Walter A. Rosenblith New Investigator Award
RESEARCH REPORT

HEALTH EFFECTS INSTITUTE

Number 152
December 2010

Evaluating Heterogeneity in Indoor and Outdoor Air Pollution Using Land-Use Regression and Constrained Factor Analysis

Jonathan I. Levy, Jane E. Clougherty, Lisa K. Baxter, E. Andres Houseman, and Christopher J. Paciorek

Evaluating Heterogeneity in Indoor and Outdoor Air Pollution Using Land-Use Regression and Constrained Factor Analysis

Jonathan I. Levy, Jane E. Clougherty, Lisa K. Baxter, E. Andres Houseman, and Christopher J. Paciorek

with a Critique by the HEI Health Review Committee

Research Report 152

Health Effects Institute

Boston, Massachusetts

Trusted Science • Cleaner Air • Better Health

Publishing history: The Web version of this document was posted at *www.healtheffects.org* in December 2010.

Citation for document:

Levy JI, Clougherty JE, Baxter LK, Houseman EA, Paciorek CJ. 2010. Evaluating Heterogeneity in Indoor and Outdoor Air Pollution Using Land-Use Regression and Constrained Factor Analysis. Research Report 152. Health Effects Institute, Boston, MA.

 Cameographics, Belfast, Me., Compositor. Printed by Recycled Paper Printing, Boston, Mass. Library of Congress Number for the HEI Report Series: WA 754 R432.

Cover paper: made with at least 55% recycled content, of which at least 30% is post-consumer waste; free of acid and elemental chlorine. Text paper: made with 100% post-consumer waste recycled content; acid free; no chlorine used in processing. The book is printed with soy-based inks and is of permanent archival quality.

CONTENTS

ABOUT HEI

The Health Effects Institute is a nonprofit corporation chartered in 1980 as an independent research organization to provide high-quality, impartial, and relevant science on the effects of air pollution on health. To accomplish its mission, the institute

- Identifies the highest-priority areas for health effects research;
- Competitively funds and oversees research projects;
- Provides intensive independent review of HEI-supported studies and related research;
- Integrates HEI's research results with those of other institutions into broader evaluations; and
- Communicates the results of HEI's research and analyses to public and private decision makers.

HEI receives half of its core funds from the U.S. Environmental Protection Agency and half from the worldwide motor vehicle industry. Frequently, other public and private organizations in the United States and around the world also support major projects or certain research programs. HEI has funded more than 280 research projects in North America, Europe, Asia, and Latin America, the results of which have informed decisions regarding carbon monoxide, air toxics, nitrogen oxides, diesel exhaust, ozone, particulate matter, and other pollutants. These results have appeared in the peer-reviewed literature and in more than 200 comprehensive reports published by HEI.

HEI's independent Board of Directors consists of leaders in science and policy who are committed to fostering the public–private partnership that is central to the organization. The Health Research Committee solicits input from HEI sponsors and other stakeholders and works with scientific staff to develop a Five-Year Strategic Plan, select research projects for funding, and oversee their conduct. The Health Review Committee, which has no role in selecting or overseeing studies, works with staff to evaluate and interpret the results of funded studies and related research.

All project results and accompanying comments by the Health Review Committee are widely disseminated through HEI's Web site (*www.healtheffects.org*), printed reports, newsletters and other publications, annual conferences, and presentations to legislative bodies and public agencies.

ABOUT THIS REPORT

Research Report 152, *Evaluating Heterogeneity in Indoor and Outdoor Air Pollution Using Land-Use Regression and Constrained Factor Analysis,* presents a research project funded by the Health Effects Institute and conducted by Dr. Jonathan I. Levy of the Harvard School of Public Health, Boston, Massachusetts, and his colleagues. This research was funded under HEI's Walter A. Rosenblith New Investigator Award Program, which provides support to promising scientists in the early stages of their careers. The report contains three main sections.

The HEI Statement, prepared by staff at HEI, is a brief, nontechnical summary of the study and its findings; it also briefly describes the Health Review Committee's comments on the study.

The Investigators' Report, prepared by Levy and colleagues, describes the scientific background, aims, methods, results, and conclusions of the study.

The Critique is prepared by members of the Health Review Committee with the assistance of HEI staff; it places the study in a broader scientific context, points out its strengths and limitations, and discusses remaining uncertainties and implications of the study's findings for public health and future research.

This report has gone through HEI's rigorous review process. When an HEI-funded study is completed, the investigators submit a draft final report presenting the background and results of the study. This draft report is first examined by outside technical reviewers and a biostatistician. The report and the reviewers' comments are then evaluated by members of the Health Review Committee, an independent panel of distinguished scientists who have no involvement in selecting or overseeing HEI studies. During the review process, the investigators have an opportunity to exchange comments with the Review Committee and, as necessary, to revise their report. The Critique reflects the information provided in the final version of the report.

HEI STATEMENT

Synopsis of Research Report 152

Evaluating Heterogeneity in Indoor and Outdoor Air Pollution Using Land-Use Regression and Constrained Factor Analysis

BACKGROUND

Epidemiologic studies of exposure to air pollution have typically relied on data from centrally located ambient air quality monitors. However, such data are not sufficient for capturing the spatial variability of pollutant concentrations at the local scale, in particular at the within-city, or intra-urban, scale at which traffic-related air pollution is both highest and most variable. The ideal approach would be to measure each individual's personal exposure to traffic-specific pollutants over time, but this is difficult, intrusive, expensive, and generally not feasible for very large populations. Investigators have consequently sought ways to predict, or to model, individual-level exposures from more readily available data.

Numerous studies have had to rely on a variety of surrogates for such exposures: measured levels of individual pollutants previously associated with traffic emissions (e.g., carbon monoxide, nitrogen dioxide [NO_2], fine particulate matter with an aerodynamic diameter $\leq$ 2.5 µm [$PM_{2.5}$], benzene, and elemental carbon [EC]) and various measures of traffic density or of proximity to traffic. More complex techniques, such as land-use regression (LUR) models, have been increasingly developed to take advantage of data available from geographic information systems (GIS) — nearby land-use patterns, traffic, physical site characteristics, housing, and other variables — that have been hypothesized to provide additional information useful for predicting concentrations of traffic-related pollutants. Each type of surrogate has limitations in predicting levels of personal exposures to traffic-related pollutants. HEI Special Report 17, a critical review of the literature on traffic-related air pollution, found, in particular, that many of the simpler surrogates do not perform well. The resulting error in individuals' exposures can in turn affect the size and significance of health outcome findings from observational epidemiologic studies.

