 9.2, SAS Inc., Cary, NC).

For this purpose, AER was modeled with the LBNL infiltration model using information specific to these RIOPA study homes (e.g. house characteristics) and sampling dates (e.g. meteorological conditions). Income status of the occupants was provided on the RIOPA Baseline Questionnaire. Floor area, a factor required to calculate NL, was inferred from the number of rooms (also on the RIOPA Baseline Questionnaire) and data relating home floor area to total number of rooms (American Housing Survey).

Wind and outdoor-temperature data were gathered from the National Climate Data Center. Because NL is calculated with “year built” and the RIOPA Baseline Questionnaire recorded this variable in multiple-year intervals, AER was calculated twice for each home: once for the older year in the year-built interval and once for the newer year. All regression analyses discussed below were performed for both AER calculations.

Twelve questions relevant to home ventilation were identified in the Baseline and Activity Questionnaires (Appendix B3). These questions concern housing characteristics and human activities that took place during AER measurements. Collinearity between questionnaire variables was tested and outliers were removed. Possible collinearities among the variables were checked based on the variance inflation factor. For any highly correlated variables, only one was retained as the input for the final regression model. The final variance inflation factors for all variables were close to one, indicating no significant collinearities were observed among the variables. Outliers were detected based on the standard influence of the observation on the predicted value (DFFITS). A value was considered as an outlier if DFFITS was greater than $2\sqrt{k/N}$, where k is the number of regression coefficients to be determined and N is the number of observations.

A total of four questionnaire variables were selected as significant predictors ($\alpha < 0.15$) of the residual between measured and modeled AER. Presence of a fireplace, a dog in the home, amount of time windows were open, and amount of time a window fan was in use during sampling were selected. The model explains 29% of the variance in the residual (Appendix B2). Increased ventilation through open windows was identified as the most significant predictor, explaining 11% of the variance in model-measurement disagreement. This suggests that modification of the LBNL infiltration model to account

for natural ventilation is likely to improve this model's capability to predict AER distributions for occupied homes. Like opening windows, the presence of a fireplace in the home is likely a contributor to variability in AER because it is a source of leakage not captured by the LBNL infiltration model. Similarly, having a dog as a pet introduces a source of home leakage due to more frequent opening of the door to let the dog out. Increased leakage in homes with dogs has been previously observed (personal communication, H. Ozkaynak). Use of a window fan introduces mechanical ventilation that was not accounted for in this residential infiltration model. However, it should be noted that the number of homes with a window fan in use or with a fireplace was small (n=4).

The LBNL Infiltration model was refined to account for open windows, the most significant predictor of variance in differences between measured and modeled AERs. When outdoor temperatures exceeded 22.5°C, it was assumed that all homes without air conditioning had open windows. The percent of homes without air conditioning in each census tract was estimated from air conditioning prevalence data for sub regions of New Jersey available from the American Housing Survey. For homes with open windows, it was assumed that indoor temperature was 90% of the outdoor temperature and that having windows open increased leakage area by 0.5 m². Hourly air exchange rates were generated with the LBNL Infiltration model and this adjustment for open windows for each case and control period and were then averaged over the same 24 hours to provide community-average AER for each MI.

Model Validation: The distribution of AERs generated with the LBNL infiltration model and the adjustment for open windows was validated against AER measurements

performed in Elizabeth, NJ as part of the Relationships of Indoor, Outdoor, and, Personal Air (RIOPA) Study.⁵ Appendix B3 shows the summary statistics of the distribution of measured AERs for Elizabeth, NJ RIOPA homes and the distribution of AERs predicted using home-specific data gathered as part of the RIOPA study and zip-code level data from the U.S. Census and American Housing Survey. Distributions of measured AERs and AERs modeled with home-specific data are in good agreement (Appendix B3). There was lesser agreement between measured AERs and those calculated using zip-code level data; however, it should be noted that RIOPA participants were selected to over-represent homes in close proximity to sources and do not reflect a random sampling of the population.

Appendix B References

1. Sherman, M. H.; Grimsrud, D. T., Measurement of infiltration using fan pressurization and weather data. Lawrence Berkeley National Laboratory Report. LBNL-10852, Berkeley, CA, **1980**.
2. Liu, D. L.; Nazaroff, W. W., Modeling particle penetration across building envelopes. *Atmos. Environ.* **2001**, **35**, 4451 - 4462.
3. Chan, W. R.; Nazaroff, W. W.; Price, P. N.; Sohn, M. D.; Gadgil, A. J., Analyzing a database of residential air leakage in the United States. *Atmos. Environ.* **2005**, **39**, 3445 - 3455.
4. Turpin, B.; Rich. D. Q.; Lunden, M. M.; Hodas, N.; Thevenet-Morrison, K.; Ohman-Strickland, P., Refined Exposure Surrogates for Ambient PM in Epidemiologic Studies: Accounting for Temporal/Spatial Variations in Infiltration, **2012**, Final Report for EPA Cooperative Agreement CR-83407201-1.
5. Weisel, C. P.; Zhang, J.; Turpin, B. J.; Morandi, M. T.; Colome, S.; Stock, T. H.; Spektor, D. M.; Korn, L.; Winer, A.; Alimokhtari, S.; Kwon, J.; Mohan, K.; Harrington, R.; Giovanetti, R.; Cui, W.; Afshar, M., Maberti, S., Shendell, D., Relationship of Indoor, Outdoor and Personal Air (RIOPA) study: study design, methods and quality assurance/control results. *J. Exposure Sci. Environ. Epidemiol.* **2005**, **15**, 123-137.

Appendix B1
Input Seasonal Air Exchange Rate (h^{-1}) Distributions for the SHEDS Model Used to
Generate Tier 2A Exposure Estimates

Season	Distribution	GM (GSD)	Min, Max
Winter	lognormal	0.49 (2.06)	0.01, 4.8
Spring	lognormal	0.6 (2.03)	0.01, 6.6
Summer	lognormal	1.11 (2.29)	0.01, 11.8
Fall	lognormal	0.32 (3.54)	0.01, 6.4

Source: Murray, D. M. and D. E. Burmaster (1995). "Residential air exchange rates in the United States: empirical and estimated parametric distributions by season and climatic region." Risk Analysis 15(4): 459-465.

Appendix B2
Input Parameters for the SHEDS Model Used to Generate Tier 2A and Tier 3 Exposure Estimates

Macroenvironment	Microenvironment	Calculation Method	Parameter Values
Inside-Home ^a		Mass-Balance	Penetration Factor = $N^c[0.91, 0.1]$, limits [-, 1] ^d Decay/Deposition rate (h^{-1}) = $N[0.79, 0.3]$, limits [0.1, -]
Inside Other ^b	Inside Office (non smoking) Store School Restaurant	Linear regression	$3.6 + 0.18 * \text{Ambient}^e + N[0, 2.9]$, limits [-, -] $9 + 0.75 * \text{Ambient} + N[0, 2.1]$, limits [-, -] $6.8 * 0.6 * \text{Ambient} + N[0, 5.4]$, limits [-, -] $.8 + 1 * \text{Ambient} + N[0, 10]$, limits [-, -]
Outside ^b		none	Ambient
In-vehicle ^b		Linear regression	$0.7125 * \text{Ambient} + N[0, 6.64]$, limits [-12, 20]

^aSource: Meng, Q. Y., B. J. Turpin, et al. (2005). "Influence of ambient (outdoor) sources on residential indoor and personal PM_{2.5} concentrations: analysis of RIOPA data." Journal of Exposure Analysis and Environmental Epidemiology 15(17-28).

^bSource: Burke, J.M., and Vendantham (2005). "SHEDS-PM (Stochastic Human Exposure and Dose Simulation for Particulate Matter): User Guide

^c N indicates a normal distribution

^d - indicates no limit

^eunits are $\mu\text{g}/\text{m}^3$

Appendix B3
Mass Median Diameter and Associated Deposition (k_{dep}) Used for Particulate Sulfate, Nitrate, Elemental carbon (EC), and Organic carbon (OC) in the Tier 2B Exposure Estimates

Species	Sulfate	Nitrate	EC		OC	
			Mode 1	Mode 2	Mode 1	Mode 2
Mass Median Diameter (μm)	0.5	0.5	0.08	0.5	0.08	0.5
k_{dep} (h^{-1})	0.09	0.09	0.05	0.09	0.05	0.09

Appendix B4
Correlations of Measured PM_{2.5} Species Mass Fractions Across Central-Site Monitors

Sulfate				
	Elizabeth	New Brunswick	Camden	Chester
Elizabeth	1	0.6016	0.1239	0.4683
New Brunswick	0.6016	1	0.6687	0.4816
Camden	0.1239	0.6687	1	0.4396
Chester	0.4683	0.4816	0.4396	1

Nitrate				
	Elizabeth	New Brunswick	Camden	Chester
Elizabeth	1	0.8403	0.7338	0.7574
New Brunswick	0.8403	1	0.7799	0.7013
Camden	0.7338	0.7799	1	0.5989
Chester	0.7574	0.7013	0.5989	1

Elemental Carbon				
	Elizabeth	New Brunswick	Camden	Chester
Elizabeth	1	0.5218	0.3287	0.2771
New Brunswick	0.5218	1	0.5458	0.415
Camden	0.3287	0.5458	1	0.3838
Chester	0.2771	0.415	0.3838	1

Organic Carbon				
	Elizabeth	New Brunswick	Camden	Chester
Elizabeth	1	0.5025	0.5098	0.2733
New Brunswick	0.5025	1	0.4613	0.2905
Camden	0.5098	0.4613	1	0.3963
Chester	0.2733	0.2905	0.3963	1

Appendix B5

Regression results for variables identified on the RIOPA Baseline and Activity questionnaires as relevant to air exchange rate

(a)	Parameter	Estimate	Standard Error	χ^2	N	R ²
	Home Type - Attached	-4.15E-02	3.51E-01	9.06E-01	50	1.10E-02
	Home Type - Detached	-7.64E-02	2.56E-01	7.66E-01	50	1.10E-02
	Number of Rooms	1.69E-02	5.57E-02	7.62E-01	50	6.52E-04
	Number of Rooms - less than 3	1.88E-01	4.23E-01	6.57E-01	50	1.02E-01
	Number of Rooms - 4	4.77E-02	3.23E-01	8.83E-01	50	1.02E-01
	Number of Rooms - 5	-2.04E-01	3.06E-01	5.04E-01	50	1.02E-01
	Number of Rooms - 6	-9.34E-02	3.23E-01	7.72E-01	50	1.02E-01
	Fireplace	-4.31E-01	5.14E-01	4.02E-01	50	3.98E-02
	Pet - No Pet	-1.99E-01	2.11E-01	3.46E-01	50	1.53E-01
	Pet - Dog as Pet	2.90E-01	2.45E-01	2.35E-01	50	1.19E-01
	Window open	4.50E-01	1.80E-01	1.24E-02	90	7.61E-02
	Time with window open	2.33E-04	7.43E-05	1.72E-03	90	1.40E-01
	Central AC in use	-5.42E-02	3.75E-01	8.85E-01	90	1.06E-05
	Time with central AC in use	-7.02E-05	1.60E-04	6.61E-01	90	2.95E-03
	Attic fan in use	3.67E+00	8.13E-01	6.26E-06	90	6.73E-02
	Time with attic fan in use	2.55E-03	5.65E-04	6.26E-06	90	6.73E-02
	Evaporative cooler in use	6.71E-02	5.82E-01	9.08E-01	90	1.54E-03
	Time with evaporative cooler in use	2.33E-05	2.02E-04	9.08E-01	90	1.54E-03
	Ceiling fan in use	1.72E-01	2.29E-01	4.52E-01	90	1.28E-02
	Time with ceiling fan in use	2.51E-04	2.41E-04	2.98E-01	90	5.55E-03
	Window fan in use	2.91E-04	2.41E-04	2.26E-01	90	3.44E-02
	Time with window fan in use	7.35E-01	3.66E-01	4.48E-02	90	1.05E-01
	Exhaust fan in use	-3.16E-01	1.99E-01	1.12E-01	90	8.53E-02
	Time with exhaust fan in use	-1.67E-03	1.13E-03	1.38E-01	90	3.62E-03
	Central heat in use	1.05E-01	3.03E-01	7.29E-01	90	6.48E-04
	Time with central heat in use	5.48E-05	1.25E-04	6.62E-01	90	1.60E-03

(b)

Parameter	Estimate	Standard Error	χ^2	N	R ²
Home Type - Attached	-1.33E-01	3.29E-01	6.86E-01	50	5.13E-03
Home Type - Detached	-1.09E-01	2.40E-01	6.51E-01	50	5.13E-03
Number of Rooms	1.03E-02	5.24E-02	8.44E-01	50	2.04E-02
Number of Rooms - less than 3	1.44E-01	3.98E-01	7.17E-01	50	5.32E-02
Number of Rooms - 4	9.15E-02	3.04E-01	7.63E-01	50	5.32E-02
Number of Rooms - 5	-1.56E-01	2.88E-01	5.88E-01	50	5.32E-02
Number of Rooms - 6	-9.16E-02	3.04E-01	7.63E-01	50	5.32E-02
Fireplace	-5.33E-01	4.80E-01	2.67E-01	50	4.86E-02
No Pet	-2.76E-01	1.84E-01	1.35E-01	50	2.40E-01
Dog as Pet	3.48E-01	2.12E-01	9.99E-02	50	2.00E-01
Window open	3.67E-01	1.70E-01	3.11E-02	90	3.50E-02
Time with window open	1.94E-04	7.04E-05	5.90E-03	90	1.19E-01
Central AC in use	-1.17E-01	3.35E-01	7.27E-01	90	6.32E-06
Time with central AC in use	-1.08E-04	1.50E-04	4.70E-01	90	5.97E-03
Attic fan in use	3.78E+00	7.26E-01	1.96E-07	90	8.27E-02
Time with attic fan in use	2.62E-03	5.04E-04	1.96E-07	90	8.27E-02
Evaporative cooler in use	-4.19E-02	5.20E-01	9.36E-01	90	5.03E-04
Time with evaporative cooler in use	-1.46E-05	1.80E-04	9.36E-01	90	5.03E-04
Ceiling fan in use	5.34E-02	1.98E-01	7.88E-01	90	6.42E-03
Time with ceiling fan in use	7.50E-05	2.10E-04	7.21E-01	90	6.80E-03
Window fan in use	2.58E-04	2.15E-04	2.29E-01	90	3.76E-02
Time with window fan in use	7.16E-01	3.25E-01	2.76E-02	90	1.27E-01
Exhaust fan in use	-3.17E-01	1.69E-01	6.17E-02	90	6.08E-02
Time with exhaust fan in use	-1.81E-03	1.01E-03	7.15E-02	90	2.76E-04
Central heat in use	1.59E-01	2.70E-01	5.57E-01	90	2.07E-03
Time with central heat in use	7.38E-05	1.12E-04	5.09E-01	90	1.11E-02

Twelve questions (25 variables) related to home ventilation were selected from the Baseline and Activity Questionnaires from the RIOPA study to be included in the regression analyses. Regressions were performed on the residual between measured air exchange rates and air exchange rates calculated with the LBNL Infiltration Model for the older year (a) and the newer year (b) in the “year built” interval reported on the RIOPA Baseline Questionnaire.

Appendix B6
LBNL Infiltration Model Evaluation Multiple Linear Regression Results

(a)

Variable	Selection Step	Parameter Estimate	Partial R ²	Model R ²	N
Intercept	-	-0.08	-	-	-
Time with windows open	1	3.93×10^{-4}	0.11	0.11	42
Pet dog	2	0.62	0.07	0.18	14
Fireplace	3	1.01	0.06	0.24	4
Time using window fan	4	2.51×10^{-3}	0.05	0.29	4

(b)

Variable	Selection Step	Parameter Estimate	Partial R ²	Model R ²	N
Intercept	-	-0.25	-	-	-
Time with windows open	1	4.08×10^{-4}	0.12	0.12	42
Fireplace	2	1.06	0.07	0.19	4
Pet dog	3	0.60	0.06	0.25	14
Time using window fan	4	2.56×10^{-3}	0.05	0.30	4

Selection summary and parameter estimates for multiple linear regression (MLR) model relating housing characteristics and human activities to the difference between LBNL-infiltration-modeled and measured AER for (a) the newer year in the year-built interval and (b) the older year in the year-built interval. Partial R² described the variance in the difference between measured and modeled AERs described by each variability independently and model R² describes the variance in this differences explained by the MLR model at each selection step.

Appendix B7
Comparison of Measured Air Exchange Rates and LBNL-Infiltration-Modeled Air Exchange Rates

	N	Mean	Median	Standard Deviation
Measured AER	155	1.19	0.88	0.93
Modeled AER (home-specific level)	94	1.17	0.93	0.77
Modeled AER (zip-code-level)	160	0.93	0.61	0.82

Summary statistics for the distribution of measured air exchange rates (AERs) for Elizabeth, NJ RIOPA homes and for AERs modeled with the LBNL Infiltration Model using home-specific data from the RIOPA questionnaires and zip-code level data gathered from the Census and American Housing Survey for RIOPA home zip codes. The zip-code level calculations are more relevant to the calculated AERs used as inputs for the Tier 2b exposure estimates. However, it should be noted that RIOPA homes were not a representative sample of the housing stock in each Elizabeth zip code.

Appendix B8
Summary Statistics for Exposure Surrogate Tiers

(a) All Tiers by location ($\mu\text{g}/\text{m}^3$)

Monitoring Area	Tier 1		Tier 2A		Tier 2B		Tier 3	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Camden	12.9 (8.3)	0 – 67.0	6.5 (4.7)	0 – 30.6	6.8 (4.5)	0 – 35.3	6.0 (3.9)	0.16 – 25.9
Elizabeth	14.5 (9.6)	0 – 61.4	7.2 (5.35)	0 – 36.0	7.7 (5.4)	0.20 – 36.6	7.0 (5.0)	0 – 31.4
Flemington	9.8 (7.6)	0 – 47.1	5.4 (4.3)	0 – 27.6	4.8 (4.0)	0.34 – 25.7	4.4 (3.3)	0 – 21.1
Jersey City	13.2 (9.3)	1 – 62.1	7.0 (5.2)	0.39 – 36.8	7.2 (5.5)	0.84 – 33.8	7.0 (5.0)	0.78 – 34.7
Millville	11.7 (7.7)	0.56 – 51.3	6.2 (5.0)	0.24 – 41.8	6.2 (4.4)	0.38 – 29.3	5.5 (4.1)	0.22 – 34.4
New Brunswick	11.1 (7.4)	0.28 – 59.8	5.5 (4.2)	0.25 – 28.6	5.7 (4.2)	0.40 – 29.7	4.8 (3.4)	0.23 – 22.2
Rahway	13.7 (7.1)	4.5 – 48.8	7.1 (4.2)	1.9 – 28.7	7.1 (4.2)	2.5 – 29.0	6.8 (3.8)	2.1 – 26.8

(b) Tier 1 exposure surrogates ($\mu\text{g}/\text{m}^3$): Summary statistics by season and location

Monitoring Area	Winter		Spring		Summer		Fall	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Camden	12.0 (6.4)	0.86 – 37.1	11.2 (6.7)	1.3 – 41.1	17.3 (10.4)	1.9 – 67.0	10.4 (6.8)	0 – 35.5
Elizabeth	13.5 (8.8)	0 – 56.4	13.2 (8.2)	0.1 – 55.6	18.5 (10.8)	0 – 61.4	12.6 (9.3)	0.03 – 50.0
Flemington	8.4 (4.9)	1.5 – 22.8	7.7 (5.4)	0 – 28.6	13.7 (9.8)	1.1 – 47.1	8.3 (6.2)	0.20 – 27.4
Jersey City	11.2 (7.6)	1.0 – 39.0	11.3 (7.6)	1.2 – 35.9	18.1 (11.3)	2.2 – 62.1	12.0 (8.5)	2.2 – 40.5
Millville	9.6 (5.2)	0.83 – 23.7	9.8 (6.2)	0.56 – 30.0	16.7 (9.8)	1.3 – 51.3	10.4 (6.2)	0.63 – 28.2
New Brunswick	9.7 (5.2)	0.98 – 26.1	9.6 (5.6)	0.28 – 41.9	15.4 (9.9)	1.1 – 59.8	9.4 (6.1)	0.53 – 38.6
Rahway	11.2 (4.0)	5.9 – 21.8	11.6 (5.3)	4.5 – 31.3	17.8 (9.0)	6.3 – 48.8	13.5 (6.1)	5.8 – 33.2

(c) Tier 2a exposure surrogates ($\mu\text{g}/\text{m}^3$): Summary statistics by season and location

Monitoring Area	Winter		Spring		Summer		Fall	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Camden	5.7 (3.1)	0.38 – 17.1	5.6 (3.5)	0.60 – 20.1	10.1 (6.1)	1.1 – 30.6	4.6 (3.0)	0 – 16.6
Elizabeth	6.5 (4.4)	0 – 27.5	6.6 (4.4)	0.05 – 24.8	10.7 (6.6)	0 – 36.0	4.9 (3.8)	0.01 – 22.8
Flemington	4.1 (2.3)	0.15 – 11.3	4.0 (2.8)	0 – 15.1	8.5 (5.8)	0.64 – 27.6	4.2 (2.8)	0.28 – 12.7
Jersey City	5.5 (3.6)	0.39 – 19.6	5.8 (3.9)	1.2 – 18.6	10.9 (6.7)	1.2 – 36.8	5.6 (3.7)	0.79 – 18.2
Millville	4.7 (2.5)	0.33 – 11.4	5.0 (3.2)	0.25 – 15.8	10.7 (7.2)	0.76 – 41.8	4.4 (2.5)	0.24 – 13.2
New Brunswick	4.7 (2.6)	0.39 – 12.7	4.8 (3.0)	0.25 – 21.6	8.8 (5.7)	0.61 – 28.6	3.9 (2.6)	0.28 – 16.1
Rahway	5.3 (1.9)	2.0 – 10.7	6.0 (2.8)	2.3 – 16.1	10.8 (5.4)	3.6 – 28.7	5.8 (2.6)	1.9 – 14.7

(d) Tier 2b exposure surrogates ($\mu\text{g}/\text{m}^3$): Summary statistics by season and location

Monitoring Area	Winter		Spring		Summer		Fall	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Camden	5.7 (2.8)	0.91 – 15.7	5.6 (3.4)	0.81 – 21.7	10.0 (5.9)	1.0 – 35.3	5.9 (3.2)	0 – 17.3
Elizabeth	6.3 (4.0)	0.22 – 21.7	6.8 (4.3)	0.28 – 29.4	11.1 (6.5)	0.20 – 36.6	6.6 (4.7)	0.24 – 24.7
Flemington	3.6 (2.0)	0.71 – 10.1	3.8 (2.7)	0.34 – 15.6	7.4 (5.4)	0.71 – 25.7	3.9 (2.8)	0.42 – 14.5
Jersey City	5.3 (3.6)	0.84 – 17.0	6.1 (4.4)	0.95 – 21.4	11.0 (6.9)	1.8 – 33.8	6.4 (4.4)	1.4 – 19.6
Millville	4.5 (2.4)	0.97 – 11.7	5.0 (2.8)	0.45 – 16.9	9.9 (5.7)	0.73 – 29.3	5.4 (3.2)	0.38 – 16.3
New Brunswick	4.4 (2.3)	0.57 – 11.7	4.6 (2.8)	0.40 – 21.4	8.8 (5.8)	0.71 – 29.7	4.6 (2.9)	0.53 – 19.5
Rahway	5.0 (1.9)	2.5 – 11.0	5.9 (3.1)	2.5 – 18.2	9.8 (5.5)	3.6 – 29.0	7.0 (3.1)	2.4 – 19.2

(e) Tier 3 exposure surrogates ($\mu\text{g}/\text{m}^3$): Summary statistics by season and location

Monitoring Area	Winter		Spring		Summer		Fall	
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Camden	5.9 (3.2)	0.73– 18.1	5.0 (3.1)	0.55 – 17.4	8.1 (5.0)	0.79 – 25.9	4.9 (3.1)	0.16 – 16.6
Elizabeth	6.7 (4.4)	0.02 – 26.8	6.2 (4.1)	0.08 – 21.9	9.6 (6.1)	0 – 31.3	5.4 (4.0)	0.03 – 23.6
Flemington	3.8 (2.2)	0.15 – 10.0	3.3 (2.4)	0 – 13.2	6.2 (4.4)	0.40 – 21.1	4.0 (2.7)	0.26 – 12.2
Jersey City	5.8 (3.5)	0.78 – 19.0	5.7 (3.9)	1.2 – 21.6	10.1 (6.6)	1.1 – 34.7	6.5 (4.2)	1.0 – 19.3
Millville	4.6 (2.4)	0.61 – 11.2	4.5 (2.9)	0.25 – 15.0	8.8 (6.1)	0.53 – 34.4	4.6 (2.6)	0.22 – 14.5
New Brunswick	4.5 (2.6)	0.48 – 12.7	4.1 (2.7)	0.23 – 18.9	6.9 (4.6)	0.53 – 34.4	3.9 (2.6)	0.43 – 16.8
Rahway	5.4 (1.8)	2.2 – 10.2	5.6 (2.6)	2.2 – 15.3	9.5 (5.1)	2.4 – 26.8	6.3 (2.9)	2.1 – 17.2

Appendix B9
LBNL Infiltration Model Air Exchange Rate Summary Statistics by Tertile and Season

Air Exchange Rates (h ⁻¹)							
Tertile	Minimum	5 th	25 th	Median	75 th	95 th	Maximum
COOL SEASON							
Low	0.24	0.28	0.36	0.41	0.44	0.46	0.47
Middle	0.47	0.48	0.50	0.54	0.57	0.60	0.60
High	0.60	0.61	0.63	0.67	0.72	0.81	1.11
WARM SEASON							
Low	0.19	0.28	0.35	0.40	0.44	0.46	0.47
Middle	0.47	0.47	0.50	0.53	0.57	0.60	0.60
High	0.60	0.62	0.67	0.75	0.87	1.05	1.35

Air exchange rate distributions for the cool season (November – April) and the warm season (May – October). Similarities between AER distributions for the warm and cool seasons can be attributed to the fact that indoor-outdoor temperature differences and human activities such as opening windows are major drivers of AER. The warm and cold seasons explored here included the transitional seasons spring and fall, respectively, which tend to have similar distributions of indoor-outdoor temperature difference. At the colder extremes of outdoor temperature, homes with heating in use have large differences in indoor and outdoor temperatures, resulting in higher AERs. This is also true for homes with air conditioning in use during warm days. While air conditioning prevalence is relatively low in the region studied, higher AERs can also occur during the warmer months because open windows allow greater air exchange.

Appendix B10
PM_{2.5} Species Mass Fractions and Ambient PM_{2.5} Concentrations

PM_{2.5} Species	Mass Fraction	
	Cool Season	Warm Season
Sulfate	0.43 ± 0.13	0.50 ± 0.14
Nitrate	0.25 ± 0.11	0.10 ± 0.06
Elemental Carbon	0.07 ± 0.03	0.07 ± 0.04
Organic Carbon	0.25 ± 0.12	0.33 ± 0.14
Mass Concentration (µg/m³)		
Total PM_{2.5}	10.2 ± 6.1	11.4 ± 8.4

Study-period average species mass fractions of the major PM_{2.5} species and total PM_{2.5} concentrations ± standard deviations in the cool (November to April) and warm (May to October) seasons measured at the New Brunswick central-site monitor.

Appendix B11
Study Subjects by Monitoring-Site Community

Central Site PM_{2.5} Monitor	AER tertile	# of MI	Percent of tertile- specific MI assigned to monitor
Camden	Low	110	21.9%
Elizabeth	Low	78	15.5%
Flemington	Low	17	3.4%
Jersey City	Low	20	4.0%
Millville	Low	20	4.0%
New Brunswick	Low	222	44.1%
Rahway	Low	36	7.2%
Camden	Medium	164	32.3%
Elizabeth	Medium	183	36.0%
Jersey City	Medium	34	6.7%
Flemington	Medium	0	0.0%
Millville	Medium	28	5.5%
New Brunswick	Medium	69	13.6%
Rahway	Medium	30	5.9%
Camden	High	44	8.6%
Elizabeth	High	320	62.4%
Flemington	High	0	0.0%
Jersey City	High	98	19.1%
Millville	High	5	1.0%
New Brunswick	High	1	0.2%
Rahway	High	45	8.8%

Number and percent of MIs assigned to each PM_{2.5}-monitoring-site-community by AER tertile

Appendix B12
Relative Odds of a Transmural Infarction Calculated at the Zip-Code Level

Tier	IQR	N	AIC	OR	95% CI	p-value
Tier 1	10.3	1561	4397.4	1.10	1.01, 1.19	0.03
Tier 2A SHEDS	5.4		4397.2	1.10	1.01, 1.20	0.03
Tier 1	10.3	1552*	4367.7	1.09	1.01, 1.19	0.04
Tier 2B APP	5.4		4366.8	1.10	1.10, 1.20	0.02
Tier 1	10.3	1561	4397.4	1.10	1.01, 1.19	0.03
Tier 3 HYBRID	5.4		4396.1	1.10	1.02, 1.20	0.01

Relative odds of a transmural infarction associated with an IQR increase in PM_{2.5} concentration, by exposure Tier. Refined exposure estimates were calculated at the zip-code level rather than at the one-value-per-monitor level presented in the main analysis. We observed no change in ORs, 95% CIs, nor AIC values for this spatial resolution compared to the main analysis.

* Subjects were excluded if there was a period of more than nine days between STN measurements for the case period or all control periods

Appendix B13
Study Population Characteristics by AER Tertile, Cool Season

	Total Transmural MI (n = 745) n(%)	Low AER Tertile (n = 244) n(%)	Middle AER Tertile (n = 302) n(%)	High AER Tertile (n=199) n(%)
Age (Years)				
18-44	58 (8)	20 (8)	22 (7)	16 (8)
45-54	152 (20)	47 (19)	58 (19)	47 (24)
55-64	184 (25)	64 (26)	69 (23)	51 (26)
65-74	151 (20)	51 (21)	61 (20)	39 (20)
75-84	141 (19)	40 (16)	70 (23)	31 (16)
≥ 85	59 (8)	22 (9)	22 (7)	15 (8)
Sex				
Male	460 (62)	157 (64)	180 (60)	123 (62)
Female	285 (38)	87 (36)	122 (40)	76 (38)
Race				
White	520 (70)	173 (71)	206 (68)	141 (71)
Black	85 (11)	18 (7)	38 (13)	29 (15)
Other	140 (19)	53 (22)	58 (19)	29 (15)
Year				
2004	209 (28)	70 (29)	90 (30)	49 (25)
2005	205 (28)	41 (17)	92 (30)	72 (36)
2006	331 (44)	133 (55)	120 (40)	78 (39)
Comorbidities				
Hypertension	411 (55)	135 (55)	169 (56)	107 (54)
Diabetes Mellitus	195 (26)	63 (26)	78 (26)	54 (27)
Type I Diabetes	4 (1)	0 (0)	3 (1)	1 (1)
Type II Diabetes	141 (19)	45 (18)	58 (19)	38 (19)
COPD	81 (11)	28 (11)	34 (11)	19 (10)
Pneumonia	31 (4)	12 (5)	9 (3)	10 (5)
Heart Disease	636 (85)	210 (86)	258 (85)	168 (84)

Appendix B14
Study Population Characteristics by AER Tertile, Warm Season

Characteristic	Total Transmural MI (n = 779) n(%)	Low AER Tertile (n = 259) n(%)	Middle AER Tertile (n = 206) n(%)	High AER Tertile (n=314) n(%)
Age (Years)				
18-44	76 (10)	25 (10)	20 (10)	31 (10)
45-54	167 (21)	51 (20)	40 (19)	76 (24)
55-64	207 (27)	71 (27)	60 (29)	76 (24)
65-74	140 (18)	52 (20)	32 (16)	56 (18)
75-84	128 (16)	41 (16)	36 (17)	51 (16)
≥ 85	61 (8)	19 (7)	18 (9)	24 (8)
Sex				
Male	493 (63)	168 (65)	123 (60)	202 (64)
Female	286 (37)	91 (35)	83 (40)	112 (36)
Race				
White	538 (69)	177 (68)	146 (71)	215 (68)
Black	87 (11)	29 (11)	21 (10)	37 (12)
Other	154 (20)	53 (20)	39 (19)	62 (20)
Year				
2004	226 (29)	65 (25)	77 (37)	84 (27)
2005	178 (23)	59 (23)	49 (24)	70 (22)
2006	375 (48)	135 (52)	80 (39)	160 (51)
Comorbidity				
Hypertension	435 (56)	146 (56)	116 (56)	173 (55)
Diabetes Mellitus	220 (28)	82 (32)	51 (25)	87 (28)
Type I Diabetes	9 (1)	3 (1)	4 (2)	1 (0.3)
Type II Diabetes	166 (21)	61 (24)	36 (17)	69 (22)
COPD	82 (10)	20 (8)	29 (14)	33 (11)
Pneumonia	26 (3)	10 (4)	5 (2)	11 (4)
Heart Disease	656 (84)	218 (84)	175 (85)	263 (84)

Frequency and percentage of study subjects included in the "AER Effect Modification" analysis and within AER tertiles in the warm season (May - October).

Appendix C: Supplemental Information for Chapter 4

Review of published PM_{2.5} species size distributions

A search was conducted in Web of Science with various combinations of keywords such as "sulfate", "elemental carbon," "black carbon," "organic matter," "organic carbon," "size-resolved species concentrations," "size distributions," "ambient PM_{2.5}," "impactor measurements," "Micro-Orifice Uniform Impactor (MOUDI)," "Aerosol Mass Spectrometer (AMS)," and "Supersite Project" in order to identify studies in which species-resolved size distribution measurements were conducted. Papers from diverse geographic locations and seasons (Appendix C3) were selected and were used to identify "typical" size distributions (i.e., number of modes, mass median diameter of each mode, and the fraction of mass in each mode) for particulate sulfate, elemental carbon, and organic carbon. These size-distribution characteristics were then used to inform the selection of particle-size-dependent model parameters (i.e. k_{dep} , P_{filter} ; Table 4-1).

Sulfate size distributions could be characterized by a bimodal size distribution with a smaller condensation mode (comprising 20% of total sulfate mass) with a peak of 0.2 μm and a larger droplet mode (comprising 80% of the sulfate mass) with a peak of 0.7 μm (Appendix C3).^{e.g. 1-3} Organic carbon (OC) was represented by a trimodal size distribution with an ultrafine mode peaking at 0.08 μm and an accumulation mode internally mixed with sulfate (i.e., 20% in the condensation mode and 80% in the droplet mode; Appendix C3).^{e.g. 4-6} We assumed the ultrafine mode of OC, which comprised 40% of OC mass,⁷ was associated with primary OC emissions and was internally mixed with EC. EC was, therefore, characterized by a unimodal distribution with a 0.08 μm peak.

RIOPA Study Measurements and Data

Measured air exchange rates (AERs) and PM_{2.5} species mass concentrations measured outside of homes sampled during the Relationships of Indoor, Outdoor, and Personal Air (RIOPA) study were used as model inputs in the main analysis to calculate indoor ambient PM_{2.5} species concentrations. Modeled indoor concentrations of ambient PM_{2.5} were then evaluated against measured indoor PM_{2.5} species mass concentrations. Questionnaires documenting the activities of RIOPA study participants and gas- and particle-phase polycyclic aromatic hydrocarbon (PAH) concentration measurements were also used to evaluate the extent to which model-measurement agreement could be improved by accounting for specific human activities and shifts in the gas-particle partitioning of ambient organics, respectively. Details regarding the RIOPA study design, measurement methods, and quality control/quality assurance are provided elsewhere.^{8,9} Below, we provide a brief description of the RIOPA data and measurements utilized in this study.

AERs were measured with a perfluorocarbon tracer method. Perfluorinated methyl cyclohexane (PMCH) was released in each home at a known emission rate over the 48 hour sampling period and was collected with a charcoal adsorption tube passive sampler. PMCH concentrations were then analyzed using gas chromatography with an electron capture detector. Because the emission rate of the PMCH was known, the measured concentration of PMCH in the home served as an indicator of the rate at which air in the home was exchanged with outdoor air.⁸ Indoor and outdoor elemental carbon (EC) and organic carbon (OC) samples were collected on quartz fiber filters with a MSP microenvironmental PM_{2.5} sampler (2.5 µm aerodynamic diameter inlet cutpoint) and EC and OC concentrations were measured by thermal-optical transmittance with a Sunset

Carbon Analyzer. OC concentrations were corrected for sampling artifacts associated with the adsorption of gas-phase semi-volatile organic compounds on quartz fiber filters and are reported as $\mu\text{gC}/\text{m}^3$. $\text{PM}_{2.5}$ filter samples collected with a Harvard Impactor with a $2.5\ \mu\text{m}$ aerodynamic diameter cutpoint were analyzed for sulfur with an energy-dispersive XRF spectrometer. Concentrations ($\mu\text{gS}/\text{m}^3$) were multiplied by 3 ($96\ \text{g mol}^{-1}$ sulfate/ $32\ \text{g mol}^{-1}$ sulfur) to obtain $\mu\text{g}/\text{m}^3$ as sulfate. Gas- and particle-phase PAHs were collected with the MSP microenvironmental $\text{PM}_{2.5}$ sampler, which was modified to include a polyurethane foam cartridge to collect gas-phase semi-volatile organic compounds.¹⁰ PAH samples were extracted and then analyzed with gas chromatography-mass spectrometry¹⁰ and reported as compounds mass (ng) per cubic meter of air. Activity questionnaires, which were designed based on the National Human Exposure Assessment Survey (NHEXAS),¹¹ documented home-occupant activities over the time period that AER and pollutant concentration measurements were being conducted.⁸ Summary statistics for measured AERs, $\text{PM}_{2.5}$ species concentrations, and gas- and particle-phase PAH measurements for the RIOPA homes included in this study (Appendix C1) are given in Appendix C2.

Use of robust regression to estimate the contribution of outdoor-generated $\text{PM}_{2.5}$ to measured indoor $\text{PM}_{2.5}$

In the main analysis, we regressed measured indoor particulate OC on measured outdoor OC using robust regression to estimate the fraction of measured indoor OC that could be attributed to outdoor sources. The slope of the resulting model represents a population average of the fraction of ambient OC that has penetrated and persisted in the indoor environment (F). To evaluate whether robust regression provides a reasonable

estimate of the population-average F , we repeated this robust regression analysis for sulfate and compared the distribution of indoor sulfate calculated using robust-regression with the distributions of measured indoor sulfate and the distribution of indoor sulfate modeled with the refined mass-balance model. Sulfate has minimal indoor sources and, thus, the distribution of indoor sulfate calculated using robust regression should agree well with distributions of measured and modeled indoor sulfate concentrations. We performed this robust regression analysis on homes with open windows, homes with central air conditioning or heating in use, and closed homes without central heating or air conditioning in use separately because we found that these human activities were significant contributors to variability in F for sulfate in the main analysis.

F values for sulfate estimated using robust regression were 0.44 for homes with central air conditioning or heating in use, 0.54 for closed homes without central air conditioning or heating in use, 0.88 for homes with open windows, and averaged 0.69 across all homes. This was in agreement with the average measured F value (0.73; calculated as the ratio of measured indoor sulfate:measured outdoor sulfate) and that estimated using the mass-balance model (0.72). Appendix C4 shows cumulative distributions of measured indoor sulfate, indoor sulfate estimated with the refined model, and indoor sulfate estimated using robust-regression. The distributions of indoor sulfate are in good agreement, suggesting that robust regression provides a reasonable estimate of the population-average F value.

Appendix C References

1. Hering, S. V.; Friedlander, S. K., Origins of sulfur size distributions in the Los Angeles basin. *Atmos. Environ.* **1982**, 11, 2647 – 2656.

2. Hering, S. V.; Eldering, A.; Seinfeld, J. H., Bimodal character of accumulation mode aerosol mass distributions in Southern California. *Atmos. Environ.* **1997**, 31, 1 – 11.
3. John, W.; Wall, S. M.; Ondo, J. L.; Winklmayr, W., Modes in the size distributions of atmospheric inorganic aerosol. *Atmos. Environ.* **1990**, 24A, 2349 - 2359.
4. Lee, S.; Murphy, D. M.; Thomson, D. S.; Middlebrook, A. M., Nitrate and oxidized organic ions in single particle mass spectra during the 1999 Atlanta Supersite Project. *J. Geophys. Res.* **1999**, 108, doi:10.1029/2001JD001455.
5. Bein, K. J.; Zhao, Y.; Wexler, A. S.; Johnston, M. V., Speciation of size-resolved individual ultrafine particles in Pittsburgh, Pennsylvania, *J. Geophys. Res.* **2005**, 110, doi:10.1029/2004JD004708.
6. Zhang, Q.; Canagaratna, M. R.; Jayne, J. T.; Worsnop, D. R.; Jimenez, J. L., Time- and size-resolved chemical composition of submicron particles in Pittsburgh: Implications for aerosol sources and processes. *J. Geophys. Res.* **2005**, 110, D07S09, doi:10.1029/2004JD004649.
7. Drewnick, F.; Jayne, J. T.; Canagaratna, M.; Worsnop, D. R.; Demerjian, K. L., Measurement of ambient aerosol composition during the PMTACS-NY 2001 using an aerosol mass spectrometer. Part II: chemically speciated mass distributions. *Aerosol Sci. Technol.* **2004**, 38, 104 - 117.
8. Weisel, C. P.; Zhang, J.; Turpin, B. J.; Morandi, M. T.; Colome, S.; Stock, T. H.; Spector, D. M.; Korn, L.; Winer, A.; Alimokhtari, S.; Kwon, J.; Mohan, K.; Harrington, R.; Giovanetti, R.; Cui, W.; Afshar, M.; Maberti, S.; Shendell, D., Relationship of Indoor, Outdoor and Personal Air (RIOPA) study: study design, methods and quality assurance/control results. *J. Exposure Sci. Environ. Epidemiol.* **2005**, 15, 123-137.
9. Turpin, B. J.; Weisel, C. P.; Morandi, M.; Colome, S.; Eisenreich, S.; Buckley, B., Relationships of Outdoor Indoor and Personal Air (RIOPA): part II. Analysis of concentrations of particulate matter species. Research Report (Health Effects Institute) **2007**, 130, 79-92.
10. Naumova, Y. Y.; Eisenreich, S. J.; Turpin, B. J.; Weisel, C. P.; Morandi, M. T.; Colome S. D.; Totten, L. A.; Stock, T. H.; Winer, A. M.; Alimokhtari, S.; Kwon, J.; Shendell, D.; Jones, J.; Maberti, S.; Wall, S. J., Polycyclic aromatic hydrocarbons in the indoor and outdoor air of three cities in the U.S. *Environ. Sci. Technol.* **2002**, 36, 2552-2559.
11. Sexton, K.; Kleffman, D. E.; Callahan, M. A., An introduction to the National Human Exposure Assessment Survey (NHEXAS) and related phase I field studies. *J. Expos. Anal. Environ. Epidemiol.* **1995**, 5, 229–232.

12. Turpin, B. J.; Saxena, P.; Allen, G.; Koutrakis, P.; McMurry, P.; Hildemann, L., Characterization of the southwestern desert aerosol, Meadview, AZ. *J. Air Waste Manag. Assoc.* **1997**, 47, 344-356.
13. Lowenthal, D. H.; Watson, J. G.; Saxena, P., Contributions to light extinction during project MOAHVE. *Atmos. Environ.* **2000**, 34, 2351-2359.
14. Hughes, L. S.; Cass, G. R.; Gone, J.; Ames, M.; Olmez, I., Physical and chemical characterization of ultrafine particles in the Los Angeles area, *Environ. Sci. Technol.* **1998**, 32, 1153-1161.
15. Kleeman, M. J.; Hughes, L. S.; Allen, J. O.; Cass, G. R., Source contributions to the size and composition distribution of atmospheric particles: Southern California in September 1996. *Environ. Sci. Technol.* **1996**, 33, 4331-4341.
16. Offenberg, J. H.; Baker, J. E., Aerosol size distributions of elemental and organic matter in urban and over-water atmosphere. *Atmos. Environ.* **2000**, 34, 1509-1517.
17. Zhaung, H.; Chan, C. K.; Fang, M.; Wexler, A. S., Size distributions of particulate sulfate, nitrate, and ammonium at a coastal site in Hong Kong. *Atmos. Environ.* **1999**, 33, 843-854.
18. Maehunt, W.; Cafmeyer, J.; Dubtsov, S.; Chi, X., Detailed mass size distributions of elements and species, and aerosol chemical mass closure during fall 1999 at Gent, Belgium. *Nuclear Instrum. Methods* **2002**, 189, 238-242.
19. Singh, M.; Jaques, P. A.; Sioutas, C., Size distribution and diurnal characteristics of particle-bound metals in source and receptor sites of the Los Angeles Basin. *Atmos. Environ.* **2002**, 36, 1675-1689.
20. Allan, J. D.; Alfarra, M. R.; Bower, K. N.; Williams, P. I.; Gallagher, M. W.; Jimenez, J. L.; McDonald, A. G.; Nemitz, E.; Canagaratna, M. R.; Jayne, J. T.; Coe, H.; Worsnop, D. R., Quantitative sampling using an Aerodyne aerosol mass spectrometer 2. Measurements of fine particulate chemical composition in two U.K. cities, *J. Geophys. Res.* **2003**, 108, doi:10.1029/2002JD002359.
21. Alfarra, M. R.; Coe, H.; Allan, J. D.; Bower, K. N.; Boudries, H.; Canagaratna, M. R.; Jimenez, J. L.; Jayne, J. T.; Garforth, A. A.; Li, S. M.; Worsnop, D. R., Characterization of urban and rural organic particulate matter in the Lower Fraser Valley using two Aerodyne Aerosol Mass Spectrometers, *Atmos. Environ.* **2004**, 38, 5745-5758.
22. Boudries, H.; Canagaratna, M. R.; Jayne, J. T.; Alfarra, M. R.; Allan, J.; Bower, K. N.; Coe, H.; Pryor, S. C.; Jimenez, J. L.; Brook, J. R.; Li, S.; Worsnop, D. R., Chemical and physical processes controlling the distribution of aerosols in the Lower

Fraser Valley, Canada, during the Pacific 2001 campaign, *Atmos. Environ.* **2004**, 38, 5759-5774.

23. Cabada, J. C.; Ress, S.; Takahama, S.; Khlystov, A.; Pandis, S. N.; Davidson, C. I.; Robinson, Al. L., Mass size distributions and size resolved chemical composition of fine particulate matter at the Pittsburgh supersite, *Atmos. Environ.* **2004**, 38, 3127-3141.

24. Miguel, A. H.; Eiguren-Fernandez, A.; Jaques, P. A.; Froines, J. R.; Grant B. L.; Mayo, P. R.; Sioutas, C., Seasonal variation of the particle size distribution of polycyclic aromatic hydrocarbons and of major aerosol species in Claremont, CA. *Atmos. Environ.* **2004**, 38, 3241-3251.

25. Chuaybamroong, P.; Cayse, K.; Wu, C. Y.; Lundgren, D., Ambient aerosol and its carbon content in Gainesville, a mid-scale city in Florida, *Environ. Monit. Assess.* **2007**, 128, 421-430.

26. Chow, J. C.; Watson, J. G.; Lowenthal, D. H.; Magliano, K. L., Size-resolved aerosol chemical concentrations at rural and urban sites in Central California, USA, *Atmos. Res.* **2008**, 90, 243-252.

Appendix C1
Number of RIOPA Study Homes Included in Chapter 4 Analyses

	Number of Homes		
	Elemental Carbon	Sulfate	Organic Carbon
Total	141	205	139
Los Angeles County, CA	41	84	41
Elizabeth, NJ	49	57	48
Houston, TX	51	64	50

Number of homes for which all data required to estimate indoor ambient $PM_{2.5}$ concentrations with the mass-balance model were available and for which measured indoor species concentrations were not considered outliers based on the student's t criteria discussed in the main analysis. Outliers are likely indicators of strong indoor sources and those homes were excluded to avoid influence of strong indoor sources of $PM_{2.5}$ on model-evaluation results, as the mass-balance model is meant to predict indoor *ambient* $PM_{2.5}$ concentrations. RIOPA study measurements were conducted in the three geographically and climatically distinct urban regions shown above in order to provide a diverse sample of air exchange rates, housing types, and ambient $PM_{2.5}$ source mix.^{8,9}

Appendix C2
Summary Statistics of RIOPA Study Measurements

	Outdoors			Indoors		
	Mean	Median	Standard Deviation	Mean	Median	Standard Deviation
Air Exchange Rate (h^{-1})	-	-	-	1.13	0.83	0.96
Elemental Carbon ($\mu\text{g}/\text{m}^3$)	1.13	0.99	0.73	1.00	0.89	0.87
Sulfate ($\mu\text{g}/\text{m}^3$)	3.26	2.98	1.87	2.37	1.84	1.60
Organic Carbon ($\mu\text{g}/\text{m}^3$)	3.61	3.30	2.38	6.08	5.16	3.77
Gas-Phase PAHs (ng/m^3)	0.04	0.02	0.10	0.02	0.02	0.02
Particle-Phase PAHs (ng/m^3)	1.88	1.12	2.34	1.53	0.77	2.71

Summary statistics of measured air exchange rate (AER), indoor and outdoor $\text{PM}_{2.5}$ species mass concentrations, and indoor and outdoor gas- and particle-phase 5 - 7 ring PAHs for the RIOPA homes studied. AER and measured outdoor $\text{PM}_{2.5}$ species concentrations were used as model inputs. Modeled indoor ambient $\text{PM}_{2.5}$ species concentrations were evaluated against measured indoor $\text{PM}_{2.5}$ species concentrations. Pooled gas-particle partitioning coefficients of the 5 - 7 ring PAHs were used as surrogates for partitioning of total ambient OC.

Appendix C3

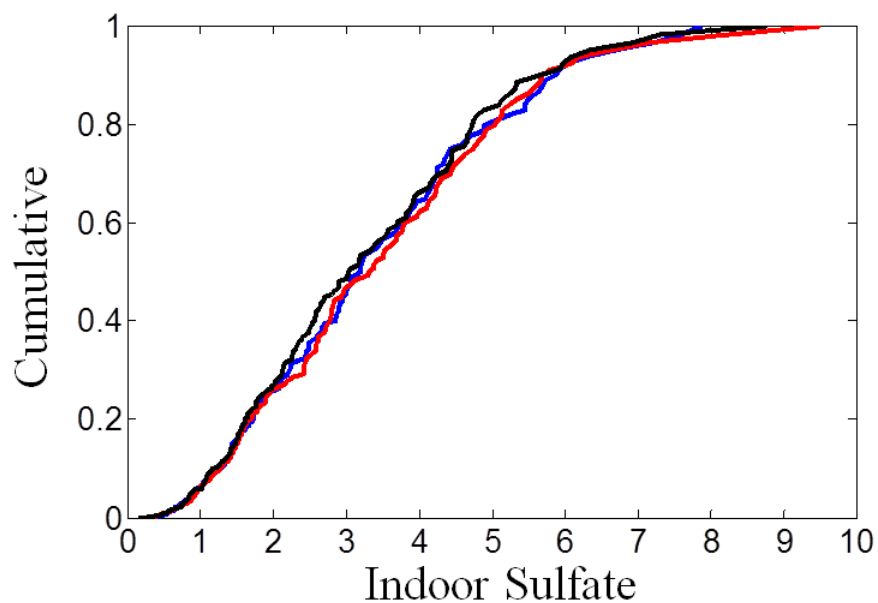
Particle Size Distribution Literature Review

Papers	Campaign/Location	Seasons	PM Species
Hering and Friedlander ¹	Pasadena, CA; Freeway Sulfur Experiment (Los Angeles Basin)	All seasons; Fall	sulfate
John et al. ³ Hering et al. ²	Southern California Air Quality Study (SCAQS; Riverside, Claremont, Long Beach, Los Angeles)	Summer, Fall	accumulation mode PM _{2.5} (sum of sulfate, nitrate, ammonium, elemental carbon (EC), organics); sulfate
Turpin et al. ¹² Lowenthal et al. ¹³	Meadview, AZ	Summer	sulfate, organics
Hughes et al. ¹⁴	Pasadena, CA	Winter	ultrafine-mode organics and elemental carbon
Kleeman et al. ¹⁵	Long Beach, Fullerton, and Riverside, CA	Fall	sulfate, organics, elemental carbon
Offenburg and Baker ¹⁶	Chicago, IL	Summer	organics, elemental carbon
Zhuang et al. ¹⁷	Hong Kong	Winter	Sulfate
Maehunt et al. ¹⁸	Gent, Belgium	Fall	sulfate, organics, elemental carbon
Singh et al. ¹⁹	Downey and Riverside, CA	All seasons	sulfate, organics, elemental carbon
Allan et al. ²⁰	Edinburgh, U.K.; Manchester, U.K	Fall; Summer, Winter	sulfate, organics
Lee et al. ⁴	Atlanta, GA Supersite Project	Summer	sulfate, organics
Alfarra et al. ²¹ Boudries et al. ²²	Lower Fraser Valley, British Columbia, Canada	Summer	sulfate, organics
Cabada et al. ²³ Bein et al. ⁵ Zhang et al. ⁶	Pittsburgh, PA Air Quality Study (PAQS)	All seasons	sulfate, organics, elemental carbon
Drewnick et al. ⁷	PM _{2.5} Technology Assessment and Characterization Study-New York (PM-TACS-NY; Queens, NY)	Summer	sulfate, organics
Miguel et al. ²⁴	Southern California Particle Center and Supersite (SCPCS; Claremont, CA)	All seasons	sulfate, organics, elemental carbon
Chuaybamroong et al. ²⁵	Gainesville, FL	Winter, Spring, Summer	organics, elemental carbon
Chow et al. ²⁶	California Regional PM ₁₀ /PM _{2.5} Air Quality Study (CRPAQS; Angiola and Fresno, CA)	Winter	sulfate, organics, elemental carbon

Summary of review of published species size distributions from diverse geographic locations and seasons conducted to determine "typical" size distributions (i.e., number of modes, mass median diameter of each mode, and the fraction of mass in each mode) for particulate sulfate, elemental carbon, and organic carbon.

Appendix C4

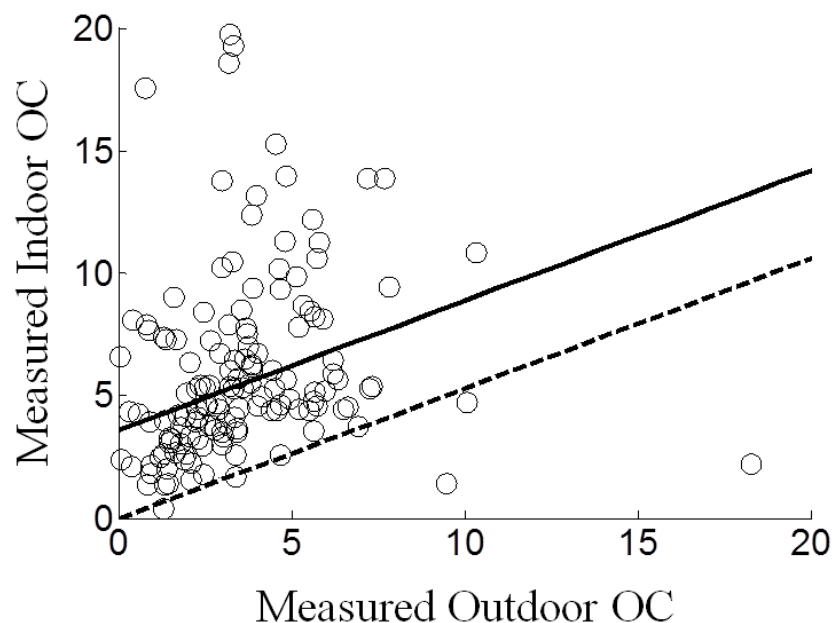
Cumulative Distributions of Indoor Sulfate



Cumulative distributions of indoor sulfate ($\mu\text{g}/\text{m}^3$): measured (blue), calculated with the refined mass-balance model (red), and estimated using the robust-regression-estimated population-average F value (black).

Appendix C5

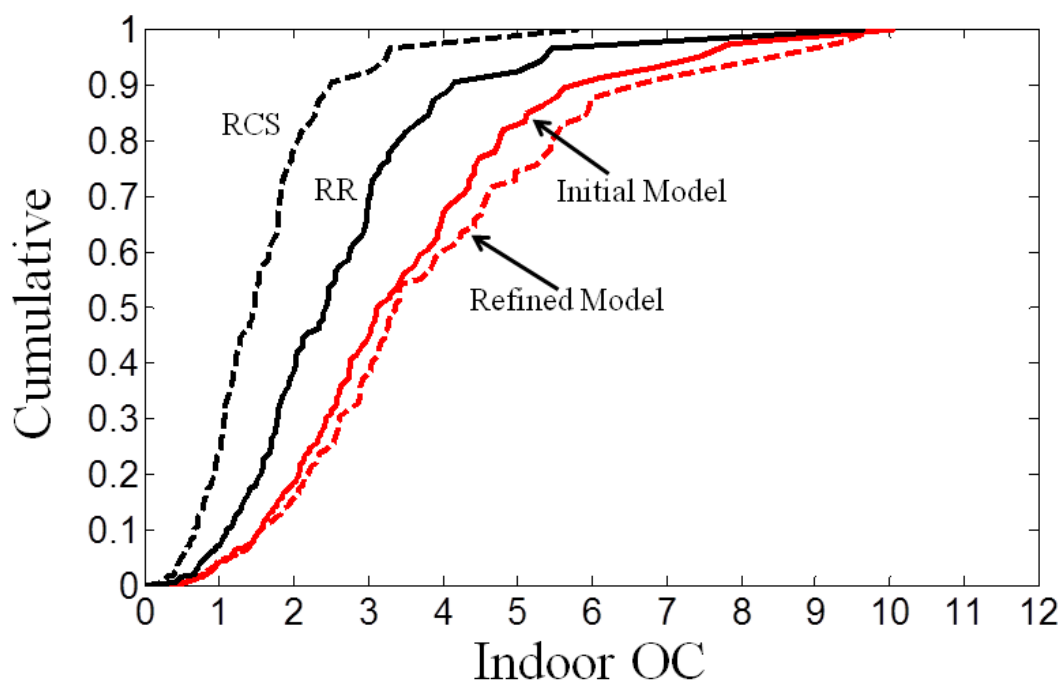
Robust Regression of Measured Indoor OC on Measured Outdoor Organic Carbon



Robust regression of measured indoor particulate organic carbon ($C_{in,OC}$) on measured outdoor particulate OC ($C_{out,OC}$), both in $\mu\text{g}/\text{m}^3$. The solid line is the regression equation ($C_{in,OC} = 3.59 \mu\text{gC}/\text{m}^3 + 0.53 * C_{out,OC}$). The intercept of the regression equation provides an average indoor source strength and the slope is a population-average estimate of the fraction of ambient OC that penetrated and persisted indoors (F). The estimated F value was multiplied by the OC concentration measured outside of each RIOPA home to calculate the distribution of measured indoor OC that could be attributed to outdoor sources (dashed line).

Appendix C6

Cumulative Distributions of Indoor Organic Carbon



Cumulative distributions of indoor organic carbon (OC) of ambient origin ($\mu\text{gC}/\text{m}^3$): calculated with the initial model (red solid), calculated with the refined model (accounting for open windows and central AC and heating use; red dashed), calculated using the robust-regression-estimated (RR) population-average F value (black solid), and calculated using the Random Component Superposition (RCS) estimated population-average F value from Polidori et al. (2006) (black dashed).