Dr. Jonathan I. Levy of the Harvard School of Public Health was awarded funding from HEI under Request for Applications 04-5, the Walter A. Rosenblith New Investigator Award. In his application, "Using Geographic Information Systems (GIS) to Evaluate Heterogeneity in Indoor and Outdoor Concentrations of Particle Constituents," Levy had proposed an approach to extend and improve upon existing GIS-based methods for predicting intra-urban exposures. An underlying goal of his study was therefore to explore ways to reduce exposure-measurement error and to improve the accuracy and precision of the associations reported in epidemiologic studies of air pollution.

APPROACH

Levy and colleagues conducted a study linked to a prospective birth cohort study of factors that might contribute to the development of asthma, the Asthma Coalition for Community, Environment, and Social Stress study in Boston, Massachusetts. Among the several factors under investigation were indoor and outdoor exposures to air pollutants, including those potentially related to traffic. Levy and colleagues collected detailed air quality measurements at a set of homes selected to reflect a range of potential exposures to traffic and of neighborhoods broadly representative of Boston. From 2003 through 2005, the investigators collected short-term NO_2 and $PM_{2.5}$ samples simultaneously indoors and outdoors at each home during two seasons. The particle

This Statement, prepared by the Health Effects Institute, summarizes a research project funded by HEI and conducted by Dr. Jonathan I. Levy at the Harvard School of Public Health, Boston, Massachusetts, and colleagues. Research Report 152 contains both the detailed Investigators' Report and a Critique of the study prepared by the Institute's Health Review Committee.

filters used to collect the samples were analyzed for EC and for individual elements using two different analytical methods. The investigators also obtained hourly NO_2, $PM_{2.5}$, EC, and meteorological data for the study period from centrally located Massachusetts Department of Environmental Protection monitors, to provide data on the variation in background pollutant levels over time.

The investigators collected several additional types of data to support development of their LUR models. They utilized existing GIS data on road networks, traffic counts, and population density to characterize proximity to and potential density of traffic in the vicinity of each home. They obtained additional data on local land use and on the age of each home, its living area, building materials, heating system, and whether or not it had air conditioning. Investigators administered a standardized questionnaire to participants at each home to obtain data on occupant behaviors and home characteristics that have been shown previously to indicate indoor sources of pollutants or to influence ventilation in the home.

Levy and his colleagues then undertook a series of systematic analytic approaches to predicting concentrations of $PM_{2.5}$, EC, and NO_2 measured at each of the homes in the study and to understanding their potential sources. Using multiple variables drawn from their indoor and outdoor residential monitoring data, GIS-based land-use data, and questionnaire data, they first developed separate GIS-based LUR models to predict concentrations of $PM_{2.5}$, EC, and NO_2 measured outdoors and indoors at the residences. Second, using constrained factor analysis, a source apportionment technique, they analyzed the particle components and NO_2 measurements to identify potential source categories for pollutants measured outdoors and indoors at the homes in the study. The third step in their approach was to apply LUR analysis to the results of their source apportionment analyses; that is, they developed additional LUR models designed specifically to help explain variability in the source categories identified from their source apportionment analysis. By evaluating the extent to which they could successfully predict sources using particular GIS, land-use, and questionnaire data collected for the study, the investigators sought to corroborate their initial interpretations of the source apportionment analysis.

Finally, using simulation techniques, they conducted an analysis to assess how a variety of possible surrogates for indoor exposures, representing different levels of exposure-measurement error, could influence epidemiologic estimates of the relationship between indoor pollutant concentrations and reports of wheeze (a possible indicator of asthma) in a child's first year of life. They compared the performance of their indoor LUR models for $PM_{2.5}$, NO_2, and EC to that of surrogates based on single variables that had performed well in their model development process ("good exposure surrogates") and to that of surrogates based on traffic indicators that had not performed well in their analysis, but that had been used in studies reported by other investigators ("poor exposure surrogates"). They ran their simulations using three scenarios for the strength of the "true" associations between individual exposure and wheeze.

RESULTS

Levy and his colleagues reported that their final multivariate outdoor LUR models performed reasonably well; the models were able to explain most of the variability in outdoor residential concentrations of EC, NO_2, and $PM_{2.5}$ (52%, 56%, and 76%, respectively). EC and NO_2 had stronger relationships with indicators for local traffic than did $PM_{2.5}$. The variation in pollutant levels over time, represented by measurements at the central site monitor and seasonal terms, explained more of the variability in $PM_{2.5}$ (68%) than in EC (30%) or NO_2 (33%), a finding consistent with other studies.

They reported less success in the ability of their LUR regression models to predict indoor concentrations of the three pollutants. The indoor LUR models could explain only 20%, 21%, and 36% of the variation in indoor NO_2, EC, and $PM_{2.5}$ levels, respectively. They found identifying traffic terms with strong explanatory power to include in their models to be particularly challenging. When ventilation terms were introduced into the models, the explanatory power of the models increased slightly, to 25%, 32%, and 40%, respectively. The investigators reported that the ratios between the indoor and outdoor concentrations of individual particle constituents varied substantially among the different constituents, which later enabled some distinctions to be made between their potential indoor and outdoor sources.

From their source apportionment analysis using outdoor pollutant concentrations, Levy and his colleagues indentified five broad source categories: long-range transport; brake wear and local traffic; diesel exhaust; fuel oil combustion; and road dust and resuspension. Their analysis of measured indoor pollutant concentrations suggested six possible source categories, three interpreted to have origins outside the home — long-range transport, fuel oil/diesel combustion, and road dust and resuspension — and three interpreted to have indoor origins — indoor combustion, indoor smoking, and indoor cleaning.

Levy and colleagues had mixed success in their efforts to use LUR models to explore more fully the potential predictors for the source categories they had identified. In general, the LUR models they developed to predict the outdoor source categories had weaker explanatory power than those they had developed earlier to predict the levels of individual pollutants. The one exception was the investigators' LUR model for long-range transport, which was able to explain 69% of the variation observed. The strong performance of this model was consistent with that of the earlier LUR model for outdoor $PM_{2.5}$, since long-range transport was most closely associated with $PM_{2.5}$ measured at the central site monitor. The LUR model for predicting long-range transport indoors also performed the best. Most of the variation (68%) in the source category was explained by a term representing a combination of $PM_{2.5}$ data from the central site monitor and a variable obtained from the questionnaires that was indicative of greater ventilation in the homes (i.e., open windows). The LUR models for the remaining indoor source factors had very little explanatory power, in most cases substantially less than the LUR models for outdoor source factors.