Appendix C7

SAS Code for Chapter 4 Calculations: Elemental Carbon

*Change to directory on accompanying flash drive with data sets containing measured PM2.5 species concentrations, measured air exchange rates, and activity data. Note: the drive label might be different for each computer;

```
libname Ch4 'G:\Chapter 4';
```

*Calculate indoor concentrations of EC with "initial" mass balance model assuming:

1. particles w/ diameter of 0.08 μm , $k_{\text{dep}} = 0.05 \text{ h}^{-1}$;

2. $P = 0.8$;

```
data EC_initial;
```

```
set Ch4.EC_AER_ACT; *data sheet with measured EC, measured AERs, and home-occupant activity data;
```

```
APP_initial = Out_Measured*((AER*0.8)/(AER+0.05)); *calculates indoor concentrations;
```

```
residual = APP_initial - In_Measured; *difference between measured and modeled values;
```

```
run;
```

```
*****
*****
```

Multiple Linear Regression of model-measurement differences on home-occupant activity variables;

*Identify Collinearities and outliers;

```
proc reg data = EC_initial;
```

```
model Residual = Hours_win Hours_cenAC Hours_cenHeat /
```

```
VIF COLLINOINT;
```

```
output out = ttest_EC /*if >= +/- 2 then an outlier*/
```

```
student = student /*if >= +/- 2 then an outlier*/
```

```
rstudent = rstudent;
```

```
run;
```

```
quit;
```

*for collinearity variance inflation>10, variance proportion >0.5 and condition indices>3.6;

*no collinearity detected. Remove outliers;

```
data EC_MLR1;
```

```
set ttest_EC;
```

```
if student>=2 or student <=-2 then delete;
```

```
drop student rstudent;
```

```
run;
```

*using this clean data set (no outliers) create the MLR model;

```
proc reg data = EC_MLR1;
```

```
model residual = Hours_win Hours_cenAC Hours_cenHeat /selection  
=stepwise;
```

```
ods output selparmest = selparmest_EC selectionsummary =
```

```
selectionsummary_EC; *generates tables with regression model selection  
summary and parameter estimates;
```

```

run;
quit;

*None of the activity variables met the P = 0.1500 significance level
for entry into the model;

*Note: 80 nm particles have high filter penetration efficiencies,
suggesting overall all P values might be larger for fine mode PM.
Recalculate indoor EC assuming overall P value 0.9;

data EC_refined;
set EC_MLR1;
APP_refined = Out_Measured*((AER*0.9)/(AER+0.05)); *calculates indoor
concentrations;
run;

*****

Paired t-test to evaluate whether pairs of measured and modeled
(initial) indoor EC are significantly different at 95% confidence
level;

proc ttest data=EC_MLR1;
paired In_Measured*APP_initial;
run;

*chi-square test to evaluate whether measured and modeled indoor EC
have the same underlying distribution at a 95% confidence level;

*put data into needed format for chi-square test;

*isolate measured indoor EC;
data EC_meas;
set EC_MLR1;
drop APP_initial; *remove modeled values from data set;
category = "Measured"; *add category for chi-square test;
Indoor = In_Measured;
drop In_Measured;
run;

*isolate modeled indoor EC;
data EC_mod;
set EC_MLR1;
drop In_Measured; *remove measured values from data set;
category = "Modeled"; *add category for chi-square test;
Indoor = APP_initial;
drop APP_initial;
run;

*merge measured and modeled indoor EC data sets;
data EC_meas_mod;
set EC_meas EC_mod;
run;

*chi-square test;

```

```

proc npar1way data = EC_meas_mod wilcoxon;
  class category;
  var indoor;
run;
*****

Paired t-test to evaluate whether pairs of measured and modeled
(refined) indoor EC are significantly different at 95% confidence
level;

proc ttest data=EC_refined;
paired In_Measured*APP_refined;
run;

*chi-square test to evaluate whether measured and modeled indoor EC
have the same underlying distribution at a 95% confidence level;

*put data into needed format for chi-square test;

*isolate measured indoor EC;
data EC_meas;
set EC_refined;
drop APP_refined; *remove modeled values from data set;
category = "Measured"; *add category for chi-square test;
Indoor = In_Measured;
drop In_Measured;
run;

*isolate modeled indoor EC;
data EC_mod_r; *r denotes "refined" model;
set EC_refined;
drop In_Measured; *remove measured values from data set;
category = "Modeled"; *add category for chi-square test;
Indoor = APP_refined;
drop APP_refined;
run;

*merge measured and modeled indoor EC data sets;
data EC_meas_mod_r; *r denotes "refined" model;
set EC_meas EC_mod_r;
run;

*chi-square test;
proc npar1way data = EC_meas_mod_r wilcoxon;
  class category;
  var indoor;
run;
*****

Paired t-test to evaluate whether pairs of measured indoor and measured
outdoor EC are significantly different at 95% confidence level;

```

```

proc ttest data= EC_MLR1;
paired In_Measured*Out_measured;
run;

*chi-square test measured outdoor vs measured indoor;
data EC_in;
set EC_MLR1;
drop Out_measured;
category = "Indoor";
Indoor = In_Measured;
drop In_Measured;
run;

data EC_out;
set EC_MLR1;
drop In_Measured;
category = "Outdoor";
Indoor = Out_measured;
drop Out_measured;
run;

data EC_in_out;
set EC_in EC_out;
run;

proc npar1way data = EC_in_out wilcoxon;
  class category;
  var indoor;
run;

```


Appendix C8

SAS Code for Chapter 4 Calculations: Sulfate

*Change to directory on accompanying flash drive with data sets containing measured PM2.5 species, concentrations, measured air exchange rates, and activity data. Note: the drive label might be different for each computer;

```
libname Ch4 'G:\Chapter 4';
```

*Calculate indoor concentrations of sulfate with "initial" mass balance model assuming:

1. 20% mass comprised of particles w/ diameter of 0.2 μm , $k_{\text{dep}} = 0.05 \text{ h}^{-1}$
2. 80% mass comprised of particles w/ diameter of 0.7 μm , $k_{\text{dep}} = 0.13 \text{ h}^{-1}$
3. $P = 0.8$;

```
data sulf_initial;
set Ch4.Sulf_AER_ACT;
APP_initial =
(0.8*Out_measured*((AER*0.8)/(AER+0.13)))+(0.2*Out_measured*((AER*0.8)/
(AER+0.05)));
residual = APP_initial - In_measured;
run;
```

```
*****
*****
```

Multiple Linear Regression of model-measurement differences on home-occupant activity variables;

```
*Identify Collinearities and outliers;
proc reg data = sulf_initial;
model Residual = Hours_win Hours_cenAC Hours_cenHeat /
VIF COLLINOINT;
output out = ttest_sulf /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;
```

*for collinearity variance inflation>10, variance proportion >0.5 and condition indices>3.6;
*no collinearity detected. Remove outliers;

```
data sulf_MLR1;
set ttest_sulf;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;
```

```
*using this clean data set (no outliers) create the MLR model;
proc reg data = sulf_MLR1;
model residual = Hours_win Hours_cenAC Hours_cenHeat /selection
=stepwise;
```

```
ods output selparmest = selparmest_sulf selectionsummary =
selectionsummary_sulf; *generates tables with regression model
selection summary and parameter estimates;
run;
quit;
```

*All activity variables selected as significant predictors of model-measurement differences for sulfate.

Calculate indoor concentrations with the "refined" mass balance model assuming:

1. 20% mass comprised of particles w/ diameter of 0.2 μm , $k_{\text{dep}} = 0.05 \text{ h}^{-1}$, $P_{\text{filter}} = 0.90$ (when AC or heating in use only);
2. 80% mass comprised of particles w/ diameter of 0.7 μm , $k_{\text{dep}} = 0.13 \text{ h}^{-1}$, $P_{\text{filter}} = 0.65$ (when AC or heating in use only);
3. $P = 0.8$ for closed homes, $P = 1.0$ for homes with open windows;

```
data sulf_refined;
set sulf_MLR1;
if Hours_win = 0 and Hours_cenAC = 0 and Hours_cenHeat = 0 then
APP_refined =
(0.8*Out_measured*((AER*0.8)/(AER+0.13)))+(0.2*Out_measured*((AER*0.8)/(
AER+0.05)));
if Hours_win > 0 and Hours_win >Hours_cenAC and Hours_win
>Hours_cenHeat then APP_refined =
(0.8*Out_measured*((AER)/(AER+0.13)))+(0.2*Out_measured*((AER)/(AER+0.0
5)));
if Hours_cenAC>0 and Hours_win <Hours_cenAC and Hours_cenAC
>Hours_cenHeat then APP_refined =
(0.8*Out_measured*((AER*0.8)/(AER+0.13))*0.65)+(0.2*Out_measured*((AER*
0.8)/(AER+0.05))*0.90);
if Hours_cenHeat>0 and Hours_win <Hours_cenHeat and Hours_cenAC
<Hours_cenHeat then APP_refined =
(0.8*Out_measured*((AER*0.8)/(AER+0.13))*0.65)+(0.2*Out_measured*((AER*
0.8)/(AER+0.05))*0.90);
run;
```

```
*****
*****
```

Paired t-test to evaluate whether pairs of measured and modeled (initial model) indoor sulfate are significantly different at 95% confidence level;

```
proc ttest data=sulf_MLR1;
paired In_Measured*APP_initial;
run;
```

*chi-square test to evaluate whether measured and modeled indoor (initial model) sulfate have the same underlying distribution at a 95% confidence level;

*put data into needed format for chi-square test;

```
*isolate measured indoor sulfate;
data sulf_meas;
set sulf_MLR1;
```

```

drop APP_Initial; *remove modeled values from data set;
category = "Measured"; *add category for chi-square test;
Indoor = In_Measured;
drop In_Measured;
run;

*isolate modeled (initial) indoor sulfate;
data sulf_mod_i; *"i" denotes initial model;
set sulf_MLR1;
drop In_Measured; *remove measured values from data set;
category = "Modeled"; *add category for chi-square test;
Indoor = APP_Initial;
drop APP_Initial;
run;

*merge measured and modeled indoor sulfate data sets;
data sulf_meas_mod_i;
set sulf_meas sulf_mod_i;
run;

*chi-square test;
proc nparway data = sulf_meas_mod_i wilcoxon;
  class category;
  var indoor;
run;

*****
*****

Paired t-test to evaluate whether pairs of measured and modeled
(refined model) indoor sulfate are significantly different at 95%
confidence level;

proc ttest data= sulf_refined;
paired In_Measured*APP_refined;
run;

*chi-square test to evaluate whether measured and modeled (refined
model) indoor sulfate have the same underlying distribution at a 95%
confidence level;

*put data into needed format for chi-square test;

*isolate measured indoor sulfate;
data sulf_meas;
set sulf_refined;
drop APP_refined; *remove modeled values from data set;
category = "Measured"; *add category for chi-square test;
Indoor = In_Measured;
drop In_Measured;
run;

*isolate modeled (refined) indoor sulfate;
data sulf_mod_r; *"r" denotes refined model;

```

```

set sulf_refined;
drop In_Measured; *remove measured values from data set;
category = "Modeled"; *add category for chi-square test;
Indoor = APP_refined;
drop APP_refined;
run;

*merge measured and modeled indoor sulfate data sets;
data sulf_meas_mod_r;
set sulf_meas sulf_mod_r;
run;

*chi-square test;
proc nparlway data = sulf_meas_mod_r wilcoxon;
  class category;
  var indoor;

*****

Paired t-test to evaluate whether pairs of measured indoor and measured
outdoor sulfate are significantly different at 95% confidence level;

proc ttest data= sulf_MLR1;
paired In_Measured*Out_measured;
run;

*chi-square test measured outdoor vs measured indoor;
data sulf_in;
set sulf_MLR1;
drop Out_measured;
category = "Indoor";
Indoor = In_Measured;
drop In_Measured;
run;

data sulf_out;
set sulf_MLR1;
drop In_Measured;
category = "Outdoor";
Indoor = Out_measured;
drop Out_measured;
run;

data sulf_in_out;
set sulf_in sulf_out;
run;

proc nparlway data = sulf_in_out wilcoxon;
  class category;
  var indoor;
run;

*****

```

Robust regression estimate of indoor sulfate of outdoor origin.

Note: regression analysis performed separately for homes with open windows, homes with central AC or heating in use, and closed homes (i.e. windows closed) w/o central AC or heating in use b/c we found that these activities are significant contributors to variability in the fraction of ambient sulfate found indoors (above);

*Break up the homes by activity before calculating population-average F with robust regression;

```
*closed homes;
data S_RR_closed;
set sulf_MLR1;
if Hours_win>0 then delete; *remove all homes except those that are
closed;
if Hours_cenAC>0 then delete;
if Hours_cenHeat>0 then delete;
run;

*robust regression for closed homes;

proc robustreg data = S_RR_closed method=LTS FWLS;
model In_measured = Out_measured;
run;

*homes w/ open windows;
data S_RR_win;
set sulf_MLR1;
if Hours_win=0 then delete; *remove all homes except those w/ open
windows;
if Hours_cenAC>Hours_win then delete;
if Hours_cenHeat>Hours_win then delete;
run;

*robust regression for homes w/ open windows;
proc robustreg data = S_RR_win method=LTS FWLS;
model In_measured = Out_measured;
run;

*homes w/ open AC or heat in use;
data S_RR_heatAC;
set sulf_MLR1;
if Hours_cenAC=0 and Hours_cenHeat = 0 then delete; *remove all homes
except those w/ central AC or heat in use;
if Hours_cenAC<Hours_win and Hours_cenHeat < Hours_win then delete;
run;

*robust regression for homes w/ heating or central AC in use;
proc robustreg data = S_RR_heatAC method=LTS FWLS;
model In_measured = Out_measured;
run;
```

```

*Now, create data set with robust regression estimates of indoor
sulfate for all homes;
data S_RobReg;
set sulf_MLR2;
if Hours_win = 0 and Hours_cenAC = 0 and Hours_cenHeat = 0 then RobReg
= Out_measured*0.54;
if Hours_win > 0 and Hours_win > Hours_cenAC and Hours_win
>Hours_cenHeat then RobReg = Out_measured*0.88;
if Hours_cenAC > 0 and Hours_win < Hours_cenAC and Hours_cenAC
>Hours_cenHeat then RobReg = Out_measured*0.44;
if Hours_cenHeat > 0 and Hours_win < Hours_cenHeat and Hours_cenAC
<Hours_cenHeat then RobReg = Out_measured*0.44;
run;

```

Appendix C9

SAS Code for Chapter 4 Calculations: Organic Carbon

*Change to directory on accompanying flash drive with data sets containing measured PM2.5 species, concentrations, measured air exchange rates, and activity data. Note: the drive label might be different for each computer;

```
libname Ch4 'G:\Chapter 4';
```

*Calculate indoor concentrations of OC with "initial" mass balance model assuming trimodal distributions with fine mode (40% mass) internally mixed with EC and accumulation mode (60% mass) internally mixed with sulfate:

1. 40% mass in fine mode with diameter 0.08, $k_{dep} = 0.05 \text{ h}^{-1}$
2. 12% mass in condensation mode with diameter 0.2, $k_{dep} = 0.05 \text{ h}^{-1}$
3. 48% mass in droplet mode with diameter 0.7, $k_{dep} = 0.13 \text{ h}^{-1}$
4. $P = 0.8$;

```
data OC_initial;
set Ch4.OC_aer_act;
APP_initial =
(0.4*Out_measured*((AER*0.8)/(AER+0.05)))+(0.12*Out_measured*((AER*0.8)
/(AER+0.05)))+(0.48*Out_measured*((AER*0.8)/(AER+0.13)));
residual = APP_initial - In_measured;
run;
```

```
*****
*****
```

Multiple Linear Regression of model-measurement differences on home-occupant activity variables;

*Identify Collinearities and outliers;

```
proc reg data = OC_initial;
model Residual = Hours_win Hours_cenAC Hours_cenHeat /
VIF COLLINOINT;
output out = ttest_OC /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;
```

*for collinearity variance inflation>10, variance proportion >0.5 and condition indices>3.6;

*no collinearity detected. Remove outliers;

```
data OC_MLR1;
set ttest_OC;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;
```

```

*using this clean data set (no outliers) create the MLR model;
proc reg data = OC_MLR1;
model residual = Hours_win Hours_cenAC Hours_cenHeat /selection
=stepwise;
ods output selparmest = selparmest_OC selectionssummary =
selectionsummary_OC; *generates tables with regression model selection
summary and parameter estimates;
run;
quit;

*None of the activity variables met the P = 0.1500 significance level
for entry into the model;

*****

Paired t-test to evaluate whether pairs of measured and modeled
(initial model) indoor OC are significantly different at 95% confidence
level;

proc ttest data=OC_MLR1;
paired In_Measured*APP_initial;
run;

*chi-square to evaluate whether measured and modeled OC have the same
underlying distribution;
data OC_meas;
set OC_MLR1;
drop APP_Initial;
category = "Measured";
Indoor = In_Measured;
drop In_Measured;
run;

data OC_mod;
set OC_MLR1;
drop In_Measured;
category = "Modeled";
Indoor = APP_Initial;
drop APP_Initial;
run;

data OC_meas_mod;
set OC_meas OC_mod;
run;

proc npar1way data = OC_meas_mod wilcoxon;
class category;
var indoor;
run;

*****

```


Paired t-test to evaluate whether pairs of measured indoor and measured outdoor OC are significantly different at 95% confidence level;

```

proc ttest data= OC_MLR1;
paired In_Measured*Out_measured;
run;

*chi-square test measured outdoor vs measured indoor;
data OC_in;
set OC_MLR1;
drop Out_measured;
category = "Indoor";
Indoor = In_Measured;
drop In_Measured;
run;

data OC_out;
set OC_MLR1;
drop In_Measured;
category = "Outdoor";
Indoor = Out_measured;
drop Out_measured;
run;

data OC_in_out;
set OC_in OC_out;
run;

proc npar1way data = OC_in_out wilcoxon;
  class category;
  var indoor;
run;

*****
*****
Robust regression to estimate OC mass measured indoors that can be
attributed to outdoor sources (i.e. reduce influence of indoor
sources);

proc robustreg data = OC_MLR1
method=LTS FWLS;
model In_measured = Out_measured;
run;

*add column with robust regression estimates of indoor OC;
data OC_RobReg;
set OC_MLR1;
Rob_Reg = Out_Measured*0.53; *Population average F value (slope of
robust regression model;
run;

*chi-square test to evaluate whether modeled and robust-regression
estimated indoor OC of outdoor origin have the same underlying
distributions;

```

```

data OC_mod;
set OC_RobReg;
drop Rob_reg;
category = "Modeled";
Indoor = APP_Initial;
drop APP_Initial;
run;

data OC_rr;
set OC_RobReg;
drop APP_Initial;
category = "RobReg";
Indoor = Rob_reg;
drop Rob_reg;
run;

data OC_mod_rr;
set OC_rr OC_mod;
run;

proc npar1way data = OC_mod_rr wilcoxon;
  class category;
  var indoor;
run;

*****
*****
Sensitivity Analyses: alternative size distributions

*Calculate indoor concentrations of OC with "initial" mass balance
model
assuming alternative size distribution 1:
1. 40% mass in fine mode with diameter 0.08, kdep = 0.05 h-1
2. 60% mass in broad accumulation mode with diameter 0.4, kdep = 0.07 h-1
3. P = 0.8;

data OC_initial_ALT1;
set OC_MLR1;
APP_initial =
(0.4*Out_measured*((AER*0.8)/(AER+0.05)))+(0.6*Out_measured*((AER*0.8)/(
(AER+0.07))));
residual = APP_initial - In_measured;
run;

*Calculate indoor concentrations of OC with "initial" mass balance
model
assuming alternative size distribution 2:
1. 40% mass in fine mode with diameter 0.08, kdep = 0.05 h-1
2. 30% mass in condensation mode with diameter 0.2, kdep = 0.05 h-1
3. 30% mass in droplet mode with diameter 0.7, kdep = 0.13 h-1;
4. P = 0.8;

```

```

data OC_initial_ALT2;
set OC_MLR1;
APP_initial =
(0.4*Out_measured*((AER*0.8)/(AER+0.05)))+(0.30*Out_measured*((AER*0.8)
/(AER+0.05)))+(0.30*Out_measured*((AER*0.8)/(AER+0.13)));
residual = APP_initial - In_measured;
run;

*****
*****
Calculation of indoor OC accounting for human activities and assuming P
= 0.9 for fine-mode PM;

data APP_Activities_and_HighP;
set OC_MLR1;
if Hours_win = 0 and Hours_cenAC = 0 and Hours_cenHeat = 0 then APP_all
=
(0.4*Out_measured*((AER*0.9)/(AER+0.05)))+(0.12*Out_measured*((AER*0.8)
/(AER+0.05)))+(0.48*Out_measured*((AER*0.8)/(AER+0.13)));
if hours_win>0 and hours_win>hours_cenAC and hours_win>hours_cenHeat
then APP_all =
(0.4*Out_measured*((AER)/(AER+0.05)))+(0.12*Out_measured*((AER)/(AER+0.
05)))+(0.48*Out_measured*((AER)/(AER+0.13)));
if hours_cenAC>0 and hours_win<hours_cenAC and
hours_cenAC>hours_cenHeat then APP_all =
((0.4*Out_measured*((AER*0.9)/(AER+0.05)))*0.90)+((0.12*Out_measured*((
AER*0.8)/(AER+0.05)))*0.90)+((0.48*Out_measured*((AER*0.8)/(AER+0.13))
)*0.65);
if hours_cenHeat>0 and hours_cenHeat>hours_cenAC and
hours_cenAC<hours_cenHeat then APP_all =
(0.4*Out_measured*((AER*0.9)/(AER+0.05)))*0.90+(0.12*Out_measured*((AER
*0.8)/(AER+0.05)))*0.90+((0.48*Out_measured*((AER*0.8)/(AER+0.13)))*0.6
5);
run;

*****
*****
Multiple linear regression analysis to evaluate the contribution of
physical losses and phase changes of OC to variance in measured indoor
OC;

*Import data sets containing PAH concentrations;

*particle phase PAHs;
PROC IMPORT OUT= particle
DATAFILE= "G:\Chapter 3\Pooled 5 7 ring PAHs.xlsx"
DBMS=EXCEL REPLACE;
SHEET="Particle-Phase";
GETNAMES=YES;
MIXED=YES;
USEDATE=YES;
SCANTIME=YES;

RUN;

```

```

*sum the particle phase concentrations;
data particle2;
set particle;
P_sum = p_BFLTs + p_BeP + p_BaP + p_PER + p_IP + p_DBA + p_BghiP +
p_COR;
keep P_sum HomeID Category;
run;

*isolate the indoor particle concentration measurements;
data particle_in;
set particle2;
if Category NE "indoor" then delete;
P_sum_in = P_sum;
drop P_sum category;
run;

*isolate outdoor particle concentration measurements;
data particle_out;
set particle2;
if Category NE "outdoor" then delete;
P_sum_out = P_sum;
drop P_sum category;
run;

*data set with gas phase PAH concentrations;
PROC IMPORT OUT= gas
            DATAFILE= "G:\Chapter 3\Pooled 5 7 ring PAHs.xlsx"
            DBMS=EXCEL REPLACE;
            SHEET="Gas-Phase";
            GETNAMES=YES;
            MIXED=YES;
            USEDATE=YES;
            SCANTIME=YES;

RUN;

*sum the gas phase concentrations;
data gas2;
set gas;
G_sum = g_BFLTs + g_BeP + g_BaP + g_PER + g_IP + g_DBA + g_BghiP +
g_COR;
keep G_sum HomeID Category;
run;

*isolate the indoor gas concentration measurements;
data gas_in;
set gas2;
if Category NE "indoor" then delete;
G_sum_in = G_sum;
drop G_sum category;
run;

*isolate the indoor gas concentration measurements;
data gas_out;
set gas2;
if Category NE "outdoor" then delete;
G_sum_out = G_sum;

```

```

drop G_sum category;
run;

*merge PAH data sets;
proc sort data = particle_in;by HomeID; run;
proc sort data = particle_out;by HomeID; run;
proc sort data = gas_in;by HomeID; run;
proc sort data = gas_out;by HomeID; run;

data PAH;
merge particle_in particle_out gas_in gas_out;
by HomeID;
run;

*Kp = (F/TSP)/A where F = concentration of PAH in particle phase
(ng/m3), A =
PAH gas phase concentration (ng/3) and TSP = total suspended PM2.5
(ug/m3);

*Import data set with total indoor PM2.5 concentrations;
PROC IMPORT OUT= TSP_in
            DATAFILE= " G:\Chapter 3\RIOPA Indoor Mass Con.xls"
            DBMS=EXCEL REPLACE;
            SHEET="Sheet1";
            GETNAMES=YES;
            MIXED=YES;
            USEDATE=YES;
            SCANTIME=YES;

RUN;

*remove invalid and un-needed data and re-name variables to match other
data sets;
data TSP_in2;
set TSP_in;
if Validation = "(Inv)" then delete;
if Validation = "(INV)" then delete;
HomeID = Home_ID;
In_PM25 = Mass_Concentration_ug_m3_;
drop Validation Comments Sample_ID Home_ID Mass_Concentration_ug_m3_
category;
run;

*Import data set with total outdoor PM2.5 concentrations;
PROC IMPORT OUT= TSP_out
            DATAFILE= " G:\Chapter 3\RIOPA Outdoor Mass Con.xls"
            DBMS=EXCEL REPLACE;
            SHEET="Sheet1";
            GETNAMES=YES;
            MIXED=YES;
            USEDATE=YES;
            SCANTIME=YES;

RUN;

*remove invalid and un-needed data and re-name variables to match other
data sets;

```

```

data TSP_out2;
set TSP_out;
if Validation = "(Inv)" then delete;
if Validation = "(INV)" then delete;
HomeID = Home_ID;
Out_PM25 = Mass_Concentration_ug_m3_;
drop Validation Comments Sample_ID Home_ID Mass_Concentration_ug_m3_
category;
run;

*merge TSP data sets with PAH data set;
proc sort data = TSP_in2;by HomeID;run;
proc sort data = TSP_out2;by HomeID;run;

data Kp;
merge PAH TSP_in2 TSP_out2;
by HomeID;
run;

*calculate Kp indoors and outdoors for PAHs and change in Kp with
indoor transport;
data Kp2;
set Kp;
Kp_in = (P_sum_in/In_PM25)/G_sum_in;
Kp_out = (P_sum_out/Out_PM25)/G_sum_out;
delta_Kp = Kp_in - Kp_out;
run;

*Merge the Kp data set with the data set containing modeled and
measured indoor OC;
proc sort data = Kp2; by HomeID;run;
proc sort data = OC_MLR1; by HomeID;run;

data OC_Kp;
merge OC_MLR1 Kp2;
by HomeID;
run;

**Now perform multiple linear regression analysis to determine if
modeled indoor OC (physical losses)and change in Kp explain the
variance in measured indoor OC;

*Identify Collinearities and outliers;
proc reg data = OC_Kp;
model In_measured = APP_initial delta_Kp/
VIF COLLINOINT;
output out = ttest_Kp /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;

*for collinearity variance inflation>10, variance proportion >0.5 and
condition indices>3.6;
*no collinearity problems detected. Remove outliers;

```

```
data OC_Kp2;
set ttest_Kp;
if student >= 2 or student <= -2 then delete;
drop student rstudent;
run;

*using this clean data set (no outliers) create the MLR model;
proc reg data = OC_Kp2;
model In_measured = APP_initial Delta_Kp / selection = stepwise;
ods output selparmest = selparmest_GPpart selectionsummary =
selectionsummary_GPpart; *creates tables with selection summary and
parameter estimates;
run;
quit;
```

Appendix D: Supporting Information for Chapter 5

Appendix D1 Volatility Distribution Parameters

	$C^*_{min} (25^\circ\text{C})$	$C^*_{max} (25^\circ\text{C})$	$C_{i,tot}$
Total OA, $\Delta H_{vap} = 50 \text{ kJ/mol}$	10^{-2}	10^3	$1.87 + 44\exp[1(\log(C^*)-3)]$
Total OA, $\Delta H_{vap} = 100 \text{ kJ/mol}$	10^{-6}	10^3	$1.65 + 19\exp[1(\log(C^*)-3)]$
HOA	10^{-6}	10^3	$0.78 + 23[0.7(\log(C^*)-3)]$
LV-OOA	10^{-7}	10^3	$0 + 0.135[-0.37(\log(C^*)-3)]$
SV-OOA	10^{-5}	10^3	$0.7 + 7 [0.3(\log(C^*)-3)]$
OOA	10^{-6}	10^3	$1.94 + 5.5[0.8(\log(C^*)-3)]$
“Other” OOA	10^{-6}	10^3	$1.65 + 19[1(\log(C^*)-3)]$

Ambient OA volatility basis set parameters¹ for 25°C and an OA loading of 17 $\mu\text{g}/\text{m}^3$ used to calculate gas-particle partitioning indoors and outdoors for RIOPA homes. C^*_{min} and C^*_{max} are the lower and upper bounds of the range of saturation vapor pressures considered in the calculation of each volatility basis set (VBS). $C_{i,tot}$ describes the relationship between total OM (gas and particle phases) and C^* . Parameters for total OA were used for "other" OA. The distribution of C^* bins for the temperature measured outside of each home was calculated assuming that the temperature-dependence of C^* can be described by the Clausius-Clapeyron equation. Adapted with permission from Cappa and Jimenez.¹

Appendix D2
Aerosol Mass Spectrometer Factor Analysis Organic Aerosol Component Mass Fractions

	HOA	LV-OOA	SV-OOA	OOA	Other OA
Houston (summer)^{a,b}	0.117	-	-	0.450	0.433
Riverside (summer)^{a,c}	0.136	0.318	0.477	-	0.068
NYC (summer)^{a,d}	0.186	0.508	0.305	-	-
NYC (winter)^{a,e}	0.458	-	-	0.542	-

^a Jimenez et al.²

^b Tanaka et al.³

^c DeCarlo et al.;⁴ Docherty et al.;⁵ Cubison et al.⁶

^d Drewnick et al.^{7,8}

^e Weimer et al.⁹

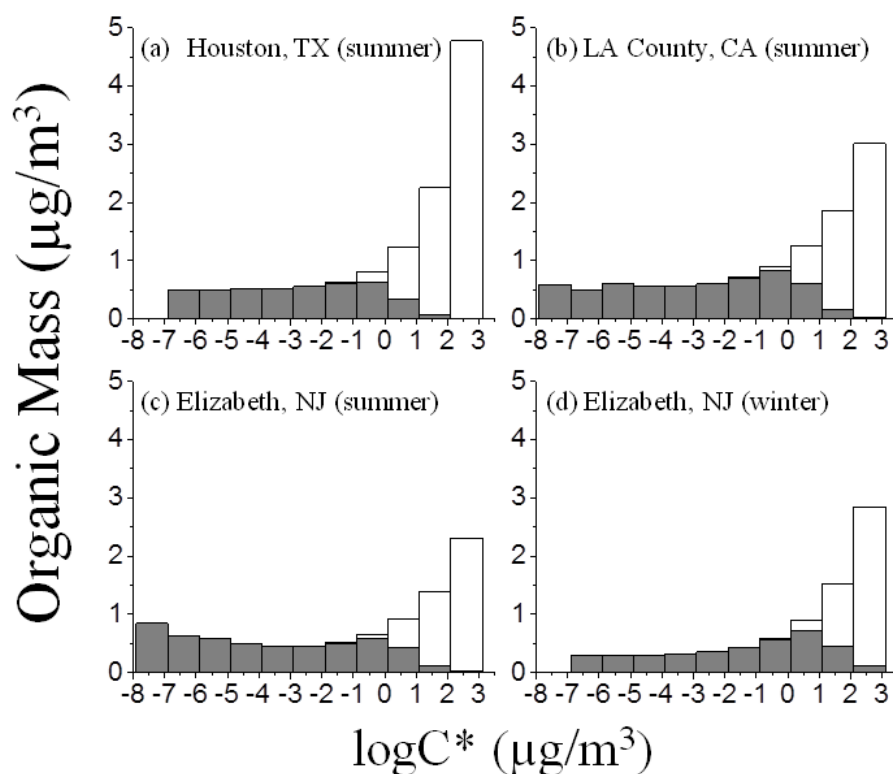
Measured mass fractions of AMS factor-analysis organic aerosol (OA) components. Measured OA component mass fractions for Riverside, New York City (NYC), and Houston were used to estimate component mass fractions for LA County, Elizabeth, and Houston RIOPA homes, respectively, in analyses that made use of OA components.

Appendix D3
Distributions of the Change in Ambient Organic Aerosol Associated with Shifts in Gas-Particle Partitioning with Outdoor-to-Indoor Transport.

	5 th Percentile	25 th Percentile	Median	75 th Percentile	95 th Percentile
$\Delta H_{vap} = 100$ kJ/mol					
ΔC_{OA} ($\mu\text{g}/\text{m}^3$)	-1.81	-0.51	-0.08	1.31	2.19
$C_{OA,in,amb}/C_{OA,out}$	0.72	0.89	0.98	1.06	1.34
$\Delta H_{vap} = 50$ kJ/mol					
ΔC_{OA} ($\mu\text{g}/\text{m}^3$)	-2.11	-0.41	0.25	0.78	1.86
$C_{OA,in,amb}/C_{OA,out}$	0.70	0.93	1.08	1.24	1.96
OA Components					
ΔC_{OA} ($\mu\text{g}/\text{m}^3$)	-1.48	-0.26	0.03	0.24	0.57
$C_{OA,in,amb}/C_{OA,out}$	0.84	0.95	1.01	1.08	1.37

Summary statistics for the distributions of the change in the ambient organic aerosol associated with shifts in gas-particle partitioning with outdoor-to-indoor transport calculated in the main analyses (ΔC_{OA}) and the same value reported as a ratio of calculated indoor ambient OA concentrations (accounting for shifts in partitioning) to the measured outdoor ambient OA concentrations ($C_{OA,in,amb}/C_{OA,out}$). These values are shown for the various assumptions made regarding the volatility distribution of ambient OA: (1) OA volatility was represented with a single volatility distribution and an enthalpy of vaporization (ΔH_{vap}) of 100 kJ/mol, (2) OA volatility was represented with a single volatility distribution and an ΔH_{vap} of 50 kJ/mol, and (3) OA volatility distributions were represented by an average of the of the OA component volatility distributions weighted by the mass fractions of the OA components in each of those seasons and locations assuming $\Delta H_{vap} = 100$ kJ/mol for all components.

Appendix D4
RIOPA-Region- and Season-Specific Aggregate Volatility Distributions Generated
for the Chemically-Resolved, Organic Aerosol Component Calculations



Region- and season-specific volatility distributions calculated assuming that ambient OA is a mixture of AMS factor-analysis components with the mass fractions shown in Appendix D2. Volatility distributions were adjusted to the average conditions measured outside the RIOPA homes in each region/season: (a) temperature = 26.1°C and OA loading = $4.28 \mu\text{g}/\text{m}^3$, (b) temperature = 21.2°C and OA loading = $5.75 \mu\text{g}/\text{m}^3$, (c) temperature = 20.6°C and OA loading = $5.17 \mu\text{g}/\text{m}^3$, and (d) temperature = 7.8°C and OA loading = $3.85 \mu\text{g}/\text{m}^3$. Volatility distributions for the OA components are provided in Figure 1 of the main text.

Appendix D5

SAS Code for Chapter 5 Calculations: Indoor and outdoor Gas-particle Partitioning for RIOPA Study Homes

*Change to directory on accompanying flash drive with data sets containing measured indoor and outdoor OC and temperature. Note: the drive label might be different for each computer;

```
libname Ch5 'G:\Chapter 5';
```

```
*****
For total ambient organic aerosol (OA) assuming delta_Hvap = 50 kJ/mol
*****
;
```

Calculate total (gas+particle) organic concentration in each bin using relationship between C and Ctot from Cappa and Jimenez 2010;

```
data Ctot_out_50;
set Ch5.OM_Temp_out; *data set with all required data (on jump drive);
C1_tot = 1.87+44*exp(1*(log10(10**-2)-3));
C2_tot = 1.87+44*exp(1*(log10(10**-1)-3));
C3_tot = 1.87+44*exp(1*(log10(10**0)-3));
C4_tot = 1.87+44*exp(1*(log10(10**1)-3));
C5_tot = 1.87+44*exp(1*(log10(10**2)-3));
C6_tot = 1.87+44*exp(1*(log10(10**3)-3));
run;
```

Calculate C bins for temperatures measured outside RIOPA homes using Clausius Clapyeron Eqn. and C* for reference temperature of 298.15 K from Cappa and Jimenez (2010);

```
data CT_out_50;
set Ctot_out_50;
C1_T_out = 10**-2*(298.15/Temp_out)*exp((( -50000/8.314)*(1/Temp_out)-(1/298.15))));
C2_T_out = 10**-1*(298.15/Temp_out)*exp((( -50000/8.314)*(1/Temp_out)-(1/298.15))));
C3_T_out = 10**0*(298.15/Temp_out)*exp((( -50000/8.314)*(1/Temp_out)-(1/298.15))));
C4_T_out = 10**1*(298.15/Temp_out)*exp((( -50000/8.314)*(1/Temp_out)-(1/298.15))));
C5_T_out = 10**2*(298.15/Temp_out)*exp((( -50000/8.314)*(1/Temp_out)-(1/298.15))));
C6_T_out = 10**3*(298.15/Temp_out)*exp((( -50000/8.314)*(1/Temp_out)-(1/298.15))));
run;
```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```
data Part_Hvap50_Out;
set CT_out_50;
C1_part_out = (1+((C1_T_out)/OM_out))**-1;
C2_part_out = (1+((C2_T_out)/OM_out))**-1;
C3_part_out = (1+((C3_T_out)/OM_out))**-1;
```

```

C4_part_out = (1+((C4_T_out)/OM_out))** -1;
C5_part_out = (1+((C5_T_out)/OM_out))** -1;
C6_part_out = (1+((C6_T_out)/OM_out))** -1;
Part_tot_out =
(C1_part_out*C1_tot+C2_part_out*C2_tot+C3_part_out*C3_tot+C4_part_out*C
4_tot+C5_part_out*C5_tot+C6_part_out*C6_tot)/(C1_tot+C2_tot+C3_tot+C4_t
ot+C5_tot+C6_tot);
run;

*Calculate C* bins for temperatures measured inside RIOPA homes using
Clausius Clapyeron Eqn.;
data CT_in_50;
set Ch5.OM_Temp_in;
C1_T_in = 10** -2*(298.15/Temp_in)*exp((( -50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C2_T_in = 10** -1*(298.15/Temp_in)*exp((( -50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C3_T_in = 10** 0*(298.15/Temp_in)*exp((( -50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C4_T_in = 10** 1*(298.15/Temp_in)*exp((( -50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C5_T_in = 10** 2*(298.15/Temp_in)*exp((( -50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C6_T_in = 10** 3*(298.15/Temp_in)*exp((( -50000/8.314)*(1/Temp_in)-
(1/298.15)))));
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data Part_Hvap50_in;
set CT_in_50;
C1_part_in = (1+((C1_T_in)/OM_in))** -1;
C2_part_in = (1+((C2_T_in)/OM_in))** -1;
C3_part_in = (1+((C3_T_in)/OM_in))** -1;
C4_part_in = (1+((C4_T_in)/OM_in))** -1;
C5_part_in = (1+((C5_T_in)/OM_in))** -1;
C6_part_in = (1+((C6_T_in)/OM_in))** -1;
run;

*merge outdoor and indoor partitioning data sets;
proc sort data = Part_Hvap50_out; by HomeID; run;
proc sort data = Part_Hvap50_in; by HomeID; run;

data Hvap50_Part;
merge Part_Hvap50_out Part_Hvap50_in;
by HomeID;
run;

*calculate total partitioning indoors;
data Hvap50_Part2;
set Hvap50_Part;
Part_tot_in =
(C1_part_in*C1_tot+C2_part_in*C2_tot+C3_part_in*C3_tot+C4_part_in*C4_to
t+C5_part_in*C5_tot+C6_part_in*C6_tot)/(C1_tot+C2_tot+C3_tot+C4_tot+C5_
tot+C6_tot);
run;

```

```

*Create data set with just HomeID, indoor and outdoor OM (measured),
indoor and outdoor temp, Total indoor and outdoor partitioning;
data Hvap50_Part3;
set Hvap50_Part2;
keep HomeID Temp_out Temp_in OM_out OM_in Part_tot_out Part_tot_in;
run;

*Calculate total OM (gas + particle phase) outdoors (Ctot_out) by
dividing measured outdoor OM by outdoor partitioning
Calculate OM in the particle phase indoors (OA_in_calc) by multiplying
Ctot_out by indoor partitioning
Calculate change in OA concentration due to partitioning shifts by
subtracting measured outdoor OM in particle phase from calculated
indoor OM in particle phase;
data Hvap50_Part4;
set Hvap50_Part3;
Ctot_out = OM_out/Part_tot_out;
OA_in_calc = Ctot_out*Part_tot_in;
delta_COA = OA_in_calc - OM_out;
run;

*This data set exported to library "Ch5" with the name "Hvap50.xls"

*****
For total ambient organic aerosol (OA) assuming delta_Hvap = 100 kJ/mol
*****

;

libname Ch5 'G:\Chapter 5';

*Calculate total (gas+particle) organic concentration in each bin using
relationship between C* and Ctot from Cappa and Jimenez 2010;
data Ctot_out_100;
set Ch5.OM_Temp_out;
C1_tot = 1.65+19*exp(1*(log10(10**-6)-3));
C2_tot = 1.65+19*exp(1*(log10(10**-5)-3));
C3_tot = 1.65+19*exp(1*(log10(10**-4)-3));
C4_tot = 1.65+19*exp(1*(log10(10**-3)-3));
C5_tot = 1.65+19*exp(1*(log10(10**-2)-3));
C6_tot = 1.65+19*exp(1*(log10(10**-1)-3));
C7_tot = 1.65+19*exp(1*(log10(10**0)-3));
C8_tot = 1.65+19*exp(1*(log10(10**1)-3));
C9_tot = 1.65+19*exp(1*(log10(10**2)-3));
C10_tot = 1.65+19*exp(1*(log10(10**3)-3));
run;

*Calculate C* bins for temperatures measured outside RIOPA homes using
Clausius Clapyeron Eqn.and C* for reference temperature of 298.15 K
from Cappa and Jimenez (2010);

```

```

data CT_out_100;
set Ctot_out_100;
C1_T_out = 10**(-6*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C2_T_out = 10**(-5*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C3_T_out = 10**(-4*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C4_T_out = 10**(-3*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C5_T_out = 10**(-2*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C6_T_out = 10**(-1*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C7_T_out = 10**(-0*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C8_T_out = 10**1*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C9_T_out = 10**2*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
C10_T_out = 10**3*(298.15/Temp_out)*exp((-100000/8.314)*((1/Temp_out)-(1/298.15))));
run;

```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```

data Part_Hvap100_Out;
set CT_out_100;
C1_part_out = (1+((C1_T_out)/OM_out))**-1;
C2_part_out = (1+((C2_T_out)/OM_out))**-1;
C3_part_out = (1+((C3_T_out)/OM_out))**-1;
C4_part_out = (1+((C4_T_out)/OM_out))**-1;
C5_part_out = (1+((C5_T_out)/OM_out))**-1;
C6_part_out = (1+((C6_T_out)/OM_out))**-1;
C7_part_out = (1+((C7_T_out)/OM_out))**-1;
C8_part_out = (1+((C8_T_out)/OM_out))**-1;
C9_part_out = (1+((C9_T_out)/OM_out))**-1;
C10_part_out = (1+((C10_T_out)/OM_out))**-1;
Part_tot_out =
(C1_part_out*C1_tot+C2_part_out*C2_tot+C3_part_out*C3_tot+C4_part_out*C
4_tot+C5_part_out*C5_tot+C6_part_out*C6_tot
+C7_part_out*C7_tot+C8_part_out*C8_tot+C9_part_out*C9_tot+C10_part_out*
C10_tot)/
(C1_tot+C2_tot+C3_tot+C4_tot+C5_tot+C6_tot+C7_tot+C8_tot+C9_tot+C10_tot
);
run;

```

*Repeat calculations for indoor environment;

Calculate C bins for temperatures measured inside RIOPA homes using Clausius Clapyeron Eqn.and C* for reference temperature of 298.15 K from Cappa and Jimenez (2010);

```

data CT_in_100;
set Ch5.OM_Temp_in;

```

```

C1_T_in = 10**-6*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C2_T_in = 10**-5*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C3_T_in = 10**-4*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C4_T_in = 10**-3*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C5_T_in = 10**-2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C6_T_in = 10**-1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C7_T_in = 10**0*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C8_T_in = 10**1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C9_T_in = 10**2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C10_T_in = 10**3*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez 2010;
data Part_Hvap100_in;
set CT_in_100;
C1_part_in = (1+((C1_T_in)/OM_in))**-1;
C2_part_in = (1+((C2_T_in)/OM_in))**-1;
C3_part_in = (1+((C3_T_in)/OM_in))**-1;
C4_part_in = (1+((C4_T_in)/OM_in))**-1;
C5_part_in = (1+((C5_T_in)/OM_in))**-1;
C6_part_in = (1+((C6_T_in)/OM_in))**-1;
C7_part_in = (1+((C7_T_in)/OM_in))**-1;
C8_part_in = (1+((C8_T_in)/OM_in))**-1;
C9_part_in = (1+((C9_T_in)/OM_in))**-1;
C10_part_in = (1+((C10_T_in)/OM_in))**-1;
run;

*merge outdoor and indoor partitioning data sets;
proc sort data = Part_Hvap100_out; by HomeID; run;
proc sort data = Part_Hvap100_in; by HomeID; run;

data Hvap100_Part;
merge Part_Hvap100_out Part_Hvap100_in;
by HomeID;
run;

*calculate total partitioning indoors;
data Hvap100_Part2;
set Hvap100_Part;
Part_tot_in =
(C1_part_in*C1_tot+C2_part_in*C2_tot+C3_part_in*C3_tot+C4_part_in*C4_tot+C5_part_in*C5_tot+C6_part_in*C6_tot+C7_part_in*C7_tot+C8_part_in*C8_tot+C9_part_in*C9_tot+C10_part_in*C10_tot)/
(C1_tot+C2_tot+C3_tot+C4_tot+C5_tot+C6_tot+C7_tot+C8_tot+C9_tot+C10_tot);

```



```

run;

*Create data set with just HomeID, indoor and outdoor OM (measured),
indoor and outdoor temp, Total indoor and outdoor partitioning;
data Hvap100_Part3;
set Hvap100_Part2;
keep HomeID Temp_out Temp_in OM_out OM_in Part_tot_out Part_tot_in;
run;

*Calculate total OM (gas + particle phase) outdoors (Ctot_out) by
dividing measured outdoor OM by outdoor partitioning
Calculate OM in the particle phase indoors (OA_in_calc) by multiplying
Ctot_out by indoor partitioning
Calculate change in OA concentration due to partitioning shifts by
subtracting measured outdoor OM in particle phase from calculated
indoor OM in particle phase;
data Hvap100_Part4;
set Hvap100_Part3;
Ctot_out = OM_out/Part_tot_out;
OA_in_calc = Ctot_out*Part_tot_in;
delta_COA = OA_in_calc - OM_out;
run;

*This data set exported to library "Ch5" with the name "Hvap100.xls"

*****
Looking at impact of indoor-outdoor temperature differences and indoor-
outdoor OA loading differences separately
*****
;

*****
Temperature Only
*****;

*Here we re-partition the ambient organics based on the change in temp
between the indoor and outdoor environments
not accounting for changes in OA concentration;

*****Assuming delta_Hvap = 50 kJ/mol*****;

*Calculate total (gas+particle) organic concentration in each bin using
relationship between C* and Ctot from Cappa and Jimenez 2010;
data Ctot_out_50;
set Ch5.OM_Temp_out;
C1_tot = 1.87+44*exp(1*(log10(10**-2)-3));
C2_tot = 1.87+44*exp(1*(log10(10**-1)-3));
C3_tot = 1.87+44*exp(1*(log10(10**0)-3));
C4_tot = 1.87+44*exp(1*(log10(10**1)-3));
C5_tot = 1.87+44*exp(1*(log10(10**2)-3));
C6_tot = 1.87+44*exp(1*(log10(10**3)-3));
run;

```

```

*Calculate how C*(Tref = 298.15 K) shifts based on outdoor temperature
using Clausius Clapyeron Eqn.;
data CT_50_out;
set Ctot_out_50;
C1_T_out = 10**-2*(298.15/Temp_out)*exp(((-50000/8.314)*(1/Temp_out)-
(1/298.15)))));
C2_T_out = 10**-1*(298.15/Temp_out)*exp(((-50000/8.314)*(1/Temp_out)-
(1/298.15)))));
C3_T_out = 10**0*(298.15/Temp_out)*exp(((-50000/8.314)*(1/Temp_out)-
(1/298.15)))));
C4_T_out = 10**1*(298.15/Temp_out)*exp(((-50000/8.314)*(1/Temp_out)-
(1/298.15)))));
C5_T_out = 10**2*(298.15/Temp_out)*exp(((-50000/8.314)*(1/Temp_out)-
(1/298.15)))));
C6_T_out = 10**3*(298.15/Temp_out)*exp(((-50000/8.314)*(1/Temp_out)-
(1/298.15)))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data Part_Hvap50_Out;
set CT_50_out;
C1_part_out = (1+((C1_T_out)/OM_out))**-1;
C2_part_out = (1+((C2_T_out)/OM_out))**-1;
C3_part_out = (1+((C3_T_out)/OM_out))**-1;
C4_part_out = (1+((C4_T_out)/OM_out))**-1;
C5_part_out = (1+((C5_T_out)/OM_out))**-1;
C6_part_out = (1+((C6_T_out)/OM_out))**-1;
Part_tot_out =
(C1_part_out*C1_tot+C2_part_out*C2_tot+C3_part_out*C3_tot+C4_part_out*C
4_tot+C5_part_out*C5_tot+C6_part_out*C6_tot)/(C1_tot+C2_tot+C3_tot+C4_t
ot+C5_tot+C6_tot);
run;

*Calculate how C*(Tref = 298.15 K) shifts based on indoor temperature
using Clausius Clapyeron Eqn.;
data CT_50_in;
set Ch5.OM_Temp_in;
C1_T_in = 10**-2*(298.15/Temp_in)*exp(((-50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C2_T_in = 10**-1*(298.15/Temp_in)*exp(((-50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C3_T_in = 10**0*(298.15/Temp_in)*exp(((-50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C4_T_in = 10**1*(298.15/Temp_in)*exp(((-50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C5_T_in = 10**2*(298.15/Temp_in)*exp(((-50000/8.314)*(1/Temp_in)-
(1/298.15)))));
C6_T_in = 10**3*(298.15/Temp_in)*exp(((-50000/8.314)*(1/Temp_in)-
(1/298.15)))));
run;

*merge indoor and outdoor data sets;
proc sort data = Part_Hvap50_Out; by HomeID; run;
proc sort data = CT_50_in; by HomeID; run;

```

```

data Part_Hvap50_in_Temp1;
merge Part_Hvap50_Out CT_50_in;
by HomeID;
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010 without accounting for change in COA;
data Part_Hvap50_in_Temp2;
set Part_Hvap50_in_Temp1;
C1_part_in = (1+((C1_T_in)/OM_out))**-1;
C2_part_in = (1+((C2_T_in)/OM_out))**-1;
C3_part_in = (1+((C3_T_in)/OM_out))**-1;
C4_part_in = (1+((C4_T_in)/OM_out))**-1;
C5_part_in = (1+((C5_T_in)/OM_out))**-1;
C6_part_in = (1+((C6_T_in)/OM_out))**-1;
Part_tot_in =
(C1_part_in*C1_tot+C2_part_in*C2_tot+C3_part_in*C3_tot+C4_part_in*C4_tot+
C5_part_in*C5_tot+C6_part_in*C6_tot)/(C1_tot+C2_tot+C3_tot+C4_tot+C5_tot+
C6_tot);
run;

*keep important data;
data Part_Hvap50_in_Temp3;
set Part_Hvap50_in_Temp2;
keep HomeID OM_out Temp_out Part_tot_out OM_in Temp_in Part_tot_in;
run;

*calculate total organic concentration outdoors and then indoor COA;
data Hvp50_Temp;
set Part_Hvap50_in_Temp3;
Ctot = OM_out/Part_tot_out;
COA_out = Part_tot_out*Ctot;
COA_in_temp = Ctot*Part_tot_in;
run;

*this data set was exported to library "Ch5" as "Temperature or OA
Loading ONLY.xls", sheet "Hvap50 Temp";

*****Assuming delta_Hvap = 100 kJ/mol*****;

*Calculate total (gas+particle) organic concentration in each bin using
relationship
between C* and Ctot from Cappa and Jimenez 2010;
data Ctot_out_100;
set Ch5.OM_Temp_out;
C1_tot = 1.65+19*exp(1*(log10(10**-6)-3));
C2_tot = 1.65+19*exp(1*(log10(10**-5)-3));
C3_tot = 1.65+19*exp(1*(log10(10**-4)-3));
C4_tot = 1.65+19*exp(1*(log10(10**-3)-3));
C5_tot = 1.65+19*exp(1*(log10(10**-2)-3));
C6_tot = 1.65+19*exp(1*(log10(10**-1)-3));
C7_tot = 1.65+19*exp(1*(log10(10**0)-3));
C8_tot = 1.65+19*exp(1*(log10(10**1)-3));
C9_tot = 1.65+19*exp(1*(log10(10**2)-3));

```

```

C10_tot = 1.65+19*exp(1*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on outdoor temperature
using Clausius Clapyeron Eqn.;
data CT_100_out;
set Ctot_out_100;
C1_T_out = 10**(-6)*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C2_T_out = 10**(-5)*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C3_T_out = 10**(-4)*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C4_T_out = 10**(-3)*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C5_T_out = 10**(-2)*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C6_T_out = 10**(-1)*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C7_T_out = 10**(-0)*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C8_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C9_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
C10_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data Part_Hvap100_Out;
set CT_100_out;
C1_part_out = (1+((C1_T_out)/OM_out))**-1;
C2_part_out = (1+((C2_T_out)/OM_out))**-1;
C3_part_out = (1+((C3_T_out)/OM_out))**-1;
C4_part_out = (1+((C4_T_out)/OM_out))**-1;
C5_part_out = (1+((C5_T_out)/OM_out))**-1;
C6_part_out = (1+((C6_T_out)/OM_out))**-1;
C7_part_out = (1+((C7_T_out)/OM_out))**-1;
C8_part_out = (1+((C8_T_out)/OM_out))**-1;
C9_part_out = (1+((C9_T_out)/OM_out))**-1;
C10_part_out = (1+((C10_T_out)/OM_out))**-1;
Part_tot_out =
(C1_part_out*C1_tot+C2_part_out*C2_tot+C3_part_out*C3_tot+C4_part_out*C
4_tot+C5_part_out*C5_tot+C6_part_out*C6_tot
+C7_part_out*C7_tot+C8_part_out*C8_tot+C9_part_out*C9_tot+C10_part_out*
C10_tot)/
(C1_tot+C2_tot+C3_tot+C4_tot+C5_tot+C6_tot+C7_tot+C8_tot+C9_tot+C10_tot
);
run;

*Calculate how C*(Tref = 298.15 K) shifts based on indoor temperature
using Clausius Clapyeron Eqn.;
data CT_100_in;
set Ch5.OM_Temp_in;;

```

```

C1_T_in = 10**-6*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C2_T_in = 10**-5*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C3_T_in = 10**-4*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C4_T_in = 10**-3*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C5_T_in = 10**-2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C6_T_in = 10**-1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C7_T_in = 10**0*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C8_T_in = 10**1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C9_T_in = 10**2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
C10_T_in = 10**3*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
run;

*merge indoor and outdoor data sets;
proc sort data = Part_Hvap100_Out; by HomeID; run;
proc sort data = CT_100_in; by HomeID; run;

data Part_Hvap100_in_Temp1;
merge Part_Hvap100_Out CT_100_in;
by HomeID;
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010 without accounting for change in COA;
data Part_Hvap100_in_Temp2;
set Part_Hvap100_in_Temp1;
C1_part_in = (1+((C1_T_in)/OM_out))**-1;
C2_part_in = (1+((C2_T_in)/OM_out))**-1;
C3_part_in = (1+((C3_T_in)/OM_out))**-1;
C4_part_in = (1+((C4_T_in)/OM_out))**-1;
C5_part_in = (1+((C5_T_in)/OM_out))**-1;
C6_part_in = (1+((C6_T_in)/OM_out))**-1;
C7_part_in = (1+((C7_T_in)/OM_out))**-1;
C8_part_in = (1+((C8_T_in)/OM_out))**-1;
C9_part_in = (1+((C9_T_in)/OM_out))**-1;
C10_part_in = (1+((C10_T_in)/OM_out))**-1;
Part_tot_in =
(C1_part_in*C1_tot+C2_part_in*C2_tot+C3_part_in*C3_tot+C4_part_in*C4_tot+
C5_part_in*C5_tot+C6_part_in*C6_tot
+C7_part_in*C7_tot+C8_part_in*C8_tot+C9_part_in*C9_tot+C10_part_in*C10_tot)/
(C1_tot+C2_tot+C3_tot+C4_tot+C5_tot+C6_tot+C7_tot+C8_tot+C9_tot+C10_tot);
run;

*keep important data;
data Part_Hvap100_in_Temp3;

```

```

set Part_Hvap100_in_Temp2;
keep HomeID OM_out Temp_out Part_tot_out OM_in Temp_in Part_tot_in;
run;

*calculate total organic concentration outdoors and then calculated
partitioned indoor COA;
data Hvp100_Temp;
set Part_Hvap100_in_Temp3;
Ctot = OM_out/Part_tot_out;
COA_out = Ctot*Part_tot_out;
COA_in_temp = Ctot*Part_tot_in;
run;

*this data set was exported to library "Ch5" as "Temperature or OA
Loading ONLY.xls," sheet "Hvap100 Temp";

*****
  OA Loading changes Only
*****;
*Here we re-partition the organics based on the change of COA between
the indoor and outdoor environments not accounting for temperature
changes;

*****Assuming delta_Hvap = 50 kJ/mol*****;

*merge data set with outdoor partitioning with data set with indoor
COA;
*merge indoor and outdoor data sets;
proc sort data = Part_Hvap50_Out; by HomeID; run;
proc sort data = Ch5.OM_Temp_in; by HomeID; run;

data Hvp50_COA_only1;
merge Part_Hvap50_Out Ch5.OM_Temp_in;
by HomeID;
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010 without accounting for change in temperature;
data Hvp50_COA_only2;
set Hvp50_COA_only1;
C1_part_in = (1+((C1_T_out)/OM_in))**-1;
C2_part_in = (1+((C2_T_out)/OM_in))**-1;
C3_part_in = (1+((C3_T_out)/OM_in))**-1;
C4_part_in = (1+((C4_T_out)/OM_in))**-1;
C5_part_in = (1+((C5_T_out)/OM_in))**-1;
C6_part_in = (1+((C6_T_out)/OM_in))**-1;
Part_tot_in =
(C1_part_in*C1_tot+C2_part_in*C2_tot+C3_part_in*C3_tot+C4_part_in*C4_to
t+C5_part_in*C5_tot+C6_part_in*C6_tot)/(C1_tot+C2_tot+C3_tot+C4_tot+C5_
tot+C6_tot);
run;

```

```

*keep important data;
data Hvap50_COA_only3;
set Hvap50_COA_only2;
keep HomeID OM_out Temp_out Part_tot_out OM_in Temp_in Part_tot_in;
run;

*calculate total organic concentration outdoors and then calculated
partitioned indoor COA;
data Hvap50_COA;
set Hvap50_COA_only3;
Ctot = OM_out/Part_tot_out;
COA_out = Part_tot_out*Ctot;
COA_in_OA = Ctot*Part_tot_in;
run;

*this sheet exported to library "Ch5" to file "Temperature or OA
Loading ONLY.xls," sheet "Hvap50 OA Load";

*****Assuming delta_Hvap = 100 kJ/mol*****;

*merge data set with outdoor partitioning with data set with indoor
COA;
*merge indoor and outdoor data sets;
proc sort data = Part_Hvap100_Out; by HomeID; run;
proc sort data = Ch5.OM_Temp_in; by HomeID; run;

data Hvap100_COA_only1;
merge Part_Hvap100_Out Ch5.OM_Temp_in;
by HomeID;
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010 without accounting for change in Temperature;
data Hvap100_COA_only2;
set Hvap100_COA_only1;
C1_part_in = (1+((C1_T_out)/OM_in))**-1;
C2_part_in = (1+((C2_T_out)/OM_in))**-1;
C3_part_in = (1+((C3_T_out)/OM_in))**-1;
C4_part_in = (1+((C4_T_out)/OM_in))**-1;
C5_part_in = (1+((C5_T_out)/OM_in))**-1;
C6_part_in = (1+((C6_T_out)/OM_in))**-1;
C7_part_in = (1+((C7_T_out)/OM_in))**-1;
C8_part_in = (1+((C8_T_out)/OM_in))**-1;
C9_part_in = (1+((C9_T_out)/OM_in))**-1;
C10_part_in = (1+((C10_T_out)/OM_in))**-1;
Part_tot_in =
(C1_part_in*C1_tot+C2_part_in*C2_tot+C3_part_in*C3_tot+C4_part_in*C4_to
t+C5_part_in*C5_tot+C6_part_in*C6_tot
+C7_part_in*C7_tot+C8_part_in*C8_tot+C9_part_in*C9_tot+C10_part_in*C10_
tot)/

```

```
(C1_tot+C2_tot+C3_tot+C4_tot+C5_tot+C6_tot+C7_tot+C8_tot+C9_tot+C10_tot
);
run;
```

```
*keep important parameters;
data Hvap100_COA_only3;
set Hvap100_COA_only2;
keep HomeID OM_out Temp_out Part_tot_out OM_in Temp_in Part_tot_in;
run;
```

```
*calculate total organic concentration outdoors and then calculated
partitioned indoor COA;
```

```
data Hvap100_COA;
set Hvap100_COA_only3;
Ctot = OM_out/Part_tot_out;
COA_out = Ctot*Part_tot_out;
COA_in_OA = Ctot*Part_tot_in;
run;
```

```
*this sheet exported to library "Ch5", file "Temperature or OA Loading
ONLY.xls," sheet "Hvap100 OA Load";
```

```
*****
Treating ambient OA as a mixture of AMS factor-analysis components.
Note: mass fractions of OA components are season- and region-specific
*****
;
```

```
libname Ch5 'G:\Chapter 5';
```

```
*First, combine data set with temp, OM concentrations with dates to
break up by season;
```

```
*Import data set with dates and HomeID;
```

```
PROC IMPORT OUT= dates
      DATAFILE= "G:\Chapter 5\RIOPA Dates.xls"
      DBMS=EXCEL REPLACE;
      SHEET="RIOPA_Dates";
      GETNAMES=YES;
      MIXED=YES;
      USEDATE=YES;
      SCANTIME=YES;

RUN;
```

```
proc sort data = dates; by HomeID;run;
```

```
data Dates_OM_Temp;
merge Ch5.OM_Temp_out Ch5.OM_Temp_in dates;
by HomeID;
```



```

if OM_out = " " then delete;
if Start_date = "." then delete;
run;