From their simulation analysis, Levy and colleagues reported that the risks of wheeze estimated using exposures to individual pollutants based on their indoor LUR models were closer to the "true" risks than those estimated using either the simpler "good" or "poor" surrogates for exposure. That is, there was less bias and less uncertainty in the predicted risk estimates relative to the known risks used in the simulation. The models for NO_2 and $PM_{2.5}$ performed better than the one for EC. The investigators inferred from this simulation analysis that their LUR models predicted individual exposure levels with less exposure-measurement error than the individual surrogate approaches, and thus enhanced the power of the simulated epidemiologic study to detect the underlying association between wheeze and the pollutants.

CONCLUSIONS

Levy and his colleagues took advantage of a small but rich data source related to a study of childhood asthma in a major U.S. city to explore important exposure questions that are of broad interest to environmental health science. They undertook a number of challenging methodologic approaches to improving predictions of personal exposure to pollution from indoor and outdoor sources and thus to improving epidemiologic estimates of the effects of traffic-related air pollution on health. Their report marks one of the first efforts to combine LUR models with source apportionment analysis to characterize potential exposures to both indoor and outdoor sources. The HEI Health Review Committee praised the evident care and competence demonstrated by the investigators in their work.

The investigators' LUR analyses of outdoor pollutant levels performed reasonably well, explaining most of the variation in concentrations of $PM_{2.5}$ and to a lesser extent in NO_2 and EC. Their results were consistent with previously published findings showing that broader-scale temporal variation, represented by measurements at the central site monitor, is an important determinant of local $PM_{2.5}$ levels. Spatially distributed factors, such as traffic, population density, and other land-use covariates, were more influential in predicting EC and NO_2 variation in the models.

Development of LUR models to predict indoor concentrations, as a closer proxy for personal exposure, proved to be a much greater challenge. The predictive value of the indoor LUR models was generally poor; however, the authors' exploration of the possible explanations and implications of this finding is thorough and informative. Their findings that the indoor LUR models' performance was poorer when important indoor sources were present is noteworthy, as is the finding that the performance of an indoor LUR model can be improved by the relatively straightforward addition of a proxy term for ventilation taken from questionnaire data (i.e., "open windows").

An ultimate, and the most innovative, goal of the study was to see if LUR modeling and source apportionment analysis together would provide additional insight about the sources contributing to outdoor and indoor concentrations of pollutants and help explain why their contributions might differ at individual homes. The analyses did provide some confirmation for the investigators' major source interpretations (long-range transport, traffic, fuel oil combustion). However, they were most successful at explaining variation in sources that are already reasonably well understood. In particular, both the outdoor and indoor LUR models performed best at predicting variability in the source identified most closely with $PM_{2.5}$ concentrations at the central site monitor, long-range transport. As for the indoor LUR models developed to predict individual pollutant concentrations, incorporating a proxy for ventilation improved the performance of the indoor LUR models designed to predict variation in this source.

The ultimate challenge for studies of this nature is to provide some demonstration that the increased sophistication of the modeling provides a sufficient improvement over simpler approaches to warrant the additional data and computational requirements it imposes. The investigators' simulation analyses, in which they explore the implications for the power of epidemiologic studies of using different surrogate measures of individual exposure, are a useful step in that direction. Their conclusion that even the relatively poor estimates of exposure provided by the LUR models might reduce measurement error and thus improve effects estimates in future studies warrants further scrutiny. Even when a poor surrogate is outperformed by a prediction model, the surrogate may be the epidemiologist's "best buy" if the extent of improved performance is outweighed by the costs of collecting the data necessary for the prediction model.

INVESTIGATORS' REPORT

Evaluating Heterogeneity in Indoor and Outdoor Air Pollution Using Land-Use Regression and Constrained Factor Analysis

Jonathan I. Levy, Jane E. Clougherty, Lisa K. Baxter, E. Andres Houseman, and Christopher J. Paciorek

Harvard School of Public Health, Department of Environmental Health (J.I.L., J.E.C., L.K.B.); Harvard School of Public Health, Department of Biostatistics (E.A.H., C.J.P.)

ABSTRACT

Previous studies have identified associations between traffic exposures and a variety of adverse health effects, but many of these studies relied on proximity measures rather than measured or modeled concentrations of specific air pollutants, complicating interpretability of the findings. An increasing number of studies have used land-use regression (LUR*) or other techniques to model small-scale variability in concentrations of specific air pollutants. However, these studies have generally considered a limited number of pollutants, focused on outdoor concentrations (or indoor concentrations of ambient origin) when indoor concentrations are better proxies for personal exposures, and have not taken full advantage of statistical methods for source apportionment that may have provided insight about the structure of the LUR models and the interpretability of model results. Given these issues, the primary objective of our study was to determine predictors of indoor and outdoor residential concentrations of multiple traffic-related air pollutants within an urban area, based on a combination of central site monitoring data; geographic information system (GIS) covariates reflecting traffic and other outdoor sources; questionnaire data reflecting indoor sources and activities that affect ventilation rates; and factor-analytic methods to better infer source contributions.

As part of a prospective birth cohort study assessing asthma etiology in urban Boston, we collected indoor and/or outdoor 3-to-4 day samples of nitrogen dioxide (NO_2) and fine particulate matter with an aerodynamic diameter $\leq$ 2.5 μm ($PM_{2.5}$) at 44 residences during multiple seasons of the year from 2003 through 2005. We performed reflectance analysis, x-ray fluorescence spectroscopy (XRF), and high-resolution inductively coupled plasma–mass spectrometry (ICP–MS) on particle filters to estimate the concentrations of elemental carbon (EC), trace elements, and water-soluble metals, respectively. We derived multiple indicators of traffic using Massachusetts Highway Department (MHD) data and traffic counts collected outside the residences where the air monitoring was conducted. We used a standardized questionnaire to collect data on home characteristics and occupant behaviors. Additional housing information was collected through property tax records. Ambient concentrations of pollutants as well as meteorological data were collected from centrally located ambient monitors.