```

*Added "Season" and state in excel. s = summer and refers to the warm season May - Oct. w = winter and refers to the cool season Nov - Apr;

```

PROC IMPORT OUT= Season
DATAFILE= " G:\Chapter 5\Season and State.xls"
DBMS=EXCEL REPLACE;
SHEET="All";
GETNAMES=YES;
MIXED=YES;
USEDATE=YES;
SCANTIME=YES;

RUN;

```

*LA County, CA RIOPA homes using Riverside, CA OA component mass fractions detailed in Jimenez et al. 2009;
 *Note the components are broken up by summer and winter, so I will break up the measured data into warm and cold seasons;

```

*create data set with CA summer dates only;
data CA_summer;
set Season;
if season NE "s" then delete;
if state NE "CA" then delete;
run;

```

```

*****
      HOA
*****

```

calculate total (gas+particle) organic concentration in each HOA bin using relationship between C and Ct_{tot} from Cappa and Jimenez 2010;

```

data CA_summerHOAtot;
set CA_summer;
HOA_C1_tot = 0.78+23*exp(0.7*(log10(10**-6)-3));
HOA_C2_tot = 0.78+23*exp(0.7*(log10(10**-5)-3));
HOA_C3_tot = 0.78+23*exp(0.7*(log10(10**-4)-3));
HOA_C4_tot = 0.78+23*exp(0.7*(log10(10**-3)-3));
HOA_C5_tot = 0.78+23*exp(0.7*(log10(10**-2)-3));
HOA_C6_tot = 0.78+23*exp(0.7*(log10(10**-1)-3));
HOA_C7_tot = 0.78+23*exp(0.7*(log10(10**0)-3));
HOA_C8_tot = 0.78+23*exp(0.7*(log10(10**1)-3));
HOA_C9_tot = 0.78+23*exp(0.7*(log10(10**2)-3));
HOA_C10_tot = 0.78+23*exp(0.7*(log10(10**3)-3));
run;

```

Calculate how C(T_{ref} = 298.15 K) shifts based on temperature using Clausius Clapyeron Eqn.;

```

data CA_summerHOA_Cout;
set CA_summerHOAtot;
HOA_C1_T_out = 10**(-6*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C2_T_out = 10**(-5*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C3_T_out = 10**(-4*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C4_T_out = 10**(-3*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C5_T_out = 10**(-2*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C6_T_out = 10**(-1*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C7_T_out = 10**(-0*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C8_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C9_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C10_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data CA_summerHOA_Part_out;
set CA_summerHOA_Cout;
HOA_C1_part_out = (1+((HOA_C1_T_out)/OM_out))**-1;
HOA_C2_part_out = (1+((HOA_C2_T_out)/OM_out))**-1;
HOA_C3_part_out = (1+((HOA_C3_T_out)/OM_out))**-1;
HOA_C4_part_out = (1+((HOA_C4_T_out)/OM_out))**-1;
HOA_C5_part_out = (1+((HOA_C5_T_out)/OM_out))**-1;
HOA_C6_part_out = (1+((HOA_C6_T_out)/OM_out))**-1;
HOA_C7_part_out = (1+((HOA_C7_T_out)/OM_out))**-1;
HOA_C8_part_out = (1+((HOA_C8_T_out)/OM_out))**-1;
HOA_C9_part_out = (1+((HOA_C9_T_out)/OM_out))**-1;
HOA_C10_part_out = (1+((HOA_C10_T_out)/OM_out))**-1;
HOA_Part_tot_out =
(HOA_C1_part_out*HOA_C1_tot+HOA_C2_part_out*HOA_C2_tot+HOA_C3_part_out*
HOA_C3_tot+HOA_C4_part_out*HOA_C4_tot+HOA_C5_part_out*HOA_C5_tot+HOA_C6
_part_out*HOA_C6_tot
+HOA_C7_part_out*HOA_C7_tot+HOA_C8_part_out*HOA_C8_tot+HOA_C9_part_out*
HOA_C9_tot+HOA_C10_part_out*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature indoors
using Clausius Clapyeron Eqn.;
data CA_summerHOA_Cin;
set CA_summerHOAtot;
HOA_C1_T_in = 10**(-6*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C2_T_in = 10**(-5*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15)))));

```

```

HOA_C3_T_in = 10**-4*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
HOA_C4_T_in = 10**-3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
HOA_C5_T_in = 10**-2*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
HOA_C6_T_in = 10**-1*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
HOA_C7_T_in = 10**0*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15))));
HOA_C8_T_in = 10**1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15))));
HOA_C9_T_in = 10**2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15))));
HOA_C10_T_in = 10**3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data CA_summerHOA_Part_in;
set CA_summerHOA_Cin;
HOA_C1_part_in = (1+((HOA_C1_T_in)/OM_in))**-1;
HOA_C2_part_in = (1+((HOA_C2_T_in)/OM_in))**-1;
HOA_C3_part_in = (1+((HOA_C3_T_in)/OM_in))**-1;
HOA_C4_part_in = (1+((HOA_C4_T_in)/OM_in))**-1;
HOA_C5_part_in = (1+((HOA_C5_T_in)/OM_in))**-1;
HOA_C6_part_in = (1+((HOA_C6_T_in)/OM_in))**-1;
HOA_C7_part_in = (1+((HOA_C7_T_in)/OM_in))**-1;
HOA_C8_part_in = (1+((HOA_C8_T_in)/OM_in))**-1;
HOA_C9_part_in = (1+((HOA_C9_T_in)/OM_in))**-1;
HOA_C10_part_in = (1+((HOA_C10_T_in)/OM_in))**-1;
HOA_Part_tot_in =
(HOA_C1_part_in*HOA_C1_tot+HOA_C2_part_in*HOA_C2_tot+HOA_C3_part_in*HOA
_C3_tot+HOA_C4_part_in*HOA_C4_tot+HOA_C5_part_in*HOA_C5_tot+HOA_C6_part
_in*HOA_C6_tot
+HOA_C7_part_in*HOA_C7_tot+HOA_C8_part_in*HOA_C8_tot+HOA_C9_part_in*HOA
_C9_tot+HOA_C10_part_in*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;

*Merge indoor and outdoor Partitioning data sets and keep only key
data;
proc sort data = CA_summerHOA_Part_out; by HomeID; run;
proc sort data = CA_summerHOA_Part_in; by HomeID; run;

data CAsummer_HOA_part;
merge CA_summerHOA_Part_out CA_summerHOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in HOA_Part_tot_out
HOA_part_tot_in;
run;

*Add column with calcuation of indoor COA for HOA based on
partitioning;
data CAsummer_HOA_part2;

```

```

set CAsummer_HOA_part;
HOA_Ctot = (0.136*OM_out)/HOA_Part_tot_out; *the 0.136 is the mass
fraction total OA that is HOA from the SI of Cappa and Jimenez 2010;
HOA_COA_out = 0.136*OM_out;
HOA_COA_in = HOA_Ctot*HOA_Part_tot_in;
run;

*****
LV-OOA
*****

*calculate total (gas+particle) organic concentration in each LV-OOA
bin using relationship between C* and Ctot from Cappa and Jimenez 2010;

data CA_summerLVOOAtot;
set CA_summer;
LVOOA_C1_tot = 0+0.135*exp(-0.37*(log10(10**-7)-3));
LVOOA_C2_tot = 0+0.135*exp(-0.37*(log10(10**-6)-3));
LVOOA_C3_tot = 0+0.135*exp(-0.37*(log10(10**-5)-3));
LVOOA_C4_tot = 0+0.135*exp(-0.37*(log10(10**-4)-3));
LVOOA_C5_tot = 0+0.135*exp(-0.37*(log10(10**-3)-3));
LVOOA_C6_tot = 0+0.135*exp(-0.37*(log10(10**-2)-3));
LVOOA_C7_tot = 0+0.135*exp(-0.37*(log10(10**-1)-3));
LVOOA_C8_tot = 0+0.135*exp(-0.37*(log10(10**0)-3));
LVOOA_C9_tot = 0+0.135*exp(-0.37*(log10(10**1)-3));
LVOOA_C10_tot = 0+0.135*exp(-0.37*(log10(10**2)-3));
LVOOA_C11_tot = 0+0.135*exp(-0.37*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts outdoors based on temperature
using Clausius Clapyeron Eqn.;
data CA_summerLVOOA_Cout;
set CA_summerLVOOAtot;
LVOOA_C1_T_out = 10**-7*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C2_T_out = 10**-6*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C3_T_out = 10**-5*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C4_T_out = 10**-4*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C5_T_out = 10**-3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C6_T_out = 10**-2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C7_T_out = 10**-1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C8_T_out = 10**0*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C9_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C10_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C11_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
run;

```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```
data CA_summerLVOOA_Part_out;
set CA_summerLVOOA_Cout;
LVOOA_C1_part_out = (1+((LVOOA_C1_T_out)/OM_out))** -1;
LVOOA_C2_part_out = (1+((LVOOA_C2_T_out)/OM_out))** -1;
LVOOA_C3_part_out = (1+((LVOOA_C3_T_out)/OM_out))** -1;
LVOOA_C4_part_out = (1+((LVOOA_C4_T_out)/OM_out))** -1;
LVOOA_C5_part_out = (1+((LVOOA_C5_T_out)/OM_out))** -1;
LVOOA_C6_part_out = (1+((LVOOA_C6_T_out)/OM_out))** -1;
LVOOA_C7_part_out = (1+((LVOOA_C7_T_out)/OM_out))** -1;
LVOOA_C8_part_out = (1+((LVOOA_C8_T_out)/OM_out))** -1;
LVOOA_C9_part_out = (1+((LVOOA_C9_T_out)/OM_out))** -1;
LVOOA_C10_part_out = (1+((LVOOA_C10_T_out)/OM_out))** -1;
LVOOA_C11_part_out = (1+((LVOOA_C11_T_out)/OM_out))** -1;
LVOOA_Part_tot_out =
(LVOOA_C1_part_out*LVOOA_C1_tot+LVOOA_C2_part_out*LVOOA_C2_tot+LVOOA_C3_
_part_out*LVOOA_C3_tot+
LVOOA_C4_part_out*LVOOA_C4_tot+LVOOA_C5_part_out*LVOOA_C5_tot+LVOOA_C6_
part_out*LVOOA_C6_tot
+LVOOA_C7_part_out*LVOOA_C7_tot+LVOOA_C8_part_out*LVOOA_C8_tot+LVOOA_C9_
_part_out*LVOOA_C9_tot+LVOOA_C10_part_out*LVOOA_C10_tot
+LVOOA_C11_part_out*LVOOA_C11_tot)/(LVOOA_C1_tot+LVOOA_C2_tot+LVOOA_C3_
tot+LVOOA_C4_tot+LVOOA_C5_tot+LVOOA_C6_tot+LVOOA_C7_tot+
LVOOA_C8_tot+LVOOA_C9_tot+LVOOA_C10_tot+LVOOA_C11_tot);
run;
```

Calculate how C(Tref = 298.15 K) shifts indoors based on temperature using Clausius Clapyeron Eqn.;

```
data CA_summerLVOOA_Cin;
set CA_summerLVOOA_tot;
LVOOA_C1_T_in = 10** -7*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C2_T_in = 10** -6*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C3_T_in = 10** -5*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C4_T_in = 10** -4*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C5_T_in = 10** -3*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C6_T_in = 10** -2*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C7_T_in = 10** -1*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C8_T_in = 10** 0*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C9_T_in = 10** 1*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C10_T_in = 10** 2*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C11_T_in = 10** 3*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
run;
```

```

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data CA_summerLVOOA_Part_in;
set CA_summerLVOOA_Cin;
LVOOA_C1_part_in = (1+((LVOOA_C1_T_in)/OM_in))**-1;
LVOOA_C2_part_in = (1+((LVOOA_C2_T_in)/OM_in))**-1;
LVOOA_C3_part_in = (1+((LVOOA_C3_T_in)/OM_in))**-1;
LVOOA_C4_part_in = (1+((LVOOA_C4_T_in)/OM_in))**-1;
LVOOA_C5_part_in = (1+((LVOOA_C5_T_in)/OM_in))**-1;
LVOOA_C6_part_in = (1+((LVOOA_C6_T_in)/OM_in))**-1;
LVOOA_C7_part_in = (1+((LVOOA_C7_T_in)/OM_in))**-1;
LVOOA_C8_part_in = (1+((LVOOA_C8_T_in)/OM_in))**-1;
LVOOA_C9_part_in = (1+((LVOOA_C9_T_in)/OM_in))**-1;
LVOOA_C10_part_in = (1+((LVOOA_C10_T_in)/OM_in))**-1;
LVOOA_C11_part_in = (1+((LVOOA_C11_T_in)/OM_in))**-1;
LVOOA_Part_tot_in =
(LVOOA_C1_part_in*LVOOA_C1_tot+LVOOA_C2_part_in*LVOOA_C2_tot+LVOOA_C3_p
art_in*LVOOA_C3_tot+
LVOOA_C4_part_in*LVOOA_C4_tot+LVOOA_C5_part_in*LVOOA_C5_tot+LVOOA_C6_pa
rt_in*LVOOA_C6_tot
+LVOOA_C7_part_in*LVOOA_C7_tot+LVOOA_C8_part_in*LVOOA_C8_tot+LVOOA_C9_p
art_in*LVOOA_C9_tot+LVOOA_C10_part_in*LVOOA_C10_tot
+LVOOA_C11_part_in*LVOOA_C11_tot)/(LVOOA_C1_tot+LVOOA_C2_tot+LVOOA_C3_t
ot+LVOOA_C4_tot+LVOOA_C5_tot+LVOOA_C6_tot+LVOOA_C7_tot+
LVOOA_C8_tot+LVOOA_C9_tot+LVOOA_C10_tot+LVOOA_C11_tot);
run;

*Merge indoor and outdoor partitioning data sets and keep only key
data;
proc sort data = CA_summerLVOOA_Part_out; by HomeID;run;
proc sort data = CA_summerLVOOA_Part_in; by HomeID;run;

data CAsummer_LVOOA_part;
merge CA_summerLVOOA_Part_out CA_summerLVOOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in LVOOA_Part_tot_out
LVOOA_part_tot_in;
run;

*Add column with calcuation of indoor COA for LV-OOA based on
partitioning;
data CAsummer_LVOOA_part2;
set CAsummer_LVOOA_part;
LVOOA_Ctot = 0.318*OM_out/LVOOA_Part_tot_out; *the 0.318 is the mass
fraction total OA that is LVOOA from the SI of Cappa and Jimenez 2010;
LVOOA_COA_out = 0.318*OM_out;
LVOOA_COA_in = LVOOA_Ctot*LVOOA_Part_tot_in;
run;

*****
SV-OOA
*****

*calculate total (gas+particle) organic concentration in each SV-OOA
bin using relationship between C* and Ctot from Cappa and Jimenez 2010;
data CA_summerSVOOAtot;

```

```

set CA_summer;
SVOOA_C1_tot = 0.7+7*exp(0.3*(log10(10**-5)-3));
SVOOA_C2_tot = 0.7+7*exp(0.3*(log10(10**-4)-3));
SVOOA_C3_tot = 0.7+7*exp(0.3*(log10(10**-3)-3));
SVOOA_C4_tot = 0.7+7*exp(0.3*(log10(10**-2)-3));
SVOOA_C5_tot = 0.7+7*exp(0.3*(log10(10**-1)-3));
SVOOA_C6_tot = 0.7+7*exp(0.3*(log10(10**0)-3));
SVOOA_C7_tot = 0.7+7*exp(0.3*(log10(10**1)-3));
SVOOA_C8_tot = 0.7+7*exp(0.3*(log10(10**2)-3));
SVOOA_C9_tot = 0.7+7*exp(0.3*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts outdoors based on temperature
using Clausius Clapyeron Eqn.;
data CA_summerSVOOA_Cout;
set CA_summerSVOOA_tot;
SVOOA_C1_T_out = 10**-5*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
SVOOA_C2_T_out = 10**-4*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
SVOOA_C3_T_out = 10**-3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
SVOOA_C4_T_out = 10**-2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
SVOOA_C5_T_out = 10**-1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
SVOOA_C6_T_out = 10**0*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
SVOOA_C7_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
SVOOA_C8_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
SVOOA_C9_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15)))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and
Jimenez 2010;
data CA_summerSVOOA_Part_out;
set CA_summerSVOOA_Cout;
SVOOA_C1_part_out = (1+((SVOOA_C1_T_out)/OM_out))**-1;
SVOOA_C2_part_out = (1+((SVOOA_C2_T_out)/OM_out))**-1;
SVOOA_C3_part_out = (1+((SVOOA_C3_T_out)/OM_out))**-1;
SVOOA_C4_part_out = (1+((SVOOA_C4_T_out)/OM_out))**-1;
SVOOA_C5_part_out = (1+((SVOOA_C5_T_out)/OM_out))**-1;
SVOOA_C6_part_out = (1+((SVOOA_C6_T_out)/OM_out))**-1;
SVOOA_C7_part_out = (1+((SVOOA_C7_T_out)/OM_out))**-1;
SVOOA_C8_part_out = (1+((SVOOA_C8_T_out)/OM_out))**-1;
SVOOA_C9_part_out = (1+((SVOOA_C9_T_out)/OM_out))**-1;
SVOOA_Part_tot_out =
(SVOOA_C1_part_out*SVOOA_C1_tot+SVOOA_C2_part_out*SVOOA_C2_tot+SVOOA_C3
_part_out*SVOOA_C3_tot+
SVOOA_C4_part_out*SVOOA_C4_tot+SVOOA_C5_part_out*SVOOA_C5_tot+SVOOA_C6
_part_out*SVOOA_C6_tot
+SVOOA_C7_part_out*SVOOA_C7_tot+SVOOA_C8_part_out*SVOOA_C8_tot+SVOOA_C9
_part_out*SVOOA_C9_tot)/

```

```
(SVOOA_C1_tot+SVOOA_C2_tot+SVOOA_C3_tot+SVOOA_C4_tot+SVOOA_C5_tot+SVOOA_C6_tot+SVOOA_C7_tot+SVOOA_C8_tot+SVOOA_C9_tot);
run;
```

Calculate how C(Tref = 298.15 K) shifts indoors based on temperature using Clausius Clapyeron Eqn.;

```
data CA_summerSVOOA_Cin;
set CA_summerSVOOA_tot;
SVOOA_C1_T_in = 10**-5*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C2_T_in = 10**-4*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C3_T_in = 10**-3*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C4_T_in = 10**-2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C5_T_in = 10**-1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C6_T_in = 10**0*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C7_T_in = 10**1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C8_T_in = 10**2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C9_T_in = 10**3*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
run;
```

*calculate the indoor partitioning based on VBS from Cappa and Jimenez 2010;

```
data CA_summerSVOOA_Part_in;
set CA_summerSVOOA_Cin;
SVOOA_C1_part_in = (1+((SVOOA_C1_T_in)/OM_in))**-1;
SVOOA_C2_part_in = (1+((SVOOA_C2_T_in)/OM_in))**-1;
SVOOA_C3_part_in = (1+((SVOOA_C3_T_in)/OM_in))**-1;
SVOOA_C4_part_in = (1+((SVOOA_C4_T_in)/OM_in))**-1;
SVOOA_C5_part_in = (1+((SVOOA_C5_T_in)/OM_in))**-1;
SVOOA_C6_part_in = (1+((SVOOA_C6_T_in)/OM_in))**-1;
SVOOA_C7_part_in = (1+((SVOOA_C7_T_in)/OM_in))**-1;
SVOOA_C8_part_in = (1+((SVOOA_C8_T_in)/OM_in))**-1;
SVOOA_C9_part_in = (1+((SVOOA_C9_T_in)/OM_in))**-1;
SVOOA_Part_tot_in =
(SVOOA_C1_part_in*SVOOA_C1_tot+SVOOA_C2_part_in*SVOOA_C2_tot+SVOOA_C3_part_in*SVOOA_C3_tot+
SVOOA_C4_part_in*SVOOA_C4_tot+SVOOA_C5_part_in*SVOOA_C5_tot+SVOOA_C6_part_in*SVOOA_C6_tot+
SVOOA_C7_part_in*SVOOA_C7_tot+SVOOA_C8_part_in*SVOOA_C8_tot+SVOOA_C9_part_in*SVOOA_C9_tot)/
(SVOOA_C1_tot+SVOOA_C2_tot+SVOOA_C3_tot+SVOOA_C4_tot+SVOOA_C5_tot+SVOOA_C6_tot+SVOOA_C7_tot+
SVOOA_C8_tot+SVOOA_C9_tot);
run;
```

*Merge indoor and outdoor Partitioning data sets and keep only key data;

```
proc sort data = CA_summerSVOOA_Part_out; by HomeID;run;
```



```

proc sort data = CA_summerSVOOA_Part_in; by HomeID;run;

data CAsummer_SVOOA_part;
merge CA_summerSVOOA_Part_out CA_summerSVOOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in SVOOA_Part_tot_out
SVOOA_part_tot_in;
run;

*Add column with calculation of indoor COA for SV-OOA based on
partitioning;
data CAsummer_SVOOA_part2;
set CAsummer_SVOOA_part;
SVOOA_Ctot = 0.477*OM_out/SVOOA_Part_tot_out; *the 0.477 is the mass
fraction total OA that is SVOOA from the SI of Cappa and Jimenez 2010;
SVOOA_COA_out = 0.477*OM_out;
SVOOA_COA_in = SVOOA_Ctot*SVOOA_Part_tot_in;
run;

*****
other OA (OTH)
*****

*calculate total (gas+particle) organic concentration in each OTH bin
using relationship between C* and Ctot from Cappa and Jimenez 2010
using the values for OA;
data CA_summerOTHtot;
set CA_summer;
OTH_C1_tot = 1.65+19*exp(1*(log10(10**-6)-3));
OTH_C2_tot = 1.65+19*exp(1*(log10(10**-5)-3));
OTH_C3_tot = 1.65+19*exp(1*(log10(10**-4)-3));
OTH_C4_tot = 1.65+19*exp(1*(log10(10**-3)-3));
OTH_C5_tot = 1.65+19*exp(1*(log10(10**-2)-3));
OTH_C6_tot = 1.65+19*exp(1*(log10(10**-1)-3));
OTH_C7_tot = 1.65+19*exp(1*(log10(10**0)-3));
OTH_C8_tot = 1.65+19*exp(1*(log10(10**1)-3));
OTH_C9_tot = 1.65+19*exp(1*(log10(10**2)-3));
OTH_C10_tot = 1.65+19*exp(1*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
data CA_summerOTH_Cout;
set CA_summerOTHtot;
OTH_C1_T_out = 10**-6*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OTH_C2_T_out = 10**-5*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OTH_C3_T_out = 10**-4*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OTH_C4_T_out = 10**-3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OTH_C5_T_out = 10**-2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OTH_C6_T_out = 10**-1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));

```

```

OTH_C7_T_out = 10**0*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OTH_C8_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OTH_C9_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OTH_C10_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data CA_summerOTH_Part_out;
set CA_summerOTH_Cout;
OTH_C1_part_out = (1+((OTH_C1_T_out)/OM_out))**-1;
OTH_C2_part_out = (1+((OTH_C2_T_out)/OM_out))**-1;
OTH_C3_part_out = (1+((OTH_C3_T_out)/OM_out))**-1;
OTH_C4_part_out = (1+((OTH_C4_T_out)/OM_out))**-1;
OTH_C5_part_out = (1+((OTH_C5_T_out)/OM_out))**-1;
OTH_C6_part_out = (1+((OTH_C6_T_out)/OM_out))**-1;
OTH_C7_part_out = (1+((OTH_C7_T_out)/OM_out))**-1;
OTH_C8_part_out = (1+((OTH_C8_T_out)/OM_out))**-1;
OTH_C9_part_out = (1+((OTH_C9_T_out)/OM_out))**-1;
OTH_C10_part_out = (1+((OTH_C10_T_out)/OM_out))**-1;
OTH_Part_tot_out =
(OTH_C1_part_out*OTH_C1_tot+OTH_C2_part_out*OTH_C2_tot+OTH_C3_part_out*
OTH_C3_tot+OTH_C4_part_out*OTH_C4_tot+OTH_C5_part_out*OTH_C5_tot+OTH_C6
_part_out*OTH_C6_tot
+OTH_C7_part_out*OTH_C7_tot+OTH_C8_part_out*OTH_C8_tot+OTH_C9_part_out*
OTH_C9_tot+OTH_C10_part_out*OTH_C10_tot)/
(OTH_C1_tot+OTH_C2_tot+OTH_C3_tot+OTH_C4_tot+OTH_C5_tot+OTH_C6_tot+OTH_
C7_tot+OTH_C8_tot+OTH_C9_tot+OTH_C10_tot);
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature indoors
using Clausius Clapyeron Eqn.;
data CA_summerOTH_Cin;
set CA_summerOTHtot;
OTH_C1_T_in = 10**6*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C2_T_in = 10**5*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C3_T_in = 10**4*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C4_T_in = 10**3*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C5_T_in = 10**2*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C6_T_in = 10**1*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C7_T_in = 10**0*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C8_T_in = 10**1*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C9_T_in = 10**2*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));

```

```

OTH_C10_T_in = 10**3*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data CA_summerOTH_Part_in;
set CA_summerOTH_Cin;
OTH_C1_part_in = (1+((OTH_C1_T_in)/OM_in))**-1;
OTH_C2_part_in = (1+((OTH_C2_T_in)/OM_in))**-1;
OTH_C3_part_in = (1+((OTH_C3_T_in)/OM_in))**-1;
OTH_C4_part_in = (1+((OTH_C4_T_in)/OM_in))**-1;
OTH_C5_part_in = (1+((OTH_C5_T_in)/OM_in))**-1;
OTH_C6_part_in = (1+((OTH_C6_T_in)/OM_in))**-1;
OTH_C7_part_in = (1+((OTH_C7_T_in)/OM_in))**-1;
OTH_C8_part_in = (1+((OTH_C8_T_in)/OM_in))**-1;
OTH_C9_part_in = (1+((OTH_C9_T_in)/OM_in))**-1;
OTH_C10_part_in = (1+((OTH_C10_T_in)/OM_in))**-1;
OTH_Part_tot_in =
(OTH_C1_part_in*OTH_C1_tot+OTH_C2_part_in*OTH_C2_tot+OTH_C3_part_in*OTH_
_C3_tot+OTH_C4_part_in*OTH_C4_tot+OTH_C5_part_in*OTH_C5_tot+OTH_C6_part
_in*OTH_C6_tot
+OTH_C7_part_in*OTH_C7_tot+OTH_C8_part_in*OTH_C8_tot+OTH_C9_part_in*OTH_
_C9_tot+OTH_C10_part_in*OTH_C10_tot)/
(OTH_C1_tot+OTH_C2_tot+OTH_C3_tot+OTH_C4_tot+OTH_C5_tot+OTH_C6_tot+OTH_
C7_tot+OTH_C8_tot+OTH_C9_tot+OTH_C10_tot);
run;

*Merge indoor and outdoor Partitioning data sets and keep only key
data;
proc sort data = CA_summerOTH_Part_out; by HomeID;run;
proc sort data = CA_summerOTH_Part_in; by HomeID;run;

data CAsummer_OTH_part;
merge CA_summerOTH_Part_out CA_summerOTH_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in OTH_Part_tot_out
OTH_part_tot_in;
run;

*Add column with calculation of indoor COA for OTH based on
partitioning;
data CAsummer_OTH_part2;
set CAsummer_OTH_part;
OTH_Ctot = (0.0682*OM_out)/OTH_Part_tot_out; *the 0.0682 is the mass
fraction total OA that is OTH from the SI of Cappa and Jimenez 2010;
OTH_COA_out = 0.0682*OM_out;
OTH_COA_in = OTH_Ctot*OTH_Part_tot_in;
run;

*Merge the HOA, LV-OOA, SV-OOA, and other OA data sets;
data CA_Summer_Part;
merge CAsummer_HOA_part2 CAsummer_LVOOA_part2 CAsummer_SVOOA_part2
CAsummer_OTH_part2;
by HomeID;
Calc_OA_Out = HOA_COA_Out + LVOOA_COA_Out+SVOOA_COA_out + OTH_COA_out;
Calc_OA_in = HOA_COA_in + LVOOA_COA_in+SVOOA_COA_in + OTH_COA_in;

```

```
run;
```

```
*This file was exported to library "Ch5" and is named "OA components by
season.xls," sheet name "CA";
```

```
*Elizabeth, NJ RIOPA homes using NYC OA component mass fractions
detailed in Jimenez et al. 2009;
*Note the components are broken up by summer and winter, so I will
break up the measured data into warm and cold seasons;
```

```
*Import data set with seasons and state;
```

```
PROC IMPORT OUT= Season
      DATAFILE= "G:\Chapter 5\Season and State.xls"
      DBMS=EXCEL REPLACE;
      SHEET="All";
      GETNAMES=YES;
      MIXED=YES;
      USEDATE=YES;
      SCANTIME=YES;
```

```
RUN;
```

```
*****
```

```
      Summer
```

```
*****
```

```
*create data set with summer dates only;
```

```
data NJ_summer;
set Season;
if season NE "s" then delete;
if state NE "NJ" then delete;
run;
```

```
*****
```

```
      HOA
```

```
*****
```

```
*calculate total (gas+particle) organic concentration in each HOA bin
using relationship between C* and Ctot from Cappa and Jimenez 2010;
```

```
data NJ_summerHOAtot;
set NJ_summer;
HOA_C1_tot = 0.78+23*exp(0.7*(log10(10**-6)-3));
HOA_C2_tot = 0.78+23*exp(0.7*(log10(10**-5)-3));
HOA_C3_tot = 0.78+23*exp(0.7*(log10(10**-4)-3));
HOA_C4_tot = 0.78+23*exp(0.7*(log10(10**-3)-3));
HOA_C5_tot = 0.78+23*exp(0.7*(log10(10**-2)-3));
HOA_C6_tot = 0.78+23*exp(0.7*(log10(10**-1)-3));
HOA_C7_tot = 0.78+23*exp(0.7*(log10(10**0)-3));
HOA_C8_tot = 0.78+23*exp(0.7*(log10(10**1)-3));
HOA_C9_tot = 0.78+23*exp(0.7*(log10(10**2)-3));
HOA_C10_tot = 0.78+23*exp(0.7*(log10(10**3)-3));
run;
```

```
*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
```

```

data NJ_summerHOA_Cout;
set NJ_summerHOAtot;
HOA_C1_T_out = 10**(-6*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C2_T_out = 10**(-5*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C3_T_out = 10**(-4*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C4_T_out = 10**(-3*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C5_T_out = 10**(-2*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C6_T_out = 10**(-1*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C7_T_out = 10**(-0*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C8_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C9_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C10_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15)))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data NJ_summerHOA_Part_out;
set NJ_summerHOA_Cout;
HOA_C1_part_out = (1+((HOA_C1_T_out)/OM_out))**-1;
HOA_C2_part_out = (1+((HOA_C2_T_out)/OM_out))**-1;
HOA_C3_part_out = (1+((HOA_C3_T_out)/OM_out))**-1;
HOA_C4_part_out = (1+((HOA_C4_T_out)/OM_out))**-1;
HOA_C5_part_out = (1+((HOA_C5_T_out)/OM_out))**-1;
HOA_C6_part_out = (1+((HOA_C6_T_out)/OM_out))**-1;
HOA_C7_part_out = (1+((HOA_C7_T_out)/OM_out))**-1;
HOA_C8_part_out = (1+((HOA_C8_T_out)/OM_out))**-1;
HOA_C9_part_out = (1+((HOA_C9_T_out)/OM_out))**-1;
HOA_C10_part_out = (1+((HOA_C10_T_out)/OM_out))**-1;
HOA_Part_tot_out =
(HOA_C1_part_out*HOA_C1_tot+HOA_C2_part_out*HOA_C2_tot+HOA_C3_part_out*
HOA_C3_tot+HOA_C4_part_out*HOA_C4_tot+HOA_C5_part_out*HOA_C5_tot+HOA_C6
_part_out*HOA_C6_tot
+HOA_C7_part_out*HOA_C7_tot+HOA_C8_part_out*HOA_C8_tot+HOA_C9_part_out*
HOA_C9_tot+HOA_C10_part_out*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature indoors
using Clausius Clapyeron Eqn.;
data NJ_summerHOA_Cin;
set NJ_summerHOAtot;
HOA_C1_T_in = 10**(-6*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C2_T_in = 10**(-5*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15)))));

```

```

HOA_C3_T_in = 10**-4*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C4_T_in = 10**-3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C5_T_in = 10**-2*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C6_T_in = 10**-1*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C7_T_in = 10**0*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15)))));
HOA_C8_T_in = 10**1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15)))));
HOA_C9_T_in = 10**2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15)))));
HOA_C10_T_in = 10**3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data NJ_summerHOA_Part_in;
set NJ_summerHOA_Cin;
HOA_C1_part_in = (1+((HOA_C1_T_in)/OM_in))**-1;
HOA_C2_part_in = (1+((HOA_C2_T_in)/OM_in))**-1;
HOA_C3_part_in = (1+((HOA_C3_T_in)/OM_in))**-1;
HOA_C4_part_in = (1+((HOA_C4_T_in)/OM_in))**-1;
HOA_C5_part_in = (1+((HOA_C5_T_in)/OM_in))**-1;
HOA_C6_part_in = (1+((HOA_C6_T_in)/OM_in))**-1;
HOA_C7_part_in = (1+((HOA_C7_T_in)/OM_in))**-1;
HOA_C8_part_in = (1+((HOA_C8_T_in)/OM_in))**-1;
HOA_C9_part_in = (1+((HOA_C9_T_in)/OM_in))**-1;
HOA_C10_part_in = (1+((HOA_C10_T_in)/OM_in))**-1;
HOA_Part_tot_in =
(HOA_C1_part_in*HOA_C1_tot+HOA_C2_part_in*HOA_C2_tot+HOA_C3_part_in*HOA
_C3_tot+HOA_C4_part_in*HOA_C4_tot+HOA_C5_part_in*HOA_C5_tot+HOA_C6_part
_in*HOA_C6_tot
+HOA_C7_part_in*HOA_C7_tot+HOA_C8_part_in*HOA_C8_tot+HOA_C9_part_in*HOA
_C9_tot+HOA_C10_part_in*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;

*Merge indoor and outdoor Partitioning data sets and keep only key
data;
proc sort data = NJ_summerHOA_Part_out; by HomeID; run;
proc sort data = NJ_summerHOA_Part_in; by HomeID; run;

data NJsummer_HOA_part;
merge NJ_summerHOA_Part_out NJ_summerHOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in HOA_Part_tot_out
HOA_part_tot_in;
run;

*Add column with calculation of indoor COA for HOA based on
partitioning;
data NJsummer_HOA_part2;

```

```

set NJsummer_HOA_part;
HOA_Ctot = 0.186*OM_out/HOA_Part_tot_out; *the 0.186 is the mass
fraction total OA that is HOA from the SI of Cappa and Jimenez 2010;
HOA_COA_out = 0.186*OM_out;
HOA_COA_in = HOA_Ctot*HOA_Part_tot_in;
run;

*****
LV-OOA
*****

*calculate total (gas+particle) organic concentration in each LV-OOA
bin using relationship between C* and Ctot from Cappa and Jimenez 2010;

data NJ_summerLVOOAtot;
set NJ_summer;
LVOOA_C1_tot = 0+0.135*exp(-0.37*(log10(10**-7)-3));
LVOOA_C2_tot = 0+0.135*exp(-0.37*(log10(10**-6)-3));
LVOOA_C3_tot = 0+0.135*exp(-0.37*(log10(10**-5)-3));
LVOOA_C4_tot = 0+0.135*exp(-0.37*(log10(10**-4)-3));
LVOOA_C5_tot = 0+0.135*exp(-0.37*(log10(10**-3)-3));
LVOOA_C6_tot = 0+0.135*exp(-0.37*(log10(10**-2)-3));
LVOOA_C7_tot = 0+0.135*exp(-0.37*(log10(10**-1)-3));
LVOOA_C8_tot = 0+0.135*exp(-0.37*(log10(10**0)-3));
LVOOA_C9_tot = 0+0.135*exp(-0.37*(log10(10**1)-3));
LVOOA_C10_tot = 0+0.135*exp(-0.37*(log10(10**2)-3));
LVOOA_C11_tot = 0+0.135*exp(-0.37*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts outdoors based on temperature
using Clausius Clapyeron Eqn.;
data NJ_summerLVOOA_Cout;
set NJ_summerLVOOAtot;
LVOOA_C1_T_out = 10**-7*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C2_T_out = 10**-6*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C3_T_out = 10**-5*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C4_T_out = 10**-4*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C5_T_out = 10**-3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C6_T_out = 10**-2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C7_T_out = 10**-1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C8_T_out = 10**0*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C9_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C10_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
LVOOA_C11_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
run;

```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```
data NJ_summerLVOOA_Part_out;
set NJ_summerLVOOA_Cout;
LVOOA_C1_part_out = (1+((LVOOA_C1_T_out)/OM_out))**-1;
LVOOA_C2_part_out = (1+((LVOOA_C2_T_out)/OM_out))**-1;
LVOOA_C3_part_out = (1+((LVOOA_C3_T_out)/OM_out))**-1;
LVOOA_C4_part_out = (1+((LVOOA_C4_T_out)/OM_out))**-1;
LVOOA_C5_part_out = (1+((LVOOA_C5_T_out)/OM_out))**-1;
LVOOA_C6_part_out = (1+((LVOOA_C6_T_out)/OM_out))**-1;
LVOOA_C7_part_out = (1+((LVOOA_C7_T_out)/OM_out))**-1;
LVOOA_C8_part_out = (1+((LVOOA_C8_T_out)/OM_out))**-1;
LVOOA_C9_part_out = (1+((LVOOA_C9_T_out)/OM_out))**-1;
LVOOA_C10_part_out = (1+((LVOOA_C10_T_out)/OM_out))**-1;
LVOOA_C11_part_out = (1+((LVOOA_C11_T_out)/OM_out))**-1;
LVOOA_Part_tot_out =
(LVOOA_C1_part_out*LVOOA_C1_tot+LVOOA_C2_part_out*LVOOA_C2_tot+LVOOA_C3
_part_out*LVOOA_C3_tot+
LVOOA_C4_part_out*LVOOA_C4_tot+LVOOA_C5_part_out*LVOOA_C5_tot+LVOOA_C6
_part_out*LVOOA_C6_tot
+LVOOA_C7_part_out*LVOOA_C7_tot+LVOOA_C8_part_out*LVOOA_C8_tot+LVOOA_C9
_part_out*LVOOA_C9_tot+LVOOA_C10_part_out*LVOOA_C10_tot
+LVOOA_C11_part_out*LVOOA_C11_tot)/(LVOOA_C1_tot+LVOOA_C2_tot+LVOOA_C3
_tot+LVOOA_C4_tot+LVOOA_C5_tot+LVOOA_C6_tot+LVOOA_C7_tot+
LVOOA_C8_tot+LVOOA_C9_tot+LVOOA_C10_tot+LVOOA_C11_tot);
run;
```

Calculate how C(Tref = 298.15 K) shifts indoors based on temperature using Clausius Clapyeron Eqn.;

```
data NJ_summerLVOOA_Cin;
set NJ_summerLVOOA_tot;
LVOOA_C1_T_in = 10**-7*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C2_T_in = 10**-6*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C3_T_in = 10**-5*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C4_T_in = 10**-4*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C5_T_in = 10**-3*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C6_T_in = 10**-2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C7_T_in = 10**-1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C8_T_in = 10**0*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C9_T_in = 10**1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C10_T_in = 10**2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
LVOOA_C11_T_in = 10**3*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-(1/298.15))));
run;
```



```

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data NJ_summerLVOOA_Part_in;
set NJ_summerLVOOA_Cin;
LVOOA_C1_part_in = (1+((LVOOA_C1_T_in)/OM_in))**-1;
LVOOA_C2_part_in = (1+((LVOOA_C2_T_in)/OM_in))**-1;
LVOOA_C3_part_in = (1+((LVOOA_C3_T_in)/OM_in))**-1;
LVOOA_C4_part_in = (1+((LVOOA_C4_T_in)/OM_in))**-1;
LVOOA_C5_part_in = (1+((LVOOA_C5_T_in)/OM_in))**-1;
LVOOA_C6_part_in = (1+((LVOOA_C6_T_in)/OM_in))**-1;
LVOOA_C7_part_in = (1+((LVOOA_C7_T_in)/OM_in))**-1;
LVOOA_C8_part_in = (1+((LVOOA_C8_T_in)/OM_in))**-1;
LVOOA_C9_part_in = (1+((LVOOA_C9_T_in)/OM_in))**-1;
LVOOA_C10_part_in = (1+((LVOOA_C10_T_in)/OM_in))**-1;
LVOOA_C11_part_in = (1+((LVOOA_C11_T_in)/OM_in))**-1;
LVOOA_Part_tot_in =
(LVOOA_C1_part_in*LVOOA_C1_tot+LVOOA_C2_part_in*LVOOA_C2_tot+LVOOA_C3_p
art_in*LVOOA_C3_tot+
LVOOA_C4_part_in*LVOOA_C4_tot+LVOOA_C5_part_in*LVOOA_C5_tot+LVOOA_C6_pa
rt_in*LVOOA_C6_tot
+LVOOA_C7_part_in*LVOOA_C7_tot+LVOOA_C8_part_in*LVOOA_C8_tot+LVOOA_C9_p
art_in*LVOOA_C9_tot+LVOOA_C10_part_in*LVOOA_C10_tot
+LVOOA_C11_part_in*LVOOA_C11_tot)/(LVOOA_C1_tot+LVOOA_C2_tot+LVOOA_C3_t
ot+LVOOA_C4_tot+LVOOA_C5_tot+LVOOA_C6_tot+LVOOA_C7_tot+
LVOOA_C8_tot+LVOOA_C9_tot+LVOOA_C10_tot+LVOOA_C11_tot);
run;

*Merge indoor and outdoor Partitioning data sets and keep only key
data;
proc sort data = NJ_summerLVOOA_Part_out; by HomeID;run;
proc sort data = NJ_summerLVOOA_Part_in; by HomeID;run;

data NJsummer_LVOOA_part;
merge NJ_summerLVOOA_Part_out NJ_summerLVOOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in LVOOA_Part_tot_out
LVOOA_part_tot_in;
run;

*Add column with calculation of indoor COA for LV-OOA based on
partitioning;
data NJsummer_LVOOA_part2;
set NJsummer_LVOOA_part;
LVOOA_Ctot = 0.508*OM_out/LVOOA_Part_tot_out; *the 0.508 is the mass
fraction total OA that is LVOOA from the SI of Cappa and Jimenez 2010;
LVOOA_COA_out = 0.508*OM_out;
LVOOA_COA_in = LVOOA_Ctot*LVOOA_Part_tot_in;
run;

*****
SV-OOA
*****

*calculate total (gas+particle) organic concentration in each SV-OOA
bin using relationship between C* and Ctot from Cappa and Jimenez 2010;
data NJ_summerSVOOAtot;
set NJ_summer;

```

```

SVOOA_C1_tot = 0.7+7*exp(0.3*(log10(10**-5)-3));
SVOOA_C2_tot = 0.7+7*exp(0.3*(log10(10**-4)-3));
SVOOA_C3_tot = 0.7+7*exp(0.3*(log10(10**-3)-3));
SVOOA_C4_tot = 0.7+7*exp(0.3*(log10(10**-2)-3));
SVOOA_C5_tot = 0.7+7*exp(0.3*(log10(10**-1)-3));
SVOOA_C6_tot = 0.7+7*exp(0.3*(log10(10**0)-3));
SVOOA_C7_tot = 0.7+7*exp(0.3*(log10(10**1)-3));
SVOOA_C8_tot = 0.7+7*exp(0.3*(log10(10**2)-3));
SVOOA_C9_tot = 0.7+7*exp(0.3*(log10(10**3)-3));
run;

```

Calculate how C(Tref = 298.15 K) shifts outdoors based on temperature using Clausius Clapyeron Eqn.;

```

data NJ_summerSVOOA_Cout;
set NJ_summerSVOOA_tot;
SVOOA_C1_T_out = 10**-5*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
SVOOA_C2_T_out = 10**-4*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
SVOOA_C3_T_out = 10**-3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
SVOOA_C4_T_out = 10**-2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
SVOOA_C5_T_out = 10**-1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
SVOOA_C6_T_out = 10**0*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
SVOOA_C7_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
SVOOA_C8_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
SVOOA_C9_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
run;

```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```

data NJ_summerSVOOA_Part_out;
set NJ_summerSVOOA_Cout;
SVOOA_C1_part_out = (1+((SVOOA_C1_T_out)/OM_out))**-1;
SVOOA_C2_part_out = (1+((SVOOA_C2_T_out)/OM_out))**-1;
SVOOA_C3_part_out = (1+((SVOOA_C3_T_out)/OM_out))**-1;
SVOOA_C4_part_out = (1+((SVOOA_C4_T_out)/OM_out))**-1;
SVOOA_C5_part_out = (1+((SVOOA_C5_T_out)/OM_out))**-1;
SVOOA_C6_part_out = (1+((SVOOA_C6_T_out)/OM_out))**-1;
SVOOA_C7_part_out = (1+((SVOOA_C7_T_out)/OM_out))**-1;
SVOOA_C8_part_out = (1+((SVOOA_C8_T_out)/OM_out))**-1;
SVOOA_C9_part_out = (1+((SVOOA_C9_T_out)/OM_out))**-1;
SVOOA_Part_tot_out =
(SVOOA_C1_part_out*SVOOA_C1_tot+SVOOA_C2_part_out*SVOOA_C2_tot+SVOOA_C3_part_out*SVOOA_C3_tot+
SVOOA_C4_part_out*SVOOA_C4_tot+SVOOA_C5_part_out*SVOOA_C5_tot+SVOOA_C6_part_out*SVOOA_C6_tot
+SVOOA_C7_part_out*SVOOA_C7_tot+SVOOA_C8_part_out*SVOOA_C8_tot+SVOOA_C9_part_out*SVOOA_C9_tot)/
(SVOOA_C1_tot+SVOOA_C2_tot+SVOOA_C3_tot+SVOOA_C4_tot+SVOOA_C5_tot+SVOOA_C6_tot+SVOOA_C7_tot+

```

```
SVOOA_C8_tot+SVOOA_C9_tot);
run;
```

```
*Calculate how C*(Tref = 298.15 K) shifts indoors based on temperature
using Clausius Clapyeron Eqn.;
```

```
data NJ_summerSVOOA_Cin;
set NJ_summerSVOOA_tot;
SVOOA_C1_T_in = 10**-5*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C2_T_in = 10**-4*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C3_T_in = 10**-3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C4_T_in = 10**-2*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C5_T_in = 10**-1*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C6_T_in = 10**0*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C7_T_in = 10**1*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C8_T_in = 10**2*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
SVOOA_C9_T_in = 10**3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
run;
```

```
*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
```

```
data NJ_summerSVOOA_Part_in;
set NJ_summerSVOOA_Cin;
SVOOA_C1_part_in = (1+((SVOOA_C1_T_in)/OM_in))**-1;
SVOOA_C2_part_in = (1+((SVOOA_C2_T_in)/OM_in))**-1;
SVOOA_C3_part_in = (1+((SVOOA_C3_T_in)/OM_in))**-1;
SVOOA_C4_part_in = (1+((SVOOA_C4_T_in)/OM_in))**-1;
SVOOA_C5_part_in = (1+((SVOOA_C5_T_in)/OM_in))**-1;
SVOOA_C6_part_in = (1+((SVOOA_C6_T_in)/OM_in))**-1;
SVOOA_C7_part_in = (1+((SVOOA_C7_T_in)/OM_in))**-1;
SVOOA_C8_part_in = (1+((SVOOA_C8_T_in)/OM_in))**-1;
SVOOA_C9_part_in = (1+((SVOOA_C9_T_in)/OM_in))**-1;
SVOOA_Part_tot_in =
(SVOOA_C1_part_in*SVOOA_C1_tot+SVOOA_C2_part_in*SVOOA_C2_tot+SVOOA_C3_p
art_in*SVOOA_C3_tot+
SVOOA_C4_part_in*SVOOA_C4_tot+SVOOA_C5_part_in*SVOOA_C5_tot+SVOOA_C6_pa
rt_in*SVOOA_C6_tot
+SVOOA_C7_part_in*SVOOA_C7_tot+SVOOA_C8_part_in*SVOOA_C8_tot+SVOOA_C9_p
art_in*SVOOA_C9_tot)/
(SVOOA_C1_tot+SVOOA_C2_tot+SVOOA_C3_tot+SVOOA_C4_tot+SVOOA_C5_tot+SVOOA
_C6_tot+SVOOA_C7_tot+
SVOOA_C8_tot+SVOOA_C9_tot);
run;
```

```
*Merge indoor and outdoor Partitioning data sets and keep only key
data;
```

```
proc sort data = NJ_summerSVOOA_Part_out; by HomeID;run;
```

```

proc sort data = NJ_summerSVOOA_Part_in; by HomeID;run;

data NJsummer_SVOOA_part;
merge NJ_summerSVOOA_Part_out NJ_summerSVOOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in SVOOA_Part_tot_out
SVOOA_part_tot_in;
run;

*Add column with calculation of indoor COA for SV-OOA based on
partitioning;
data NJsummer_SVOOA_part2;
set NJsummer_SVOOA_part;
SVOOA_Ctot = 0.305*OM_out/SVOOA_Part_tot_out; *the 0.305 is the mass
fraction total OA that is SVOOA from the SI of Cappa and Jimenez 2010;
SVOOA_COA_out = 0.305*OM_out;
SVOOA_COA_in = SVOOA_Ctot*SVOOA_Part_tot_in;
run;

*Merge the HOA, LV-OOA, and SV-OOA data sets;
data NJ_Summer_Part;
merge NJsummer_HOA_part2 NJsummer_LVOOA_part2 NJsummer_SVOOA_part2;
by HomeID;
Calc_OA_Out = HOA_COA_Out + LVOOA_COA_Out+SVOOA_COA_out;
Calc_OA_in = HOA_COA_in + LVOOA_COA_in+SVOOA_COA_in;
run;

*This file was exported to library "Ch5" and is named "OA components by
season.xls," sheet "NJ_Summer";

*****
Winter
*****

*create data set with NJ winter dates only;
data NJ_winter;
set Season;
if season NE "w" then delete;
if state NE "NJ" then delete;
run;

*****
HOA
*****

*calculate total (gas+particle) organic concentration in each HOA bin
using relationship between C* and Ctot from Cappa and Jimenez 2010;
data NJ_winterHOAtot;
set NJ_winter;
HOA_C1_tot = 0.78+23*exp(0.7*(log10(10**-6)-3));
HOA_C2_tot = 0.78+23*exp(0.7*(log10(10**-5)-3));
HOA_C3_tot = 0.78+23*exp(0.7*(log10(10**-4)-3));
HOA_C4_tot = 0.78+23*exp(0.7*(log10(10**-3)-3));

```

```

HOA_C5_tot = 0.78+23*exp(0.7*(log10(10**-2)-3));
HOA_C6_tot = 0.78+23*exp(0.7*(log10(10**-1)-3));
HOA_C7_tot = 0.78+23*exp(0.7*(log10(10**0)-3));
HOA_C8_tot = 0.78+23*exp(0.7*(log10(10**1)-3));
HOA_C9_tot = 0.78+23*exp(0.7*(log10(10**2)-3));
HOA_C10_tot = 0.78+23*exp(0.7*(log10(10**3)-3));
run;

```

Calculate how C(Tref = 298.15 K) shifts based on temperature using Clausius Clapyeron Eqn.;

```

data NJ_winterHOA_Cout;
set NJ_winterHOA_tot;
HOA_C1_T_out = 10**-6*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C2_T_out = 10**-5*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C3_T_out = 10**-4*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C4_T_out = 10**-3*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C5_T_out = 10**-2*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C6_T_out = 10**-1*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C7_T_out = 10**0*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C8_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C9_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
HOA_C10_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
run;

```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```

data NJ_winterHOA_Part_out;
set NJ_winterHOA_Cout;
HOA_C1_part_out = (1+((HOA_C1_T_out)/OM_out))**-1;
HOA_C2_part_out = (1+((HOA_C2_T_out)/OM_out))**-1;
HOA_C3_part_out = (1+((HOA_C3_T_out)/OM_out))**-1;
HOA_C4_part_out = (1+((HOA_C4_T_out)/OM_out))**-1;
HOA_C5_part_out = (1+((HOA_C5_T_out)/OM_out))**-1;
HOA_C6_part_out = (1+((HOA_C6_T_out)/OM_out))**-1;
HOA_C7_part_out = (1+((HOA_C7_T_out)/OM_out))**-1;
HOA_C8_part_out = (1+((HOA_C8_T_out)/OM_out))**-1;
HOA_C9_part_out = (1+((HOA_C9_T_out)/OM_out))**-1;
HOA_C10_part_out = (1+((HOA_C10_T_out)/OM_out))**-1;
HOA_Part_tot_out =
(HOA_C1_part_out*HOA_C1_tot+HOA_C2_part_out*HOA_C2_tot+HOA_C3_part_out*
HOA_C3_tot+HOA_C4_part_out*HOA_C4_tot+HOA_C5_part_out*HOA_C5_tot+HOA_C6
_part_out*HOA_C6_tot
+HOA_C7_part_out*HOA_C7_tot+HOA_C8_part_out*HOA_C8_tot+HOA_C9_part_out*
HOA_C9_tot+HOA_C10_part_out*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;

```

```

*Calculate how C*(Tref = 298.15 K) shifts based on temperature indoors
using Clausius Clapyeron Eqn.;
data NJ_winterHOA_Cin;
set NJ_winterHOA_tot;
HOA_C1_T_in = 10**-6*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C2_T_in = 10**-5*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C3_T_in = 10**-4*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C4_T_in = 10**-3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C5_T_in = 10**-2*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C6_T_in = 10**-1*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C7_T_in = 10**0*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15)))));
HOA_C8_T_in = 10**1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15)))));
HOA_C9_T_in = 10**2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15)))));
HOA_C10_T_in = 10**3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data NJ_winterHOA_Part_in;
set NJ_winterHOA_Cin;
HOA_C1_part_in = (1+((HOA_C1_T_in)/OM_in))**-1;
HOA_C2_part_in = (1+((HOA_C2_T_in)/OM_in))**-1;
HOA_C3_part_in = (1+((HOA_C3_T_in)/OM_in))**-1;
HOA_C4_part_in = (1+((HOA_C4_T_in)/OM_in))**-1;
HOA_C5_part_in = (1+((HOA_C5_T_in)/OM_in))**-1;
HOA_C6_part_in = (1+((HOA_C6_T_in)/OM_in))**-1;
HOA_C7_part_in = (1+((HOA_C7_T_in)/OM_in))**-1;
HOA_C8_part_in = (1+((HOA_C8_T_in)/OM_in))**-1;
HOA_C9_part_in = (1+((HOA_C9_T_in)/OM_in))**-1;
HOA_C10_part_in = (1+((HOA_C10_T_in)/OM_in))**-1;
HOA_Part_tot_in =
(HOA_C1_part_in*HOA_C1_tot+HOA_C2_part_in*HOA_C2_tot+HOA_C3_part_in*HOA
_C3_tot+HOA_C4_part_in*HOA_C4_tot+HOA_C5_part_in*HOA_C5_tot+HOA_C6_part
_in*HOA_C6_tot
+HOA_C7_part_in*HOA_C7_tot+HOA_C8_part_in*HOA_C8_tot+HOA_C9_part_in*HOA
_C9_tot+HOA_C10_part_in*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;

*Merge indoor and outdoor partitioning data sets and keep only key
data;
proc sort data = NJ_winterHOA_Part_out; by HomeID;run;
proc sort data = NJ_winterHOA_Part_in; by HomeID;run;

data NJwinter_HOA_part;

```

```

merge NJ_winterHOA_Part_out NJ_winterHOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in HOA_Part_tot_out
HOA_part_tot_in;
run;

*Add column with calculation of indoor COA for HOA based on
partitioning;
data NJwinter_HOA_part2;
set NJwinter_HOA_part;
HOA_Ctot = 0.458*OM_out/HOA_Part_tot_out; *the 0.458 is the mass
fraction total OA that is HOA from the SI of Cappa and Jimenez 2010;
HOA_COA_out = 0.458*OM_out;
HOA_COA_in = HOA_Ctot*HOA_Part_tot_in;
run;

*****
OOA
*****

*calculate total (gas+particle) organic concentration in each OOA bin
using relationship between C* and Ctot from Cappa and Jimenez 2010;
data NJ_winterOOAtot;
set NJ_winter;
OOA_C1_tot = 1.94+5.5*exp(0.8*(log10(10**-6)-3));
OOA_C2_tot = 1.94+5.5*exp(0.8*(log10(10**-5)-3));
OOA_C3_tot = 1.94+5.5*exp(0.8*(log10(10**-4)-3));
OOA_C4_tot = 1.94+5.5*exp(0.8*(log10(10**-3)-3));
OOA_C5_tot = 1.94+5.5*exp(0.8*(log10(10**-2)-3));
OOA_C6_tot = 1.94+5.5*exp(0.8*(log10(10**-1)-3));
OOA_C7_tot = 1.94+5.5*exp(0.8*(log10(10**0)-3));
OOA_C8_tot = 1.94+5.5*exp(0.8*(log10(10**1)-3));
OOA_C9_tot = 1.94+5.5*exp(0.8*(log10(10**2)-3));
OOA_C10_tot = 1.94+5.5*exp(0.8*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
data NJ_winterOOA_Cout;
set NJ_winterOOAtot;
OOA_C1_T_out = 10**-6*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C2_T_out = 10**-5*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C3_T_out = 10**-4*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C4_T_out = 10**-3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C5_T_out = 10**-2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C6_T_out = 10**-1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C7_T_out = 10**-0*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));

```

```

OOA_C8_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
OOA_C9_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
OOA_C10_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data NJ_winterOOA_Part_out;
set NJ_winterOOA_Cout;
OOA_C1_part_out = (1+((OOA_C1_T_out)/OM_out))**-1;
OOA_C2_part_out = (1+((OOA_C2_T_out)/OM_out))**-1;
OOA_C3_part_out = (1+((OOA_C3_T_out)/OM_out))**-1;
OOA_C4_part_out = (1+((OOA_C4_T_out)/OM_out))**-1;
OOA_C5_part_out = (1+((OOA_C5_T_out)/OM_out))**-1;
OOA_C6_part_out = (1+((OOA_C6_T_out)/OM_out))**-1;
OOA_C7_part_out = (1+((OOA_C7_T_out)/OM_out))**-1;
OOA_C8_part_out = (1+((OOA_C8_T_out)/OM_out))**-1;
OOA_C9_part_out = (1+((OOA_C9_T_out)/OM_out))**-1;
OOA_C10_part_out = (1+((OOA_C10_T_out)/OM_out))**-1;
OOA_Part_tot_out =
(OOA_C1_part_out*OOA_C1_tot+OOA_C2_part_out*OOA_C2_tot+OOA_C3_part_out*
OOA_C3_tot+OOA_C4_part_out*OOA_C4_tot+OOA_C5_part_out*OOA_C5_tot+OOA_C6
_part_out*OOA_C6_tot
+OOA_C7_part_out*OOA_C7_tot+OOA_C8_part_out*OOA_C8_tot+OOA_C9_part_out*
OOA_C9_tot+OOA_C10_part_out*OOA_C10_tot)/
(OOA_C1_tot+OOA_C2_tot+OOA_C3_tot+OOA_C4_tot+OOA_C5_tot+OOA_C6_tot+OOA
_C7_tot+OOA_C8_tot+OOA_C9_tot+OOA_C10_tot);
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature indoors
using Clausius Clapyeron Eqn.;
data NJ_winterOOA_Cin;
set NJ_winterOOA_tot;
OOA_C1_T_in = 10**-6*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C2_T_in = 10**-5*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C3_T_in = 10**-4*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C4_T_in = 10**-3*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C5_T_in = 10**-2*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C6_T_in = 10**-1*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C7_T_in = 10**0*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C8_T_in = 10**1*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C9_T_in = 10**2*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C10_T_in = 10**3*(298.15/Temp_in)*exp(((100000/8.314)*(1/Temp_in)-(1/298.15))));
run;