We used GIS-based LUR models to explain spatial and temporal variability in residential outdoor concentrations of $PM_{2.5}$, EC, and NO_2. We subsequently derived latent-source factors for residential outdoor concentrations using confirmatory factor analysis constrained to nonnegative loadings. We developed LUR models to determine whether GIS covariates and other predictors explain factor variability and thereby support initial factor interpretations. To evaluate indoor concentrations, we developed physically interpretable regression models that explored the relationship between measured indoor and outdoor concentrations, relying on

This Investigators' Report is one part of Health Effects Institute Research Report 152, which also includes a Critique by the Health Review Committee and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr. Jonathan I. Levy, 715 Albany St., Talbot 4W, Boston, MA 02118; jonlevy@bu.edu.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83234701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

* A list of abbreviations and other terms appears at the end of the Investigators' Report.

questionnaire data to characterize indoor sources and activities. Because outdoor pollutant concentrations measured directly outside of homes are unlikely to be available for most large epidemiologic studies, we developed regression models to explain indoor concentrations of $PM_{2.5}$, EC, and NO_2 as a function of other, more readily available data: GIS covariates, questionnaire data reflecting both sources and ventilation, and central site monitoring data. As we did for outdoor concentrations, we then derived latent-source factors for residential indoor concentrations and developed regression models explaining variability in these indoor latent-source factors. Finally, to provide insight about the effects of improved characterization of exposures for the results of subsequent epidemiologic investigations, we developed a simulation framework to quantitatively compare the implications of using exposure models derived from validation studies with the use of other surrogate models with varying amounts of measurement error.

The concentrations of outdoor $PM_{2.5}$ were strongly associated with the central site monitor data, whereas EC concentrations showed greater spatial variability, especially during colder months, and were predicted by the length of roadway within 200 m of the home. Outdoor NO_2 also showed significant spatial variability, predicted in part by population density and roadway length within 50 m of the home. Our constrained factor analysis of outdoor concentrations produced loadings indicating long-range transport, brake wear and traffic exhaust, diesel exhaust, fuel oil combustion, and resuspended road dust as sources; corresponding LUR models largely corroborated these factor interpretations through covariate significance. For example, long-range transport was predicted by central site $PM_{2.5}$ and season, brake wear and traffic exhaust and resuspended road dust by traffic and residential density, diesel exhaust by the percentage of diesel traffic on the nearest major road, and fuel oil combustion by population density.

Our modeling of the concentrations of indoor pollutants demonstrated substantial variability in indoor–outdoor relationships across constituents, helping to separate constituents dominated by outdoor sources (e.g., S, Se, and V) from those dominated by indoor sources (e.g., Ca and Si). Regression models indicated that indoor $PM_{2.5}$ was not influenced substantially by local traffic but had significant indoor sources (cooking activity and occupant density), while EC was associated with distance to the nearest designated truck route, and NO_2 was associated with both traffic density within 50 m of the home and gas stove usage. Our constrained factor analysis of indoor concentrations helped to separate outdoor-dominated factors from indoor-dominated factors, though some factors appeared to be influenced by both indoor and outdoor sources. Subsequent factor analyses of the indoor-attributable fractions from indoor–outdoor regression models provided generally consistent interpretations of indoor-dominated factors. The use of regression models on indoor factors demonstrated the limited predictive power of questionnaire data related to indoor sources, but reinforced the viability of modeling indoor concentrations of pollutants of ambient origin. In spite of the relatively weak predictive power of some of the indoor-concentration regression models, our epidemiologic simulations illustrated that exposure models with fairly modest R^2 values (in the range of 0.3 through 0.4, corresponding with the regression models for $PM_{2.5}$ and NO_2) yielded substantial improvements in epidemiologic study performance relative to the use of exposure proxies that could be applied in the absence of validation studies.

In spite of limitations related to sample size and available covariate data, our study demonstrated significant outdoor spatial variability within an urban area in NO_2 and in several constituents of airborne particles. LUR techniques combined with constrained factor analysis helped to disentangle the contributions to temporal variability of local, long-range transport, and other sources, ultimately allowing exposures from defined source categories to be investigated in epidemiologic studies. For the indoor residential environment, we demonstrated substantial variability in indoor–outdoor relationships among particle constituents; then, using information from public databases and focused questionnaire data, we were able to predict indoor concentrations for a subset of key pollutants. Constrained factor analysis methods applied to the indoor environment helped to separate indoor sources from outdoor sources. The corresponding indoor regression models had limited predictive power, reinforcing the complexity of characterizing the indoor environment when only limited information about key predictors is available. This finding also underscores the likelihood that these regression models might characterize indoor concentrations of pollutants with ambient origins better than they can the indoor concentrations from all sources. Our findings provide direction for future studies characterizing indoor exposure sources and patterns, and our epidemiologic simulation reinforced the importance of reducing measurement error in a context where many traffic-related air pollutants are influenced by both indoor and outdoor sources. The combination of analytical techniques used in our study could ultimately allow for more refined exposure characterization and evaluation of the relative contributions of various sources to health outcomes in epidemiologic studies.

INTRODUCTION

There is a growing body of literature demonstrating the use of GIS to relate traffic exposure to asthma and other respiratory outcomes. In the United States and Europe, children living or attending school near truck routes and highways show more asthma symptoms (Brauer et al. 2002; Zmirou et al. 2004; Gordian et al. 2006), asthma hospitalizations (Edwards et al. 1994; Lin et al. 2002), respiratory illness (Brauer et al. 2002), allergic rhinitis (Duhme et al. 1996), and reduced lung function (Brunekreef et al. 1997). Other studies have linked proximity to major roads with excess risk of mortality from all causes (Finkelstein et al. 2004), from stroke (Maheswaran and Elliott 2003), or from cardiopulmonary disease (Hoek et al. 2002).

While this literature has been consistent and robust, it has a few key limitations. First, proximity measures can represent a variety of air pollutants or other near-roadway exposures, including noise, and can also be correlated with socioeconomic status and other nonphysical environmental stressors. Without identifying specific agents or sources, it is difficult to infer causality or determine appropriate interventions. The same proximity measure may represent different air pollutants in different locations, or may represent the same pollutant at different levels of exposure, making it difficult to determine generalized relationships. In addition, there are many proximity measures (e.g., distance to nearest major road, length of road segments within a defined radius, vehicle miles traveled within a defined radius), and it remains unclear which are most appropriate for capturing variability in any specific pollutant or location, especially in the absence of validation sampling.