```



```

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data NJ_winterOOA_Part_in;
set NJ_winterOOA_Cin;
OOA_C1_part_in = (1+((OOA_C1_T_in)/OM_in))**-1;
OOA_C2_part_in = (1+((OOA_C2_T_in)/OM_in))**-1;
OOA_C3_part_in = (1+((OOA_C3_T_in)/OM_in))**-1;
OOA_C4_part_in = (1+((OOA_C4_T_in)/OM_in))**-1;
OOA_C5_part_in = (1+((OOA_C5_T_in)/OM_in))**-1;
OOA_C6_part_in = (1+((OOA_C6_T_in)/OM_in))**-1;
OOA_C7_part_in = (1+((OOA_C7_T_in)/OM_in))**-1;
OOA_C8_part_in = (1+((OOA_C8_T_in)/OM_in))**-1;
OOA_C9_part_in = (1+((OOA_C9_T_in)/OM_in))**-1;
OOA_C10_part_in = (1+((OOA_C10_T_in)/OM_in))**-1;
OOA_Part_tot_in =
(OOA_C1_part_in*OOA_C1_tot+OOA_C2_part_in*OOA_C2_tot+OOA_C3_part_in*OOA_
_C3_tot+OOA_C4_part_in*OOA_C4_tot+OOA_C5_part_in*OOA_C5_tot+OOA_C6_part
_in*OOA_C6_tot
+OOA_C7_part_in*OOA_C7_tot+OOA_C8_part_in*OOA_C8_tot+OOA_C9_part_in*OOA_
_C9_tot+OOA_C10_part_in*OOA_C10_tot)/
(OOA_C1_tot+OOA_C2_tot+OOA_C3_tot+OOA_C4_tot+OOA_C5_tot+OOA_C6_tot+OOA_
C7_tot+OOA_C8_tot+OOA_C9_tot+OOA_C10_tot);
run;

*Merge indoor and outdoor Partitioning data sets and keep only key
data;
proc sort data = NJ_winterOOA_Part_out; by HomeID; run;
proc sort data = NJ_winterOOA_Part_in; by HomeID; run;

data NJwinter_OOA_part;
merge NJ_winterOOA_Part_out NJ_winterOOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in OOA_Part_tot_out
OOA_part_tot_in;
run;

*Add column with calcuation of indoor COA for OOA based on
partitioning;
data NJwinter_OOA_part2;
set NJwinter_OOA_part;
OOA_Ctot = 0.542*OM_out/OOA_Part_tot_out; *the 0.542 is the mass
fraction total OA that is OOA from the SI of Cappa and Jimenez 2010;
OOA_COA_out = 0.542*OM_out;
OOA_COA_in = OOA_Ctot*OOA_Part_tot_in;
run;

*Merge the HOA and OOA data sets;
data NJ_winter_Part;
merge NJwinter_HOA_part2 NJwinter_OOA_part2;
by HomeID;
Calc_OA_Out = HOA_COA_Out + OOA_COA_Out;
Calc_OA_in = HOA_COA_in + OOA_COA_in;
run;

*This file was exported to library "Ch5" and is named "OA components by
season.xls," sheet "NJ_Winter";

```

```

*Houston, TX RIOPA homes using Houston OA component mass fractions
detailed in Jimenez et al. 2009;
*Note the components are broken up by summer and winter, so I will
break up the measured data into warm and cold seasons;

*Import data set with seasons and state;
PROC IMPORT OUT= Season
    DATAFILE= "G:\Chapter 5\Season and State.xls"
    DBMS=EXCEL REPLACE;
    SHEET="All";
    GETNAMES=YES;
    MIXED=YES;
    USEDATE=YES;
    SCANTIME=YES;

RUN;

*****
        Summer
*****

*create data set with summer dates only;
data TX_summer;
set Season;
if season NE "s" then delete;
if state NE "TX" then delete;
run;

*****
        HOA
*****

*calculate total (gas+particle) organic concentration in each HOA bin
using relationship between C* and Ctot from Cappa and Jimenez 2010;
data TX_summerHOAtot;
set TX_summer;
HOA_C1_tot = 0.78+23*exp(0.7*(log10(10**-6)-3));
HOA_C2_tot = 0.78+23*exp(0.7*(log10(10**-5)-3));
HOA_C3_tot = 0.78+23*exp(0.7*(log10(10**-4)-3));
HOA_C4_tot = 0.78+23*exp(0.7*(log10(10**-3)-3));
HOA_C5_tot = 0.78+23*exp(0.7*(log10(10**-2)-3));
HOA_C6_tot = 0.78+23*exp(0.7*(log10(10**-1)-3));
HOA_C7_tot = 0.78+23*exp(0.7*(log10(10**0)-3));
HOA_C8_tot = 0.78+23*exp(0.7*(log10(10**1)-3));
HOA_C9_tot = 0.78+23*exp(0.7*(log10(10**2)-3));
HOA_C10_tot = 0.78+23*exp(0.7*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
data TX_summerHOA_Cout;
set TX_summerHOAtot;
HOA_C1_T_out = 10**-6*(298.15/Temp_out)*exp(((100000/8.314)*(1/Temp_out)-(1/298.15))));

```

```

HOA_C2_T_out = 10**-5*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C3_T_out = 10**-4*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C4_T_out = 10**-3*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C5_T_out = 10**-2*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C6_T_out = 10**-1*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C7_T_out = 10**0*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C8_T_out = 10**1*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C9_T_out = 10**2*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
HOA_C10_T_out = 10**3*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data TX_summerHOA_Part_out;
set TX_summerHOA_Cout;
HOA_C1_part_out = (1+((HOA_C1_T_out)/OM_out))**-1;
HOA_C2_part_out = (1+((HOA_C2_T_out)/OM_out))**-1;
HOA_C3_part_out = (1+((HOA_C3_T_out)/OM_out))**-1;
HOA_C4_part_out = (1+((HOA_C4_T_out)/OM_out))**-1;
HOA_C5_part_out = (1+((HOA_C5_T_out)/OM_out))**-1;
HOA_C6_part_out = (1+((HOA_C6_T_out)/OM_out))**-1;
HOA_C7_part_out = (1+((HOA_C7_T_out)/OM_out))**-1;
HOA_C8_part_out = (1+((HOA_C8_T_out)/OM_out))**-1;
HOA_C9_part_out = (1+((HOA_C9_T_out)/OM_out))**-1;
HOA_C10_part_out = (1+((HOA_C10_T_out)/OM_out))**-1;
HOA_Part_tot_out =
(HOA_C1_part_out*HOA_C1_tot+HOA_C2_part_out*HOA_C2_tot+HOA_C3_part_out*
HOA_C3_tot+HOA_C4_part_out*HOA_C4_tot+HOA_C5_part_out*HOA_C5_tot+HOA_C6
_part_out*HOA_C6_tot
+HOA_C7_part_out*HOA_C7_tot+HOA_C8_part_out*HOA_C8_tot+HOA_C9_part_out*
HOA_C9_tot+HOA_C10_part_out*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature indoors
using Clausius Clapyeron Eqn.;
data TX_summerHOA_Cin;
set TX_summerHOAtot;
HOA_C1_T_in = 10**-6*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C2_T_in = 10**-5*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C3_T_in = 10**-4*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
HOA_C4_T_in = 10**-3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));

```

```

HOA_C5_T_in = 10**-2*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
HOA_C6_T_in = 10**-1*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
HOA_C7_T_in = 10**0*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15))));
HOA_C8_T_in = 10**1*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15))));
HOA_C9_T_in = 10**2*(298.15/Temp_in)*exp(((-100000/8.314)*(1/Temp_in)-
(1/298.15))));
HOA_C10_T_in = 10**3*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15))));
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data TX_summerHOA_Part_in;
set TX_summerHOA_Cin;
HOA_C1_part_in = (1+((HOA_C1_T_in)/OM_in))**-1;
HOA_C2_part_in = (1+((HOA_C2_T_in)/OM_in))**-1;
HOA_C3_part_in = (1+((HOA_C3_T_in)/OM_in))**-1;
HOA_C4_part_in = (1+((HOA_C4_T_in)/OM_in))**-1;
HOA_C5_part_in = (1+((HOA_C5_T_in)/OM_in))**-1;
HOA_C6_part_in = (1+((HOA_C6_T_in)/OM_in))**-1;
HOA_C7_part_in = (1+((HOA_C7_T_in)/OM_in))**-1;
HOA_C8_part_in = (1+((HOA_C8_T_in)/OM_in))**-1;
HOA_C9_part_in = (1+((HOA_C9_T_in)/OM_in))**-1;
HOA_C10_part_in = (1+((HOA_C10_T_in)/OM_in))**-1;
HOA_Part_tot_in =
(HOA_C1_part_in*HOA_C1_tot+HOA_C2_part_in*HOA_C2_tot+HOA_C3_part_in*HOA_
_C3_tot+HOA_C4_part_in*HOA_C4_tot+HOA_C5_part_in*HOA_C5_tot+HOA_C6_part_
_in*HOA_C6_tot
+HOA_C7_part_in*HOA_C7_tot+HOA_C8_part_in*HOA_C8_tot+HOA_C9_part_in*HOA_
_C9_tot+HOA_C10_part_in*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA_
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;

*Merge indoor and outdoor partitioning data sets and keep only key
data;
proc sort data = TX_summerHOA_Part_out; by HomeID;run;
proc sort data = TX_summerHOA_Part_in; by HomeID;run;

data TXsummer_HOA_part;
merge TX_summerHOA_Part_out TX_summerHOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in HOA_Part_tot_out
HOA_part_tot_in;
run;

*Add column with calculation of indoor COA for HOA based on
partitioning;
data TXsummer_HOA_part2;
set TXsummer_HOA_part;
HOA_Ctot = 0.117*OM_out/HOA_Part_tot_out; *the 0.117 is the mass
fraction total OA that is HOA from the SI of Cappa and Jimenez 2010;
HOA_COA_out = 0.117*OM_out;

```

```

HOA_COA_in = HOA_Ctot*HOA_Part_tot_in;
run;

*****
OOA
*****

*calculate total (gas+particle) organic concentration in each OOA bin
using relationship between C* and Ctot from Cappa and Jimenez 2010;
data TX_summerOOA_tot;
set TX_summer;
OOA_C1_tot = 1.94+5.5*exp(0.8*(log10(10**-6)-3));
OOA_C2_tot = 1.94+5.5*exp(0.8*(log10(10**-5)-3));
OOA_C3_tot = 1.94+5.5*exp(0.8*(log10(10**-4)-3));
OOA_C4_tot = 1.94+5.5*exp(0.8*(log10(10**-3)-3));
OOA_C5_tot = 1.94+5.5*exp(0.8*(log10(10**-2)-3));
OOA_C6_tot = 1.94+5.5*exp(0.8*(log10(10**-1)-3));
OOA_C7_tot = 1.94+5.5*exp(0.8*(log10(10**0)-3));
OOA_C8_tot = 1.94+5.5*exp(0.8*(log10(10**1)-3));
OOA_C9_tot = 1.94+5.5*exp(0.8*(log10(10**2)-3));
OOA_C10_tot = 1.94+5.5*exp(0.8*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
data TX_summerOOA_Cout;
set TX_summerOOA_tot;
OOA_C1_T_out = 10**-6*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C2_T_out = 10**-5*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C3_T_out = 10**-4*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C4_T_out = 10**-3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C5_T_out = 10**-2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C6_T_out = 10**-1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C7_T_out = 10**0*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C8_T_out = 10**1*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C9_T_out = 10**2*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
OOA_C10_T_out = 10**3*(298.15/Temp_out)*exp(((100000/8.314)*((1/Temp_out)-(1/298.15))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data TX_summerOOA_Part_out;
set TX_summerOOA_Cout;
OOA_C1_part_out = (1+((OOA_C1_T_out)/OM_out))**-1;
OOA_C2_part_out = (1+((OOA_C2_T_out)/OM_out))**-1;
OOA_C3_part_out = (1+((OOA_C3_T_out)/OM_out))**-1;
OOA_C4_part_out = (1+((OOA_C4_T_out)/OM_out))**-1;

```

```

OOA_C5_part_out = (1+((OOA_C5_T_out)/OM_out))** -1;
OOA_C6_part_out = (1+((OOA_C6_T_out)/OM_out))** -1;
OOA_C7_part_out = (1+((OOA_C7_T_out)/OM_out))** -1;
OOA_C8_part_out = (1+((OOA_C8_T_out)/OM_out))** -1;
OOA_C9_part_out = (1+((OOA_C9_T_out)/OM_out))** -1;
OOA_C10_part_out = (1+((OOA_C10_T_out)/OM_out))** -1;
OOA_Part_tot_out =
(OOA_C1_part_out*OOA_C1_tot+OOA_C2_part_out*OOA_C2_tot+OOA_C3_part_out*
OOA_C3_tot+OOA_C4_part_out*OOA_C4_tot+OOA_C5_part_out*OOA_C5_tot+OOA_C6
_part_out*OOA_C6_tot
+OOA_C7_part_out*OOA_C7_tot+OOA_C8_part_out*OOA_C8_tot+OOA_C9_part_out*
OOA_C9_tot+OOA_C10_part_out*OOA_C10_tot)/
(OOA_C1_tot+OOA_C2_tot+OOA_C3_tot+OOA_C4_tot+OOA_C5_tot+OOA_C6_tot+OOA_
C7_tot+OOA_C8_tot+OOA_C9_tot+OOA_C10_tot);
run;

```

Calculate how C(Tref = 298.15 K) shifts based on temperature indoors using Clausius Clapyeron Eqn.;

```

data TX_summerOOA_Cin;
set TX_summerOOA_tot;
OOA_C1_T_in = 10** -6*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C2_T_in = 10** -5*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C3_T_in = 10** -4*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C4_T_in = 10** -3*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C5_T_in = 10** -2*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C6_T_in = 10** -1*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
OOA_C7_T_in = 10** 0*(298.15/Temp_in)*exp((( -100000/8.314)*(1/Temp_in)-
(1/298.15))));
OOA_C8_T_in = 10** 1*(298.15/Temp_in)*exp((( -100000/8.314)*(1/Temp_in)-
(1/298.15))));
OOA_C9_T_in = 10** 2*(298.15/Temp_in)*exp((( -100000/8.314)*(1/Temp_in)-
(1/298.15))));
OOA_C10_T_in = 10** 3*(298.15/Temp_in)*exp((( -
100000/8.314)*(1/Temp_in)-(1/298.15))));
run;

```

*calculate the indoor partitioning based on VBS from Cappa and Jimenez 2010;

```

data TX_summerOOA_Part_in;
set TX_summerOOA_Cin;
OOA_C1_part_in = (1+((OOA_C1_T_in)/OM_in))** -1;
OOA_C2_part_in = (1+((OOA_C2_T_in)/OM_in))** -1;
OOA_C3_part_in = (1+((OOA_C3_T_in)/OM_in))** -1;
OOA_C4_part_in = (1+((OOA_C4_T_in)/OM_in))** -1;
OOA_C5_part_in = (1+((OOA_C5_T_in)/OM_in))** -1;
OOA_C6_part_in = (1+((OOA_C6_T_in)/OM_in))** -1;
OOA_C7_part_in = (1+((OOA_C7_T_in)/OM_in))** -1;
OOA_C8_part_in = (1+((OOA_C8_T_in)/OM_in))** -1;
OOA_C9_part_in = (1+((OOA_C9_T_in)/OM_in))** -1;
OOA_C10_part_in = (1+((OOA_C10_T_in)/OM_in))** -1;

```

```

OOA_Part_tot_in =
(OOA_C1_part_in*OOA_C1_tot+OOA_C2_part_in*OOA_C2_tot+OOA_C3_part_in*OOA
_C3_tot+OOA_C4_part_in*OOA_C4_tot+OOA_C5_part_in*OOA_C5_tot+OOA_C6_part
_in*OOA_C6_tot
+OOA_C7_part_in*OOA_C7_tot+OOA_C8_part_in*OOA_C8_tot+OOA_C9_part_in*OOA
_C9_tot+OOA_C10_part_in*OOA_C10_tot)/
(OOA_C1_tot+OOA_C2_tot+OOA_C3_tot+OOA_C4_tot+OOA_C5_tot+OOA_C6_tot+OOA
_C7_tot+OOA_C8_tot+OOA_C9_tot+OOA_C10_tot);
run;

*Merge indoor and outdoor Partitioning data sets and keep only key
data;
proc sort data = TX_summerOOA_Part_out; by HomeID;run;
proc sort data = TX_summerOOA_Part_in; by HomeID;run;

data TXsummer_OOA_part;
merge TX_summerOOA_Part_out TX_summerOOA_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in OOA_Part_tot_out
OOA_part_tot_in;
run;

*Add column with calculation of indoor COA for OOA based on
partitioning;
data TXsummer_OOA_part2;
set TXsummer_OOA_part;
OOA_Ctot = 0.450*OM_out/OOA_Part_tot_out; *the 0.450 is the mass
fraction total OA that is OOA from the SI of Cappa and Jimenez 2010;
OOA_COA_out = 0.450*OM_out;
OOA_COA_in = OOA_Ctot*OOA_Part_tot_in;
run;

*****
other OA (OTH)
*****

*calculate total (gas+particle) organic concentration in each OTH bin
using relationship between C* and Ctot from Cappa and Jimenez 2010
using the values for OA;
data TX_summerOTHtot;
set TX_summer;
OTH_C1_tot = 1.65+19*exp(1*(log10(10**-6)-3));
OTH_C2_tot = 1.65+19*exp(1*(log10(10**-5)-3));
OTH_C3_tot = 1.65+19*exp(1*(log10(10**-4)-3));
OTH_C4_tot = 1.65+19*exp(1*(log10(10**-3)-3));
OTH_C5_tot = 1.65+19*exp(1*(log10(10**-2)-3));
OTH_C6_tot = 1.65+19*exp(1*(log10(10**-1)-3));
OTH_C7_tot = 1.65+19*exp(1*(log10(10**0)-3));
OTH_C8_tot = 1.65+19*exp(1*(log10(10**1)-3));
OTH_C9_tot = 1.65+19*exp(1*(log10(10**2)-3));
OTH_C10_tot = 1.65+19*exp(1*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
data TX_summerOTH_Cout;
set TX_summerOTHtot;

```

```

OTH_C1_T_out = 10**-6*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C2_T_out = 10**-5*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C3_T_out = 10**-4*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C4_T_out = 10**-3*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C5_T_out = 10**-2*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C6_T_out = 10**-1*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C7_T_out = 10**0*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C8_T_out = 10**1*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C9_T_out = 10**2*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
OTH_C10_T_out = 10**3*(298.15/Temp_out)*exp(((-
100000/8.314)*(1/Temp_out)-(1/298.15)))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez
2010;
data TX_summerOTH_Part_out;
set TX_summerOTH_Cout;
OTH_C1_part_out = (1+((OTH_C1_T_out)/OM_out))**-1;
OTH_C2_part_out = (1+((OTH_C2_T_out)/OM_out))**-1;
OTH_C3_part_out = (1+((OTH_C3_T_out)/OM_out))**-1;
OTH_C4_part_out = (1+((OTH_C4_T_out)/OM_out))**-1;
OTH_C5_part_out = (1+((OTH_C5_T_out)/OM_out))**-1;
OTH_C6_part_out = (1+((OTH_C6_T_out)/OM_out))**-1;
OTH_C7_part_out = (1+((OTH_C7_T_out)/OM_out))**-1;
OTH_C8_part_out = (1+((OTH_C8_T_out)/OM_out))**-1;
OTH_C9_part_out = (1+((OTH_C9_T_out)/OM_out))**-1;
OTH_C10_part_out = (1+((OTH_C10_T_out)/OM_out))**-1;
OTH_Part_tot_out =
(OTH_C1_part_out*OTH_C1_tot+OTH_C2_part_out*OTH_C2_tot+OTH_C3_part_out*
OTH_C3_tot+OTH_C4_part_out*OTH_C4_tot+OTH_C5_part_out*OTH_C5_tot+OTH_C6
_part_out*OTH_C6_tot
+OTH_C7_part_out*OTH_C7_tot+OTH_C8_part_out*OTH_C8_tot+OTH_C9_part_out*
OTH_C9_tot+OTH_C10_part_out*OTH_C10_tot)/
(OTH_C1_tot+OTH_C2_tot+OTH_C3_tot+OTH_C4_tot+OTH_C5_tot+OTH_C6_tot+OTH_
C7_tot+OTH_C8_tot+OTH_C9_tot+OTH_C10_tot);
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature indoors
using Clausius Clapyeron Eqn.;
data TX_summerOTH_Cin;
set TX_summerOTHtot;
OTH_C1_T_in = 10**-6*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
OTH_C2_T_in = 10**-5*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));
OTH_C3_T_in = 10**-4*(298.15/Temp_in)*exp(((-
100000/8.314)*(1/Temp_in)-(1/298.15)))));

```



```

OTH_C4_T_in = 10**3*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C5_T_in = 10**2*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C6_T_in = 10**1*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C7_T_in = 10**0*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C8_T_in = 10**1*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C9_T_in = 10**2*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
OTH_C10_T_in = 10**3*(298.15/Temp_in)*exp(((100000/8.314)*((1/Temp_in)-(1/298.15))));
run;

*calculate the indoor partitioning based on VBS from Cappa and Jimenez
2010;
data TX_summerOTH_Part_in;
set TX_summerOTH_Cin;
OTH_C1_part_in = (1+((OTH_C1_T_in)/OM_in))**-1;
OTH_C2_part_in = (1+((OTH_C2_T_in)/OM_in))**-1;
OTH_C3_part_in = (1+((OTH_C3_T_in)/OM_in))**-1;
OTH_C4_part_in = (1+((OTH_C4_T_in)/OM_in))**-1;
OTH_C5_part_in = (1+((OTH_C5_T_in)/OM_in))**-1;
OTH_C6_part_in = (1+((OTH_C6_T_in)/OM_in))**-1;
OTH_C7_part_in = (1+((OTH_C7_T_in)/OM_in))**-1;
OTH_C8_part_in = (1+((OTH_C8_T_in)/OM_in))**-1;
OTH_C9_part_in = (1+((OTH_C9_T_in)/OM_in))**-1;
OTH_C10_part_in = (1+((OTH_C10_T_in)/OM_in))**-1;
OTH_Part_tot_in =
(OTH_C1_part_in*OTH_C1_tot+OTH_C2_part_in*OTH_C2_tot+OTH_C3_part_in*OTH
_C3_tot+OTH_C4_part_in*OTH_C4_tot+OTH_C5_part_in*OTH_C5_tot+OTH_C6_part
_in*OTH_C6_tot
+OTH_C7_part_in*OTH_C7_tot+OTH_C8_part_in*OTH_C8_tot+OTH_C9_part_in*OTH
_C9_tot+OTH_C10_part_in*OTH_C10_tot)/
(OTH_C1_tot+OTH_C2_tot+OTH_C3_tot+OTH_C4_tot+OTH_C5_tot+OTH_C6_tot+OTH
_C7_tot+OTH_C8_tot+OTH_C9_tot+OTH_C10_tot);
run;

*Merge indoor and outdoor Partitioning data sets and keep only key
data;
proc sort data = TX_summerOTH_Part_out; by HomeID; run;
proc sort data = TX_summerOTH_Part_in; by HomeID; run;

data TXsummer_OTH_part;
merge TX_summerOTH_Part_out TX_summerOTH_Part_in;
by HomeID;
keep HomeID OM_out OM_in Temp_out Temp_in OTH_Part_tot_out
OTH_part_tot_in;
run;

*Add column with calculation of indoor COA for OTH based on
partitioning;
data TXsummer_OTH_part2;
set TXsummer_OTH_part;

```

```

OTH_Ctot = 0.433*OM_out/OTH_Part_tot_out; *the 0.433 is the mass
fraction total OA that is OTH from the SI of Cappa and Jimenez 2010;
OTH_COA_out = 0.433*OM_out;
OTH_COA_in = OTH_Ctot*OTH_Part_tot_in;
run;

*Merge the HOA, OOA, and OTH data sets;
data TX_summer_Part;
merge TXsummer_HOA_part2 TXsummer_OOA_part2 TXsummer_OTH_part2;
by HomeID;
Calc_OA_Out = HOA_COA_Out + OOA_COA_Out+OTH_COA_Out;
Calc_OA_in = HOA_COA_in + OOA_COA_in + OTH_COA_in;
run;

*This file was exported to library "Ch5" and is named "OA components by
season.xls," sheet "TX";

```

Appendix D6

SAS Code for Chapter 5 Calculations: Multiple Linear Regression Analyses

*Program to explore whether changes in OA concentrations (delta_COA) due to shifts in partitioning are driven predominantly by indoor-outdoor temperature differences (delta_T) or indoor-outdoor differences in OA loading (delta_OAload) using multiple linear regression (MLR);

```
libname Ch5 'C:\Chapter 5';
```

```
*****
Delta_Hvap = 50 kJ/mol
*****
```

*read in the data set with calculated partitioning and variables to be included in MLR;

```
PROC IMPORT OUT= Hvap50
  DATAFILE= "G:\Chapter 5\Hvap50 for MLR.xls"
  DBMS=EXCEL REPLACE;
  SHEET="All";
  GETNAMES=YES;
  MIXED=YES;
  USEDATE=YES;
  SCANTIME=YES;
```

```
RUN;
```

*Identify Collinearities and outliers;

```
proc reg data = Hvap50;
model delta_COA = delta_T delta_OMload/
VIF COLLINOINT;
output out = ttest_50 /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;
```

*for collinearity variance inflation>10, variance proportion >0.5 and condition indices>3.6;

*no collinearity problems detected. Remove outliers;

```
data Hvap50_2;
set ttest_50;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;
```

*using this clean data set (no outliers) create the MLR model;

```
proc reg data = Hvap50_2;
model delta_COA = delta_T delta_OMload/selection =stepwise;
ods output selparmest = Ch5.selparmest_50 selectionsummary =
Ch5.selectionsummary_50;
run;
quit;
```

```

*****
For each RIOPA region separately
*****;

*LA County;

*read in the data set with calculated partitioning and variables to be
included in MLR;
PROC IMPORT OUT= Hvap50_CA
    DATAFILE= "G:\Chapter 5\Hvap50 for MLR.xls"
    DBMS=EXCEL REPLACE;
    SHEET="CA";
    GETNAMES=YES;
    MIXED=YES;
    USEDATE=YES;
    SCANTIME=YES;

RUN;

*Identify Collinearities and outliers;
proc reg data = Hvap50_CA;
model delta_COA = delta_T delta_OMload/
VIF COLLINOINT;
output out = ttest_50_CA /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;

*for collinearity variance inflation>10, variance proportion >0.5 and
condition indices>3.6;
*no collinearity problems detected. Remove outliers;

data Hvap50_CA2;
set ttest_50_CA;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;

*using this clean data set (no outliers) create the MLR model;
proc reg data = Hvap50_CA2;
model delta_COA = delta_T delta_OMload/selection =stepwise;
ods output selparmest = Ch5.selparmest_50_CA selectionsummary =
Ch5.selectionsummary_50_CA;
run;
quit;

*Elizabeth

*read in the data set with calculated partitioning and variables to be
included in MLR;
PROC IMPORT OUT= Hvap50_NJ
    DATAFILE= "G:\Chapter 5\Hvap50 for MLR.xls"
    DBMS=EXCEL REPLACE;
    SHEET="NJ";
    GETNAMES=YES;

```

```

        MIXED=YES;
        USEDATE=YES;
        SCANTIME=YES;

RUN;

*Identify Collinearities and outliers;
proc reg data = Hvap50_NJ;
model delta_COA = delta_T delta_OMload/
VIF COLLINOINT;
output out = ttest_50_NJ /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;

*for collinearity variance inflation>10, variance proportion >0.5 and
condition indices>3.6;
*no collinearity problems detected. Remove outliers;
data Hvap50_NJ2;
set ttest_50_NJ;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;

*using this clean data set (no outliers) create the MLR model;
proc reg data = Hvap50_NJ2;
model delta_COA = delta_T delta_OMload/selection =stepwise;
ods output selparmest = Ch5.selparmest_50_NJ selectionsummary =
Ch5.selectionsummary_50_NJ;
run;
quit;

*Houston

*read in the data set with calculated partitioning and variables to be
included in MLR;
PROC IMPORT OUT= Hvap50_TX
        DATAFILE= "G:\Chapter 5\Hvap50 for MLR.xls"
        DBMS=EXCEL REPLACE;
        SHEET="TX";
        GETNAMES=YES;
        MIXED=YES;
        USEDATE=YES;
        SCANTIME=YES;

RUN;

*Identify Collinearities and outliers;
proc reg data = Hvap50_TX;
model delta_COA = delta_T delta_OMload/
VIF COLLINOINT;
output out = ttest_50_TX /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;

```

```

*for collinearity variance inflation>10, variance proportion >0.5 and
condition indices>3.6;
*no collinearity problems detected. Remove outliers;
data Hvap50_TX2;
set ttest_50_TX;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;

*using this clean data set (no outliers) create the MLR model;
proc reg data = Hvap50_TX2;
model delta_COA = delta_T delta_OMload/selection =stepwise;
ods output selparmest = Ch5.selparmest_50_TX selectionsummary =
Ch5.selectionsummary_50_TX;
run;
quit;

*****
Delta_Hvap = 100 kJ/mol
*****

*read in the data set with calculated partitioning and variables to be
included in MLR;
PROC IMPORT OUT= Hvap100
        DATAFILE= "G:\Chapter 5\Hvap100 for MLR.xls"
        DBMS=EXCEL REPLACE;
        SHEET="All";
        GETNAMES=YES;
        MIXED=YES;
        USEDATE=YES;
        SCANTIME=YES;

RUN;

*Identify Collinearities and outliers;
proc reg data = Hvap100;
model delta_COA = delta_T delta_OMload/
VIF COLLINOINT;
output out = ttest_100 /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;

*for collinearity variance inflation>10, variance proportion >0.5 and
condition indices>3.6;
*no collinearity problems detected. Remove outliers;
data Hvap100_2;
set ttest_100;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;

*using this clean data set (no outliers) create the MLR model;
proc reg data = Hvap100_2;
model delta_COA = delta_T delta_OMload/selection =stepwise;

```

```

ods output selparmest = Ch5.selparmest_100 selectionsummary =
Ch5.selectionsummary_100;
run;
quit;

*****
For each RIOPA region separately
*****;

*LA County

*read in the data set with calculated partitioning and variables to be
included in MLR;
PROC IMPORT OUT= Hvap100_CA
      DATAFILE= "G:\Chapter 5\Hvap100 for MLR.xls"
      DBMS=EXCEL REPLACE;
      SHEET="CA";
      GETNAMES=YES;
      MIXED=YES;
      USEDATE=YES;
      SCANTIME=YES;

RUN;

*Identify Collinearities and outliers;
proc reg data = Hvap100_CA;
model delta_COA = delta_T delta_OMload/
VIF COLLINOINT;
output out = ttest_100_CA /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;

*for collinearity variance inflation>10, variance proportion >0.5 and
condition indices>3.6;
*no collinearity problems detected. Remove outliers;

data Hvap100_CA2;
set ttest_100_CA;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;

*using this clean data set (no outliers) create the MLR model;
proc reg data = Hvap100_CA2;
model delta_COA = delta_T delta_OMload/selection =stepwise;
ods output selparmest = Ch5.selparmest_100_CA selectionsummary =
Ch5.selectionsummary_100_CA;
run;
quit;

*Elizabeth

*read in the data set with calculated partitioning and variables to be
included in MLR;
PROC IMPORT OUT= Hvap100_NJ