Some studies have moved beyond proximity measures by using LUR or other techniques to model smaller scale variability in pollutant concentrations within urban areas (Jerrett et al. 2005). These studies generally involve the simultaneous collection of pollution measurements at multiple sites (in many cases, focusing on integrated NO_2, given its association with traffic exhaust and the availability of low-cost passive samplers); they also involve using covariates related to traffic density and meteorology as predictors of the measured concentrations. For NO_2, models have generally been able to explain a majority of its spatial variability with a subset of key traffic predictors (Brauer et al. 2003; Gilbert et al. 2005, 2007; Ross et al. 2006, 2007; Sahsuvaroglu et al. 2006; Arain et al. 2007; Rosenlund et al. 2008; Su et al. 2008). For example, in Hamilton, Canada, an LUR study of 100 samplers placed concurrently during a single 2-week period reported that covariates for traffic density, roadway proximity, and land use explained 76% of the variability in NO_2 concentrations (Sahsuvaroglu et al. 2006). Similarly, 79% of the variability in NO_2 concentrations for 39 samplers placed concurrently in San Diego, California, was explained by traffic density, length of road segments, and distance to the coast (Ross et al. 2006), and 54% of the variability in NO_2 concentrations across 67 samplers placed concurrently in Montreal, Canada, was explained by distance to roadways, length of road segments, and population density (Gilbert et al. 2005). Given the influences of meteorology and atmospheric chemistry, these relationships between NO_2 and land-use characteristics can vary over time, so other NO_2 studies have analyzed measurements taken at the same locations over multiple sampling sessions. However, rather than focusing on temporal variability, these studies developed LUR models that predicted average concentrations across seasons (Arain et al. 2007; Jerrett et al. 2007; Madsen et al. 2007; Rosenlund et al. 2008).

A growing number of studies have moved beyond NO_2 and have considered other pollutants that are potentially related to traffic. Nitric oxide (NO) has been measured in studies using similar LUR methods (Beelen et al. 2007; Henderson et al. 2007; Madsen et al. 2007; Su et al. 2008), and, as anticipated, its concentrations have generally displayed more spatial heterogeneity than the concentrations of NO_2. EC has been considered in fewer studies (Beelen et al. 2007; Henderson et al. 2007; Morgenstern et al. 2007; Nethery et al. 2008), most of which were conducted within Europe, where diesel vehicles are more common, and produced a stronger association with EC than found in the United States. Other LUR models have considered $PM_{2.5}$ (Henderson et al. 2007; Moore et al. 2007; Morgenstern et al. 2007; Ross et al. 2007; Nethery et al. 2008). Although $PM_{2.5}$ is generally influenced more by regional transport than by local combustion, some models of annual average concentrations found traffic density and land-use factors to be significant, indicating some spatial variability after adjustment for temporal factors.

Although these studies have helped to explain the spatial variability of multiple traffic-related pollutants, there are some notable limitations in the literature. First, most studies either collected all measurements at a single point in time or collected a small number of repeated measurements at the same sites. These study designs reduce complexities in separating spatial and temporal variability and are practical approaches for ambient monitoring with passive samplers. However, for multipollutant studies that rely on active samplers at residences, simultaneous sampling at all sites may not be possible. The number of instruments available may be limited, and residential monitoring can be time-consuming and labor-intensive.

More generally, the pollutants measured in LUR studies may or may not be the causal agents for health outcomes, and without additional monitoring, it is not clear whether the pollutants more commonly measured (such as NO_2 and NO) are reasonable proxies for the causal agents in traffic exhaust. While many traffic-related pollutants are highly correlated, relationships among them can differ by site and setting, given differences in traffic composition and in how specific pollutant concentrations vary with distance from sources (i.e., distance-dependent relationships) (Zhou and Levy 2007).

A further limitation of proximity-to-roadway measures and LUR studies is the implicit assumption that residential outdoor (or near-roadway) concentrations of pollutants reasonably approximate personal exposure to pollutants. Residential indoor concentrations are consistently better predictors of personal exposure than are residential outdoor concentrations (Clayton et al. 1993; Ozkaynak et al. 1996; Koistinen et al. 2004; Johannesson et al. 2007), given the time people spend indoors and the strong influence of ventilation and indoor sources on indoor concentrations. The predictive power of residential outdoor concentrations can be related to a number of factors, including the ventilation characteristics of the homes being sampled, the presence and type of indoor sources, the traffic intensity, and the variability in each of these factors across a given cohort. For pollutants with strong indoor sources, such as NO_2 (which is greatly affected by gas stove use) (Levy et al. 1998; Zota et al. 2005), residential outdoor concentrations may be a weaker proxy than for pollutants lacking indoor sources. For example, one recent study found no differences between the personal exposures to $PM_{2.5}$ and NO_2 of adults living on high- or low-traffic streets because of the influence on exposures of time-activity patterns, indoor sources, and ventilation (Van Roosbroeck et al. 2008). Significant differences in exposure to EC were seen, but as this study was conducted in Europe, where diesel-powered vehicles are more prevalent than in the United States, the results may not generalize to the United States. A related study of 14 Dutch children found that the contrast in exposure, expressed as the ratio of the geometric mean personal exposures for children living near major roadways to those of children living at urban background locations, was greater for EC than for NO_2 (van Roosbroeck et al. 2006).

Unlike for outdoor concentrations of pollutants, few studies have attempted to use LUR-style models to predict indoor concentrations or personal exposures. A recent model developed to explore personal exposures for pregnant women (Nethery et al. 2008) found that land-use terms did not predict EC or $PM_{2.5}$ exposures. While land-use terms partially predicted NO_2 exposures, the presence of a gas stove explained far more variability in personal exposures. As part of the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air), Cohen and colleagues (2009) are using indoor–outdoor monitoring coupled with GIS information to develop personal exposure models, but only for exposures to air pollution of ambient origin. For young children who spend a substantial amount of time inside the home, indoor residential concentrations may be a good surrogate of personal exposure, but it is unclear which combination of outdoor-source, indoor-source, and ventilation factors would best predict indoor concentrations of various pollutants.

Because of the importance of factors that can modify indoor concentrations, such as ventilation and indoor sources, the exposure misclassification induced by using proximity measures in an epidemiologic study could be significant, as these represent uncertain proxies for concentrations of specific outdoor air pollutants, which themselves represent uncertain proxies of personal exposures. This exposure misclassification reduces the ability of epidemiologic studies to detect significant effects. Moreover, the misclassification may be differential with respect to particular subpopulations if the outdoor concentration–personal exposure relationship differs by housing characteristics that are in turn linked to socioeconomic status or other health risk factors. While numerous studies have evaluated the influence of exposure misclassification on epidemiologic results (Wacholder et al. 1993; Brenner 1996; Spiegelman et al. 1997), exposure misclassification has not been thoroughly evaluated in proximity-to-roadway or LUR studies.