```

```

        DATAFILE= "G:\Chapter 5\Hvap100 for MLR.xls"
        DBMS=EXCEL REPLACE;
        SHEET="NJ";
        GETNAMES=YES;
        MIXED=YES;
        USEDATE=YES;
        SCANTIME=YES;

RUN;

*Identify Collinearities and outliers;
proc reg data = Hvp100_NJ;
model delta_COA = delta_T delta_OMload/
VIF COLLINOINT;
output out = ttest_100_NJ /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/
rstudent = rstudent;
run;
quit;

*for collinearity variance inflation>10, variance proportion >0.5 and
condition indices>3.6;
*no collinearity problems detected. Remove outliers;
data Hvp100_NJ2;
set ttest_100_NJ;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;

*using this clean data set (no outliers) create the MLR model;
proc reg data = Hvp100_NJ2;
model delta_COA = delta_T delta_OMload/selection =stepwise;
ods output selparmest = Ch5.selparmest_100_NJ selectionsummary =
Ch5.selectionsummary_100_NJ;
run;
quit;

*Houston

*read in the data set with calculated partitioning and variables to be
included in MLR;
PROC IMPORT OUT= Hvp100_TX
        DATAFILE= "G:\Chapter 5\Hvap100 for MLR.xls"
        DBMS=EXCEL REPLACE;
        SHEET="TX";
        GETNAMES=YES;
        MIXED=YES;
        USEDATE=YES;
        SCANTIME=YES;

RUN;

*Identify Collinearities and outliers;
proc reg data = Hvp100_TX;
model delta_COA = delta_T delta_OMload/
VIF COLLINOINT;
output out = ttest_100_TX /*if >= +/- 2 then an outlier*/
student = student /*if >= +/- 2 then an outlier*/

```



```

rstudent = rstudent;
run;
quit;

*for collinearity variance inflation>10, variance proportion >0.5 and
condition indices>3.6;
*no collinearity problems detected. Remove outliers;
data Hvap100_TX2;
set ttest_100_TX;
if student>=2 or student <=-2 then delete;
drop student rstudent;
run;

*using this clean data set (no outliers) create the MLR model;
proc reg data = Hvap100_TX2;
model delta_COA = delta_T delta_OMload/selection =stepwise;
ods output selparmest = Ch5.selparmest_100_TX selectionsummary =
Ch5.selectionsummary_100_TX;
run;
quit;

```

Appendix D7

SAS Code for Chapter 5 Calculations: Average RIOPA Conditions Volatility Basis Sets

```

*Program to generate VBS for average conditions measured during RIOPA
(outside) for the homes included in Chapter 5 calculations;

*Average conditions:
Temperature = 291.32 K = 18.17 degrees C
Measured outdoor OM = 4.87 ug/m3

*for delta_Hvap = 50 kJ/mol;

*Calculate total (gas+particle) organic concentration in each bin using
relationship
between C* and Ctot from Cappa and Jimenez 2010;

data Hvap_50;
C1_tot = 1.87+44*exp(1*(log10(10**-2)-3));
C2_tot = 1.87+44*exp(1*(log10(10**-1)-3));
C3_tot = 1.87+44*exp(1*(log10(10**0)-3));
C4_tot = 1.87+44*exp(1*(log10(10**1)-3));
C5_tot = 1.87+44*exp(1*(log10(10**2)-3));
C6_tot = 1.87+44*exp(1*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
data Hvap_502;
set Hvap_50;
C1_T_out = 10**-2*(298.15/291.32)*exp((( -50000/8.314)*(1/291.32) -
(1/298.15))));
C2_T_out = 10**-1*(298.15/291.32)*exp((( -50000/8.314)*(1/291.32) -
(1/298.15))));
C3_T_out = 10**-0*(298.15/291.32)*exp((( -50000/8.314)*(1/291.32) -
(1/298.15))));
C4_T_out = 10**1*(298.15/291.32)*exp((( -50000/8.314)*(1/291.32) -
(1/298.15))));
C5_T_out = 10**2*(298.15/291.32)*exp((( -50000/8.314)*(1/291.32) -
(1/298.15))));
C6_T_out = 10**3*(298.15/291.32)*exp((( -50000/8.314)*(1/291.32) -
(1/298.15))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and
Jimenez 2010;

data Hvap_503;
set Hvap_502;
C1_part_out = (1+((C1_T_out)/4.87))**-1;
C2_part_out = (1+((C2_T_out)/4.87))**-1;
C3_part_out = (1+((C3_T_out)/4.87))**-1;
C4_part_out = (1+((C4_T_out)/4.87))**-1;
C5_part_out = (1+((C5_T_out)/4.87))**-1;
C6_part_out = (1+((C6_T_out)/4.87))**-1;

```

```

Part_tot_out =
(C1_part_out*C1_tot+C2_part_out*C2_tot+C3_part_out*C3_tot+C4_part_out*C
4_tot+C5_part_out*C5_tot+C6_part_out*C6_tot)/(C1_tot+C2_tot+C3_tot+C4_t
ot+C5_tot+C6_tot);
run;

*for delta_Hvap = 100 kJ/mol;

*Calculate total (gas+particle) organic concentration in each bin using
relationship
between C* and Ctot from Cappa and Jimenez 2010;

data Hvap100;
C1_tot = 1.65+19*exp(1*(log10(10**-6)-3));
C2_tot = 1.65+19*exp(1*(log10(10**-5)-3));
C3_tot = 1.65+19*exp(1*(log10(10**-4)-3));
C4_tot = 1.65+19*exp(1*(log10(10**-3)-3));
C5_tot = 1.65+19*exp(1*(log10(10**-2)-3));
C6_tot = 1.65+19*exp(1*(log10(10**-1)-3));
C7_tot = 1.65+19*exp(1*(log10(10**0)-3));
C8_tot = 1.65+19*exp(1*(log10(10**1)-3));
C9_tot = 1.65+19*exp(1*(log10(10**2)-3));
C10_tot = 1.65+19*exp(1*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
data Hvap100_2;
set Hvap100;
C1_T_out = 10**-6*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C2_T_out = 10**-5*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C3_T_out = 10**-4*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C4_T_out = 10**-3*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C5_T_out = 10**-2*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C6_T_out = 10**-1*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C7_T_out = 10**-0*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C8_T_out = 10**1*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C9_T_out = 10**2*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
C10_T_out = 10**3*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(
1/298.15))));
run;

*calculate the outdoor partitioning based on VBS from Cappa and
Jimenez 2010;
data Hvap100_3;
set Hvap100_2;

```

```

C1_part_out = (1+((C1_T_out)/4.87))** -1;
C2_part_out = (1+((C2_T_out)/4.87))** -1;
C3_part_out = (1+((C3_T_out)/4.87))** -1;
C4_part_out = (1+((C4_T_out)/4.87))** -1;
C5_part_out = (1+((C5_T_out)/4.87))** -1;
C6_part_out = (1+((C6_T_out)/4.87))** -1;
C7_part_out = (1+((C7_T_out)/4.87))** -1;
C8_part_out = (1+((C8_T_out)/4.87))** -1;
C9_part_out = (1+((C9_T_out)/4.87))** -1;
C10_part_out = (1+((C10_T_out)/4.87))** -1;
Part_tot_out =
(C1_part_out*C1_tot+C2_part_out*C2_tot+C3_part_out*C3_tot+C4_part_out*C
4_tot+C5_part_out*C5_tot+C6_part_out*C6_tot
+C7_part_out*C7_tot+C8_part_out*C8_tot+C9_part_out*C9_tot+C10_part_out*
C10_tot)/
(C1_tot+C2_tot+C3_tot+C4_tot+C5_tot+C6_tot+C7_tot+C8_tot+C9_tot+C10_tot
);
run;

*HOA;
data HOA;
HOA_C1_tot = 0.78+23*exp(0.7*(log10(10**-6)-3));
HOA_C2_tot = 0.78+23*exp(0.7*(log10(10**-5)-3));
HOA_C3_tot = 0.78+23*exp(0.7*(log10(10**-4)-3));
HOA_C4_tot = 0.78+23*exp(0.7*(log10(10**-3)-3));
HOA_C5_tot = 0.78+23*exp(0.7*(log10(10**-2)-3));
HOA_C6_tot = 0.78+23*exp(0.7*(log10(10**-1)-3));
HOA_C7_tot = 0.78+23*exp(0.7*(log10(10**0)-3));
HOA_C8_tot = 0.78+23*exp(0.7*(log10(10**1)-3));
HOA_C9_tot = 0.78+23*exp(0.7*(log10(10**2)-3));
HOA_C10_tot = 0.78+23*exp(0.7*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
data HOA2;
set HOA;
HOA_C1_T_out = 10**-6*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));
HOA_C2_T_out = 10**-5*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));
HOA_C3_T_out = 10**-4*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));
HOA_C4_T_out = 10**-3*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));
HOA_C5_T_out = 10**-2*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));
HOA_C6_T_out = 10**-1*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));
HOA_C7_T_out = 10**-0*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));
HOA_C8_T_out = 10**1*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));
HOA_C9_T_out = 10**2*(298.15/291.32)*exp((( -100000/8.314)*( (1/291.32) -
(1/298.15))));

```

```
HOA_C10_T_out = 10**3*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
run;
```

```
*calculate the outdoor partitioning based on VBS from Cappa and
Jimenez 2010;
```

```
data HOA3;
set HOA2;
HOA_C1_part_out = (1+((HOA_C1_T_out)/4.87))**-1;
HOA_C2_part_out = (1+((HOA_C2_T_out)/4.87))**-1;
HOA_C3_part_out = (1+((HOA_C3_T_out)/4.87))**-1;
HOA_C4_part_out = (1+((HOA_C4_T_out)/4.87))**-1;
HOA_C5_part_out = (1+((HOA_C5_T_out)/4.87))**-1;
HOA_C6_part_out = (1+((HOA_C6_T_out)/4.87))**-1;
HOA_C7_part_out = (1+((HOA_C7_T_out)/4.87))**-1;
HOA_C8_part_out = (1+((HOA_C8_T_out)/4.87))**-1;
HOA_C9_part_out = (1+((HOA_C9_T_out)/4.87))**-1;
HOA_C10_part_out = (1+((HOA_C10_T_out)/4.87))**-1;
HOA_Part_tot_out =
(HOA_C1_part_out*HOA_C1_tot+HOA_C2_part_out*HOA_C2_tot+HOA_C3_part_out*
HOA_C3_tot+HOA_C4_part_out*HOA_C4_tot+HOA_C5_part_out*HOA_C5_tot+HOA_C6
_part_out*HOA_C6_tot
+HOA_C7_part_out*HOA_C7_tot+HOA_C8_part_out*HOA_C8_tot+HOA_C9_part_out*
HOA_C9_tot+HOA_C10_part_out*HOA_C10_tot)/
(HOA_C1_tot+HOA_C2_tot+HOA_C3_tot+HOA_C4_tot+HOA_C5_tot+HOA_C6_tot+HOA
_C7_tot+HOA_C8_tot+HOA_C9_tot+HOA_C10_tot);
run;
```

```
*LV-OOA;
```

```
*calculate total (gas+particle) organic concentration in each LV-OOA
bin using relationship
between C* and Ctot from Cappa and Jimenez 2010;
```

```
data LVOOA;
LVOOA_C1_tot = 0+0.135*exp(-0.37*(log10(10**-7)-3));
LVOOA_C2_tot = 0+0.135*exp(-0.37*(log10(10**-6)-3));
LVOOA_C3_tot = 0+0.135*exp(-0.37*(log10(10**-5)-3));
LVOOA_C4_tot = 0+0.135*exp(-0.37*(log10(10**-4)-3));
LVOOA_C5_tot = 0+0.135*exp(-0.37*(log10(10**-3)-3));
LVOOA_C6_tot = 0+0.135*exp(-0.37*(log10(10**-2)-3));
LVOOA_C7_tot = 0+0.135*exp(-0.37*(log10(10**-1)-3));
LVOOA_C8_tot = 0+0.135*exp(-0.37*(log10(10**0)-3));
LVOOA_C9_tot = 0+0.135*exp(-0.37*(log10(10**1)-3));
LVOOA_C10_tot = 0+0.135*exp(-0.37*(log10(10**2)-3));
LVOOA_C11_tot = 0+0.135*exp(-0.37*(log10(10**3)-3));
run;
```

```
*Calculate how C*(Tref = 298.15 K) shifts outdoors based on temperature
using Clausius Clapyeron Eqn.;
```

```
data LVOOA2;
set LVOOA;
LVOOA_C1_T_out = 10**-7*(298.15/291.32)*exp((( -
100000/8.314)*(1/291.32) - (1/298.15))));
```

```

LVOOA_C2_T_out = 10**-6*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C3_T_out = 10**-5*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C4_T_out = 10**-4*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C5_T_out = 10**-3*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C6_T_out = 10**-2*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C7_T_out = 10**-1*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C8_T_out = 10**0*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C9_T_out = 10**1*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C10_T_out = 10**2*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
LVOOA_C11_T_out = 10**3*(298.15/291.32)*exp(((-
100000/8.314)*(1/291.32)-(1/298.15)))));
run;

```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```

data LVOOA3;
set LVOOA2;
LVOOA_C1_part_out = (1+((LVOOA_C1_T_out)/4.87))**-1;
LVOOA_C2_part_out = (1+((LVOOA_C2_T_out)/4.87))**-1;
LVOOA_C3_part_out = (1+((LVOOA_C3_T_out)/4.87))**-1;
LVOOA_C4_part_out = (1+((LVOOA_C4_T_out)/4.87))**-1;
LVOOA_C5_part_out = (1+((LVOOA_C5_T_out)/4.87))**-1;
LVOOA_C6_part_out = (1+((LVOOA_C6_T_out)/4.87))**-1;
LVOOA_C7_part_out = (1+((LVOOA_C7_T_out)/4.87))**-1;
LVOOA_C8_part_out = (1+((LVOOA_C8_T_out)/4.87))**-1;
LVOOA_C9_part_out = (1+((LVOOA_C9_T_out)/4.87))**-1;
LVOOA_C10_part_out = (1+((LVOOA_C10_T_out)/4.87))**-1;
LVOOA_C11_part_out = (1+((LVOOA_C11_T_out)/4.87))**-1;
LVOOA_Part_tot_out =
(LVOOA_C1_part_out*LVOOA_C1_tot+LVOOA_C2_part_out*LVOOA_C2_tot+LVOOA_C3
_part_out*LVOOA_C3_tot+
LVOOA_C4_part_out*LVOOA_C4_tot+LVOOA_C5_part_out*LVOOA_C5_tot+LVOOA_C6
_part_out*LVOOA_C6_tot
+LVOOA_C7_part_out*LVOOA_C7_tot+LVOOA_C8_part_out*LVOOA_C8_tot+LVOOA_C9
_part_out*LVOOA_C9_tot+LVOOA_C10_part_out*LVOOA_C10_tot
+LVOOA_C11_part_out*LVOOA_C11_tot)/(LVOOA_C1_tot+LVOOA_C2_tot+LVOOA_C3
_tot+LVOOA_C4_tot+LVOOA_C5_tot+LVOOA_C6_tot+LVOOA_C7_tot+
LVOOA_C8_tot+LVOOA_C9_tot+LVOOA_C10_tot+LVOOA_C11_tot);
run;

```

*SV-OOA

calculate total (gas+particle) organic concentration in each SV-OOA bin using relationship between C and Ctot from Cappa and Jimenez 2010;

```

data SVOOA;

```

```

SV00A_C1_tot = 0.7+7*exp(0.3*(log10(10**-5)-3));
SV00A_C2_tot = 0.7+7*exp(0.3*(log10(10**-4)-3));
SV00A_C3_tot = 0.7+7*exp(0.3*(log10(10**-3)-3));
SV00A_C4_tot = 0.7+7*exp(0.3*(log10(10**-2)-3));
SV00A_C5_tot = 0.7+7*exp(0.3*(log10(10**-1)-3));
SV00A_C6_tot = 0.7+7*exp(0.3*(log10(10**0)-3));
SV00A_C7_tot = 0.7+7*exp(0.3*(log10(10**1)-3));
SV00A_C8_tot = 0.7+7*exp(0.3*(log10(10**2)-3));
SV00A_C9_tot = 0.7+7*exp(0.3*(log10(10**3)-3));
run;

```

Calculate how C(Tref = 298.15 K) shifts outdoors based on temperature using Clausius Clapyeron Eqn.;

```

data SV00A2;
set SV00A;
SV00A_C1_T_out = 10**-5*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
SV00A_C2_T_out = 10**-4*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
SV00A_C3_T_out = 10**-3*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
SV00A_C4_T_out = 10**-2*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
SV00A_C5_T_out = 10**-1*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
SV00A_C6_T_out = 10**0*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
SV00A_C7_T_out = 10**1*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
SV00A_C8_T_out = 10**2*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
SV00A_C9_T_out = 10**3*(298.15/291.32)*exp(((100000/8.314)*(1/291.32)-(1/298.15))));
run;

```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```

data SV00A3;
set SV00A2;
SV00A_C1_part_out = (1+((SV00A_C1_T_out)/4.87))**-1;
SV00A_C2_part_out = (1+((SV00A_C2_T_out)/4.87))**-1;
SV00A_C3_part_out = (1+((SV00A_C3_T_out)/4.87))**-1;
SV00A_C4_part_out = (1+((SV00A_C4_T_out)/4.87))**-1;
SV00A_C5_part_out = (1+((SV00A_C5_T_out)/4.87))**-1;
SV00A_C6_part_out = (1+((SV00A_C6_T_out)/4.87))**-1;
SV00A_C7_part_out = (1+((SV00A_C7_T_out)/4.87))**-1;
SV00A_C8_part_out = (1+((SV00A_C8_T_out)/4.87))**-1;
SV00A_C9_part_out = (1+((SV00A_C9_T_out)/4.87))**-1;
SV00A_Part_tot_out =
(SV00A_C1_part_out*SV00A_C1_tot+SV00A_C2_part_out*SV00A_C2_tot+SV00A_C3
_part_out*SV00A_C3_tot+
SV00A_C4_part_out*SV00A_C4_tot+SV00A_C5_part_out*SV00A_C5_tot+SV00A_C6
_part_out*SV00A_C6_tot

```

```
+SVOOA_C7_part_out*SVOOA_C7_tot+SVOOA_C8_part_out*SVOOA_C8_tot+SVOOA_C9
_part_out*SVOOA_C9_tot)/
(SVOOA_C1_tot+SVOOA_C2_tot+SVOOA_C3_tot+SVOOA_C4_tot+SVOOA_C5_tot+SVOOA
_C6_tot+SVOOA_C7_tot+
SVOOA_C8_tot+SVOOA_C9_tot);
run;
```

```
*OOA;
```

```
*calculate total (gas+particle) organic concentration in each OOA bin
using relationship
between C* and Ctot from Cappa and Jimenez 2010;
```

```
data OOA;
```

```
OOA_C1_tot = 1.94+5.5*exp(0.8*(log10(10**-6)-3));
OOA_C2_tot = 1.94+5.5*exp(0.8*(log10(10**-5)-3));
OOA_C3_tot = 1.94+5.5*exp(0.8*(log10(10**-4)-3));
OOA_C4_tot = 1.94+5.5*exp(0.8*(log10(10**-3)-3));
OOA_C5_tot = 1.94+5.5*exp(0.8*(log10(10**-2)-3));
OOA_C6_tot = 1.94+5.5*exp(0.8*(log10(10**-1)-3));
OOA_C7_tot = 1.94+5.5*exp(0.8*(log10(10**0)-3));
OOA_C8_tot = 1.94+5.5*exp(0.8*(log10(10**1)-3));
OOA_C9_tot = 1.94+5.5*exp(0.8*(log10(10**2)-3));
OOA_C10_tot = 1.94+5.5*exp(0.8*(log10(10**3)-3));
run;
```

```
*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;
```

```
data OOA2;
```

```
set OOA;
```

```
OOA_C1_T_out = 10**-6*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C2_T_out = 10**-5*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C3_T_out = 10**-4*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C4_T_out = 10**-3*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C5_T_out = 10**-2*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C6_T_out = 10**-1*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C7_T_out = 10**-0*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C8_T_out = 10**1*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C9_T_out = 10**2*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OOA_C10_T_out = 10**3*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
run;
```

```
*calculate the outdoor partitioning based on VBS from Cappa and
Jimenez 2010;
```

```
data OOA3;
```



```

set OOA2;
OOA_C1_part_out = (1+((OOA_C1_T_out)/4.87))** -1;
OOA_C2_part_out = (1+((OOA_C2_T_out)/4.87))** -1;
OOA_C3_part_out = (1+((OOA_C3_T_out)/4.87))** -1;
OOA_C4_part_out = (1+((OOA_C4_T_out)/4.87))** -1;
OOA_C5_part_out = (1+((OOA_C5_T_out)/4.87))** -1;
OOA_C6_part_out = (1+((OOA_C6_T_out)/4.87))** -1;
OOA_C7_part_out = (1+((OOA_C7_T_out)/4.87))** -1;
OOA_C8_part_out = (1+((OOA_C8_T_out)/4.87))** -1;
OOA_C9_part_out = (1+((OOA_C9_T_out)/4.87))** -1;
OOA_C10_part_out = (1+((OOA_C10_T_out)/4.87))** -1;
OOA_Part_tot_out =
(OOA_C1_part_out*OOA_C1_tot+OOA_C2_part_out*OOA_C2_tot+OOA_C3_part_out*
OOA_C3_tot+OOA_C4_part_out*OOA_C4_tot+OOA_C5_part_out*OOA_C5_tot+OOA_C6
_part_out*OOA_C6_tot
+OOA_C7_part_out*OOA_C7_tot+OOA_C8_part_out*OOA_C8_tot+OOA_C9_part_out*
OOA_C9_tot+OOA_C10_part_out*OOA_C10_tot)/
(OOA_C1_tot+OOA_C2_tot+OOA_C3_tot+OOA_C4_tot+OOA_C5_tot+OOA_C6_tot+OOA_
C7_tot+OOA_C8_tot+OOA_C9_tot+OOA_C10_tot);
run;

*Other OOA;

*calculate total (gas+particle) organic concentration in each OTH bin
using relationship
between C* and Ctot from Cappa and Jimenez 2010 using the values for
OA;

data OTH;
OTH_C1_tot = 1.65+19*exp(1*(log10(10**-6)-3));
OTH_C2_tot = 1.65+19*exp(1*(log10(10**-5)-3));
OTH_C3_tot = 1.65+19*exp(1*(log10(10**-4)-3));
OTH_C4_tot = 1.65+19*exp(1*(log10(10**-3)-3));
OTH_C5_tot = 1.65+19*exp(1*(log10(10**-2)-3));
OTH_C6_tot = 1.65+19*exp(1*(log10(10**-1)-3));
OTH_C7_tot = 1.65+19*exp(1*(log10(10**0)-3));
OTH_C8_tot = 1.65+19*exp(1*(log10(10**1)-3));
OTH_C9_tot = 1.65+19*exp(1*(log10(10**2)-3));
OTH_C10_tot = 1.65+19*exp(1*(log10(10**3)-3));
run;

*Calculate how C*(Tref = 298.15 K) shifts based on temperature using
Clausius Clapyeron Eqn.;

data OTH2;
set OTH;
OTH_C1_T_out = 10**-6*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OTH_C2_T_out = 10**-5*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OTH_C3_T_out = 10**-4*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OTH_C4_T_out = 10**-3*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));
OTH_C5_T_out = 10**-2*(298.15/291.32)*exp((( -100000/8.314)*(1/291.32) -
(1/298.15))));

```

```

OTH_C6_T_out = 10**-1*(298.15/291.32)*exp(((-100000/8.314)*(1/291.32)-(1/298.15))));
OTH_C7_T_out = 10**-0*(298.15/291.32)*exp(((-100000/8.314)*(1/291.32)-(1/298.15))));
OTH_C8_T_out = 10**1*(298.15/291.32)*exp(((-100000/8.314)*(1/291.32)-(1/298.15))));
OTH_C9_T_out = 10**2*(298.15/291.32)*exp(((-100000/8.314)*(1/291.32)-(1/298.15))));
OTH_C10_T_out = 10**3*(298.15/291.32)*exp(((-100000/8.314)*(1/291.32)-(1/298.15))));
run;

```

*calculate the outdoor partitioning based on VBS from Cappa and Jimenez 2010;

```

data OTH3;
set OTH2;
OTH_C1_part_out = (1+((OTH_C1_T_out)/4.87))**-1;
OTH_C2_part_out = (1+((OTH_C2_T_out)/4.87))**-1;
OTH_C3_part_out = (1+((OTH_C3_T_out)/4.87))**-1;
OTH_C4_part_out = (1+((OTH_C4_T_out)/4.87))**-1;
OTH_C5_part_out = (1+((OTH_C5_T_out)/4.87))**-1;
OTH_C6_part_out = (1+((OTH_C6_T_out)/4.87))**-1;
OTH_C7_part_out = (1+((OTH_C7_T_out)/4.87))**-1;
OTH_C8_part_out = (1+((OTH_C8_T_out)/4.87))**-1;
OTH_C9_part_out = (1+((OTH_C9_T_out)/4.87))**-1;
OTH_C10_part_out = (1+((OTH_C10_T_out)/4.87))**-1;
OTH_Part_tot_out =
(OTH_C1_part_out*OTH_C1_tot+OTH_C2_part_out*OTH_C2_tot+OTH_C3_part_out*
OTH_C3_tot+OTH_C4_part_out*OTH_C4_tot+OTH_C5_part_out*OTH_C5_tot+OTH_C6
_part_out*OTH_C6_tot
+OTH_C7_part_out*OTH_C7_tot+OTH_C8_part_out*OTH_C8_tot+OTH_C9_part_out*
OTH_C9_tot+OTH_C10_part_out*OTH_C10_tot)/
(OTH_C1_tot+OTH_C2_tot+OTH_C3_tot+OTH_C4_tot+OTH_C5_tot+OTH_C6_tot+OTH_
C7_tot+OTH_C8_tot+OTH_C9_tot+OTH_C10_tot);
run;

```

*Data available in library "Ch5" in file "Average VBS.xlsx"

Appendix D References

1. Cappa C. D.; Jimenez J., Quantitative estimates of the volatility of ambient organic aerosol. *Atmos. Chem. Phys.* **2010**, 10, 5409–5424.
2. Jimenez, J. L.; Canagaratna, M. R.; Donahue, N. M.; Prevot, A. S. H.; Zhang, Q.; Kroll, J. H.; DeCarlo, P. F.; Allan, J. D.; Coe, H.; Ng, N. L.; Aiken, A. C.; Docherty, K. S.; Ulbrich, I. M.; Grieshop, A. P.; Robinson, A. L.; Duplissy, J.; Smith, J. D.; Wilson, K. R.; Lanz, V. A.; Hueglin, C.; Sun, Y. L.; Tian, J.; Laaksonen, A.; Raatikainen, T.; Rautiainen, J.; Vaattovaara, P.; Ehn, M.; Kulmala, M.; Tomlinson, J. M.; Collins, D. R.; Cubison, M. J.; Dunlea, E. J.; Huffman, J. A.; Onasch, T. B.; Alfarra, M. R.; Williams, P. I.; Bower, K.; Kondo, Y.; Schneider, J.; Drewnick, F.; Borrmann, S.; Weimer, S.; Demerjian, K.; Salcedo, D.; Cottrell, L.; Griffin, R.; Takami, A.; Miyoshi, T.; Hatakeyama, S.; Shimono, A.; Sun, J. Y.; Zhang, Y. M.; Dzepina, K.; Kimmel, J. R.; Sueper, D.; Jayne, J. T.; Herndon, S. C.; Trimborn, A. M.; Williams, L. R.; Wood, E. C.; Middlebrook, A. M.; Kolb, C. E.; Baltensperger, U.; Worsnop, D. R., Evolution of organic aerosol in the atmosphere. *Science* **2009**, 326, 1525 - 1529.
3. Tanaka, P.; Riemer, D. D.; Chang, S.; Yarwood, G.; McDonald-Buller, E. C.; Apel, E. C.; Orlando, J. J.; Silva, P. J.; Jimenez, J. L.; Canagaratna, M. R.; Neece, J. D.; Mullins, C. B.; Allen, D. T., Direct evidence for chlorine-enhanced urban ozone formation in Houston, Texas. *Atmos. Environ.* **2003**, 37, 1393-1400.
4. DeCarlo, P. F.; Kimmel, J. R.; Trimborn, A.; Northway, M. J.; Jayne, J. T.; Aiken, A. C.; Gonin, M.; Fuhrer, K.; Horvath, T.; Docherty, K. S.; Worsnop, D. R.; Jimenez, J. L., Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. *Anal. Chem.* **2006**, 78, 8281-8289.
5. Docherty, K. S.; Stone, E. A.; Ulbrich, I. M.; DeCarlo, P. F.; Snyder, D. C.; Schauer, J. J.; Peltier, R. E.; Weber, R. J.; Murphy, S. M.; Seinfeld, J. H.; Grover, B. D.; Eatough, D. J.; Jimenez, J. L., Apportionment of primary and secondary aerosols in Southern California during the 2005 Study of Organic Aerosols in Riverside (SOAR-1). *Environ. Sci. Technol.* **2008**, 42, 7655-7662.
6. Cubbison, M. J.; Alfarra, M. R.; Allan, J.; Bower, K. N.; Coe, H.; McFiggans, G. B.; Whitehead, J. D.; Williams, P. I.; Zhang, Q.; Jimenez, J. L.; Hopkins, J.; Lee, J., The characterisation of pollution aerosol in a changing photochemical environment. *Atmos. Chem. Phys.* **2008**, 6, 5573 - 5588.
7. Drewnick, F.; Schwab, J. J.; Jayne, J. T.; Canagaratna, M.; Worsnop, D. R.; Demerjian, K. L., Measurement of ambient aerosol composition during the PMTACS-NY using an aerosol mass spectrometer. Part I: mass concentrations. *Aerosol Sci. Technol.*, **2004**, 38, 92 - 103.
8. Drewnick, F.; Jayne, J. T.; Canagaratna, M.; Worsnop, D. R.; Demerjian, K. L., Measurement of ambient aerosol composition during the PMTACS-NY 2001 using an

aerosol mass spectrometer. Part II: chemically speciated mass distributions. *Aerosol Sci. Technol.* **2004**, 38, 104 - 117.

9. Weimer, S.; Drewnick, F.; Högrefe, O.; Schwab, J. J.; Rhoads, K.; Orsini, D.; Canagaratna, M.; Worsnop, D. R.; Demerjian, K. L., Size-selective nonrefractory ambient aerosol measurements during the Particulate Matter Technology Assessment and Characterization Study–New York 2004 Winter Intensive in New York City. *J. Geophys. Res.*, **2006**, 111, D18305, doi:10.1029/2006JD007215.