A primary aim of LUR is to examine the contribution of local (versus urban- or regional-scale) sources to measured concentrations of pollutants. The approach provides insights similar to those of source apportionment, differentiating the contribution of local traffic from that of other sources. A common approach to source apportionment involves the use of factor analysis and related statistical techniques, often applied to particle-constituent data (Thurston and Spengler 1985; Laden et al. 2000; Lall and Thurston 2006; Ogulei et al. 2006; Pekney et al. 2006; Zhao et al. 2006; Rizzo and Scheff 2007; Lee et al. 2008). Unlike LUR studies, many source apportionment studies rely on temporally dense measures at a limited number of fixed monitoring sites, in which common temporal variations provide insight about predominant sources. While this approach facilitates determination of key source categories, it does not provide insight about intraurban spatial variability, as does LUR. Some recent source apportionment studies have begun to incorporate more sampling sites, to

account for some spatial and temporal variability in $PM_{2.5}$ composition. However, most have relied on a small number of monitoring sites spread over a large geographic area (Ito et al. 2004; Chan et al. 2008), and source apportionment studies covering a greater number of sites have not drawn inferences about spatial variability in source contributions (Larson et al. 2004). Thus, while factor analysis studies have developed key insights about broad source contributions, they have generally not been formally integrated with LUR or other techniques that could facilitate the determination of source-specific exposures across individuals for epidemiologic studies. A recent study (Ryan et al. 2007) used receptor modeling to determine diesel-source signatures, which were then used in LUR models, but did not consider other source categories or capture the residential indoor environment. Combining GIS-based LUR with factor analysis could increase the interpretability of each method, providing greater insight about how source contributions may vary over space and time within an urban area. This combined analysis can, in principle, apply to both outdoor and indoor environments. In the latter case, the analysis would involve applying both land-use terms and information from a home-activities questionnaire to outputs from a factor analysis in order to differentiate the contributions of outdoor and indoor sources, respectively.

SPECIFIC AIMS

Given these issues, the primary objectives of our study were to examine distributions of indoor and outdoor residential concentrations of multiple traffic-related air pollutants in an urban setting, using GIS-based covariates, home-activities-questionnaire data, and LUR and factor analytic methods to better infer source contributions. Our specific aims were as follows:

1. Using monitoring data collected as part of a birth cohort study in Boston, Massachusetts, to develop GIS-based LUR models with which to explain spatial and temporal variability in residential outdoor concentrations of $PM_{2.5}$, EC, and NO_2.
2. Using constrained factor analysis on particle-constituent concentrations and NO_2 measures, to determine source types contributing to variability in residential outdoor concentrations.
3. To apply LUR modeling techniques to factor scores estimated through constrained factor analysis on residential outdoor concentrations, to determine whether GIS covariates and other predictors explain factor variability and thereby support initial factor interpretations.
4. Using indoor and outdoor residential monitoring data, to develop physically interpretable regression models exploring indoor–outdoor relationships for EC, $PM_{2.5}$, NO_2, and particle constituents, relying on questionnaire data to characterize indoor sources and ventilation.
5. To develop GIS-based LUR models including terms for indoor sources and ventilation that would be available for all participants in an epidemiologic cohort and that could be used to explain spatial and temporal variability in residential indoor concentrations of $PM_{2.5}$, EC, and NO_2.
6. Using constrained factor analysis on particle-constituent concentrations and NO_2 measures, to determine source types contributing to variability in residential indoor concentrations.
7. To apply LUR modeling techniques to factor scores estimated using constrained factor analysis on residential indoor concentrations, to determine whether GIS covariates and other predictors explain factor variability and thereby support initial interpretations of factor loadings with respect to indoor- and outdoor-source types (i.e., initial factor interpretations).
8. To use a simulation framework to assess the degree to which the exposure misclassification resulting from the use of a proxy variable for exposure might influence epidemiologic study findings.

We hypothesized that (1) we would observe significant outdoor spatial variability in EC and NO_2 but not in $PM_{2.5}$; (2) the combined application of LUR and factor analysis to outdoor concentrations would provide interpretable LUR models for multiple latent sources (including traffic and nontraffic terms); (3) indoor–outdoor relationships would vary across homes and pollutants and would be partly explained by questionnaire data; and (4) combined LUR and factor analysis approaches would be applicable to indoor residential concentrations measures and could help to separate the contribution of indoor and outdoor sources to indoor residential concentrations.

METHODS AND STUDY DESIGN

PARTICIPANT SELECTION

The air pollution measurements and home characteristics and occupant behavior data were collected from 2003 to 2005 as part of the Asthma Coalition for Community,

Environment, and Social Stress (ACCESS) study, a prospective birth cohort study that recruited pregnant women throughout the Boston metropolitan area with the overarching goal of assessing the potential role of social factors (maternal stress, exposure to violence), genetic factors, and environmental factors (traffic-related pollution and indoor allergens) in the etiology of asthma. The study was approved by the Human Studies Committees at the Brigham and Women's Hospital and the Harvard School of Public Health, and all participants signed consent forms for both the overall cohort study and for participation in the air pollution substudy. All consent and information materials were made available in both English and Spanish.

For the air pollution portion of the study, sampling was conducted at a subset of homes, with the aim of developing models to explain between-home variability and, ultimately, to permit extrapolation of findings to the full cohort (or other geographically and demographically similar populations). The target sample size was 40 homes, each to be sampled in both the heating season (December–March) and the cooling season (May–October). This target sample size was based on both logistical considerations and the sample size in previous LUR studies that have demonstrated robust regression models linking outdoor concentrations of the pollutants of interest with traffic characteristics (Briggs et al. 2000; Brauer et al. 2003). Ultimately, 44 homes were included in the air pollution substudy.

To ensure that a range of exposures were observed within our subsample and to maximize our statistical power, we selected participants following a predesignated sampling scheme. The homes of all participants enrolled in the cohort study at the time our sampling began were first individually geocoded with ArcGIS 9.1 software (Redlands, CA) using the year 2000 U.S. Census TIGER (topologically integrated geographic encoding and referencing system) files and City of Boston 2002 street parcels data. We then calculated GIS-based estimates of traffic density, with the goal of choosing participants who would capture heterogeneity in traffic exposures both across all participants and within neighborhoods of interest. Using the Spatial Analyst extension of ArcGIS 9.1, traffic-density scores were calculated for 50-m raster cells, using traffic-volume data (provided by the MHD) for all road segments within 100 m of the participant home. For kernel-weighted traffic scores, a kernel technique (the quadratic inverse-distance weighting function) was applied to more heavily weight road segments near the cell's center. The result was a 100-m, kernel-weighted traffic-density score for each candidate home. While alternative traffic measures could have been explored and were ultimately used in our LUR modeling, we focused on a single measure for selecting participants that was commonly used in prior studies and that was expected to capture key aspects of proximity to traffic. Based on their traffic density scores, homes were divided into tertiles and selected to represent a range of traffic densities and neighborhoods, with an aim to oversample in neighborhoods where recruitment was occurring, but also to include representative sites across Boston to ensure generalizability of the study.

Cohort recruitment for the ACCESS study was not complete when air pollution sampling began in 2003. In particular, certain neighborhoods targeted for future recruitment (through community health centers, hospitals, and Women, Infants, and Children [WIC] programs) were not yet adequately represented in the database. To capture these additional neighborhoods and to expand the available sample size, we added a convenience sample of participants from outside of the cohort (42% of our participants). As shown in Figure 1, the noncohort participants filled in some spatial gaps and helped us avoid concentrating a disproportionate number of measurements in East Boston, where early cohort recruitment was focused. There were

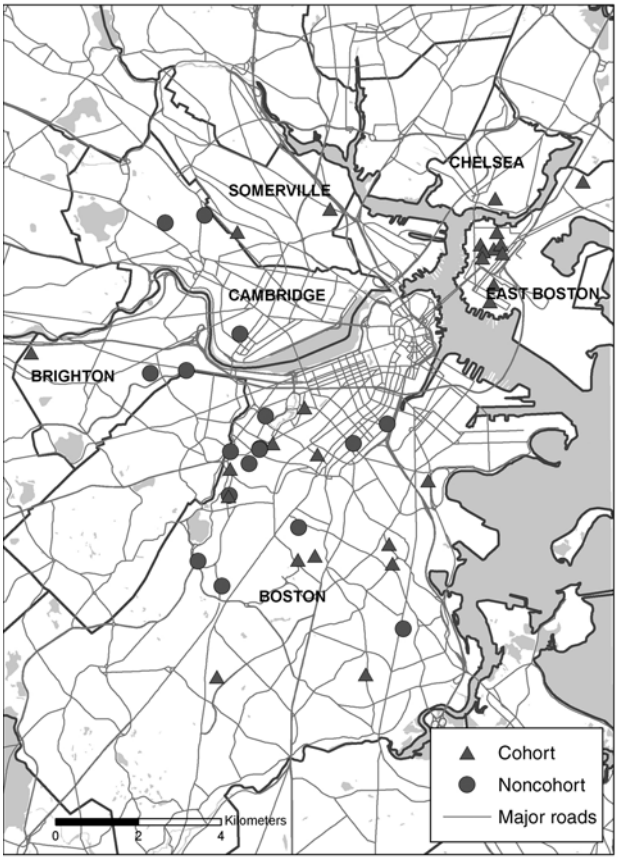

Figure 1. Map of the air pollution sampling sites with both indoor and outdoor concentration data.

some demographic and behavioral differences between cohort and noncohort participants, which are described in more detail in the Results section, under Participant Characteristics. Because the cohort recruitment ultimately expanded beyond the original design, to capture a broader array of neighborhoods and a range of socioeconomic statuses, we expected that our inclusion of noncohort participants, whose demographic and behavioral characteristics differed from those of the original cohort participants, would make our models more, rather than less, applicable to the full cohort and to other populations of interest.

SAMPLING PROTOCOL

The sampling was conducted in both the cooling (May–October) and heating (December–March) seasons to capture seasonal effects on concentrations. Indoor and outdoor 3-to-4-day samples of NO_2 and $PM_{2.5}$ were collected simultaneously at each home in both seasons. This sampling period was selected empirically to avoid limit-of-detection (LOD) issues and sampler overload. When possible, two consecutive measurements were collected at each home in each season, providing 1-week average concentrations and minimizing weekday versus weekend effects. Thus, four indoor samples and four outdoor samples would optimally be collected for each home; however, because of losses to follow-up and logistical issues, two consecutive measurements were not always collected. Not all homes could be sampled in the same time period because of limited sampling equipment. Note that in this report, we define "sampling session" as a measurement period within a given season, for which there is either one or two measurements, and for which questionnaire data describing activities during the session are available.

$PM_{2.5}$ samples were collected using a Harvard Personal Environmental Monitor (PEM), a size-selective inertial impactor attached to a Medo linear-piston vacuum pump (Hanover Park, IL) running at 4 L/min. Particles were captured on 37-mm Teflon filters, and an elutriator (10 cm long, 5 cm in diameter) was attached to the inlet of each PEM. Sampling, preparation, and analysis procedures followed standard protocols, which have been described previously (Marple et al. 1987). NO_2 samples were collected using Yanagisawa (Avantec MFS, Dublin, CA) passive filter badges, which absorb NO_2 in a triethanolamine solution on a cellulose fiber filter (Yanagisawa and Nishimura 1982).

Indoor samples were taken in the main living space of the home, away from windows, stoves, and heat sources. Outdoor samples were taken on a freestanding tripod whenever possible, located away from the home and immediate sources of emissions (e.g., driveways, grills, locations where smokers gather). When an appropriate outdoor space was not available, to avoid the building envelope, samplers were extended from a window on a three-foot sampling arm made of polyvinyl chloride piping. A four-inch-deep stainless steel rain dish was placed over both indoor and outdoor samplers to reduce interference by curious children living in the homes and by rain, snow, and wind. Pumps and continuous-monitoring equipment were placed inside a soundproofed plastic enclosure, minimizing the intrusiveness of the sampling, important given the placement of the samplers. Additional measurements included indoor and outdoor temperature and humidity monitored with a HOBO device (Onset Computer Corporation, Pocasset, MA) and continuous traffic counts recorded directly on the most heavily trafficked road within 100 m of the home using a TRAX I Plus traffic counter (JAMAR Technologies, Horsham, PA). Traffic counts were generally conducted in the cooling season only, as the tubing used by the TRAX I Plus does not perform well in snow (given potential moisture entry and damage caused by snowplows).

LABORATORY ANALYSIS

NO_2 samples were analyzed using spectrophotometry. EC concentrations were estimated using reflectance analysis on the Teflon particle filters, a nondestructive process that provides measurements highly correlated with concentrations measured using thermal-optical methods (Kinney et al. 2000). We used the M43D smoke stain reflectometer (Diffusion Systems Ltd., London, UK) with the absorption coefficient calculated in accordance with ISO 9835, as described previously (Cyrys et al. 2003).

To allow for source apportionment, elemental analysis was conducted by XRF on the particle filters (Dzubay et al. 1988). The analyses were conducted according to standard operating procedures at the Desert Research Institute laboratories (DRI; Reno, NV) and included their quality-control and quality-assurance measures (Chow and Watson 1998; Watson et al. 1999). Subsequent to XRF analysis, the filters were analyzed by high-resolution ICP–MS, to estimate concentrations of water-soluble metals (PQ Excell; Thermo Elemental, Franklin, MA). The Wisconsin State Laboratory of Hygiene (Madison, WI) conducted the ICP–MS analysis following previously documented protocols (ESS INO Method 400.4; U.S. Environmental Protection Agency [U.S. EPA] Method 1638) as described in more detail elsewhere (Sutton and Caruso 1999). The ICP–MS analysis was conducted to yield constituent concentrations at lower LODs (though capturing only the water-soluble fraction of each constituent) and to provide a comparison with XRF outputs. Note that the ICP–MS analyses were run on a

magnetic-sector ICP–MS, using appropriate element/isotope-specific mass resolution settings that ensured that all recognized isobaric interferences were mass resolved.

QUESTIONNAIRE DATA

A standardized questionnaire was administered at the end of each sampling period to gather housing characteristics and occupant-behavior data. Our questionnaire was derived from a questionnaire used in the Inner-City Air Pollution substudy within the Inner-City Asthma Study (Wallace et al. 2003) and the U.S. EPA's Residence Survey and Daily Follow-Up Questionnaire (Williams et al. 2003). Home information included location and building type, year built, and type of heating and cooking fuel. Occupant activities included smoking, opening windows, time spent cooking, cleaning activities, and use of candles, air fresheners, and air conditioning. Each factor indicates possible indoor pollution sources or influences on air exchange, penetration efficiencies, or resuspension of pollutants. The questionnaire administered to all air pollution study participants is provided as Appendix D (available on the HEI Web site). Additional information (including heating fuel type, number of rooms in the home, and building age) was gathered from the prenatal questionnaire administered to all ACCESS cohort members; the same questionnaire was also administered to noncohort participants in the air pollution study (provided as Appendix E, available on the HEI Web site).

ADDITIONAL DATA COLLECTION

During each sampling period, field staff completed field logs noting the locations of the indoor and outdoor monitors, the building type, the presence of key sources within 100 m of the monitor (e.g., idling vehicles, construction activities, evidence of grilling or smoking), carpeting or clutter in the apartment, and other observed factors that may have influenced indoor or outdoor concentrations or resuspension. Digital photos were taken of all monitors to corroborate monitor-placement data.

For our regression analyses, we needed data to account for temporal variability in background concentrations of pollutants and for meteorology because we could only sample at a limited number of homes at a time. We therefore gathered central site monitoring data from a Massachusetts Department of Environmental Protection (MassDEP) monitor in the Roxbury neighborhood of Boston, centrally located within our sampling zone (Figure 1). This was also the only monitor in Boston that had hourly data for NO_2, $PM_{2.5}$, and EC; however, we used data from other area monitors to test the sensitivity of our findings to choice of central site monitor. Hourly NO_2 was measured at the Roxbury MassDEP monitor using the TECO 42C with a chemiluminescence analyzer. Hourly $PM_{2.5}$ was measured using the Met-One BAM (Met One Instruments, Grants Pass, OR) with a $PM_{2.5}$ sharp cut cyclone beta-attenuation mass monitor. The MassDEP monitor measured EC by optical absorption with the Magee Scientific AE22ER aethalometer (Magee Scientific Co., Berkeley, CA), whereas we used reflectance analysis to estimate EC from our particle filters. The relationship between aethalometer and reflectance measures of EC has previously been shown to differ by season in Boston, with aethalometers reading higher in the summer and lower in the winter (Allen 2006). We consider this potential seasonal bias formally in our regression models.

We gathered meteorological data from the same central site, including hourly wind speed and direction, to derive indicators of meteorological conditions that may modify source–concentration relationships. To provide covariates more applicable to the traffic sources of interest, we estimated mean wind speed and direction during daytime hours (6 am–9 pm) when higher traffic volume was expected. We additionally created several wind parameters in relation to traffic sources (e.g., percent of sampling hours when the home is downwind from the nearest road), such that the statistical significance of the wind parameter in

Table 1. Meteorological Covariates and Other Potential Modifiers of Traffic–Pollution Relationships in LUR Models

	Units	Distribution[a]
Home Characteristics		
Obstructed from road	Yes / No	57% / 43%
Obstructed from major road	Yes / No	84% / 16%
Sampling Period Characteristics		
Hours downwind from major road	Percent (%)	0% / 50% / 100%
Average wind speed during daytime sampling hours (6AM–9PM)	m/sec	3.0 / 5.2 / 11
Daytime hours with winds < 2 m/sec	Percent (%)	0% / 0% / 37%
Weekend sampling days	Percent (%)	0% / 40% / 67%
Floor (monitor height)	(Categorical: 1, 2, 3+)	34% / 34% / 31%
Snow during sampling period	Yes / No	3% / 97%

[a] Values represent minimum/median/maximum for continuous variables or the percentage within various categories for categorical variables, estimated across measurements.

Table 2. Traffic Indicators Examined for GIS-Based LUR Models[a]

Indicator	Minimum	Median	Maximum
Cumulative Density Scores (Vehicle-Meters/m^2/Day)			
Unweighted traffic density			
Within 50-m buffer	4.1	43.9	197.5
Within 100-m buffer	3.8	55.9	186.7
Within 200-m buffer	10.7	64.3	570.7
Within 300-m buffer	12.7	73.1	479.2
Within 500-m buffer	26.7	80.9	347.3
Kernel-weighted traffic density			
Within 50-m buffer	4.3	29.5	194.0
Within 100-m buffer	5.8	52.8	168.3
Within 200-m buffer	11.7	59.0	479.4
Within 300-m buffer	12.1	66.0	927.8
Within 500-m buffer	23.1	73.5	440.4
Density of larger roads (> 8500 cars/day) within 200-m buffer	0	5.2	21.9
Summary Measures			
Total roadway length (m)			
Within 50-m buffer	63.3	427.9	940.6
Within 100-m buffer	498.4	1074.3	1,931.4
Within 200-m buffer	1950.5	3395.3	7,468.9
Within 300-m buffer	4529.3	6632.2	16,353.6
Total ADT × length within 200-m buffer (vehicle-meters/day)	1.24E6	8.97E6	7.31E7
Distance-Based Measures (m)			
Distance to nearest larger road (> 8500 cars/day)	15.0	64.2	353.8
Distance to nearest major road (> 13,000 cars/day)	16.7	66.4	373.4
Distance to nearest highway (> 19,000 cars/day)	34.9	216.4	1,394.6
Distance to nearest designated truck route	91.7	608.6	2,951.1
Characteristics of Nearest Major Road			
ADT (vehicles/day)	8900	8,900	117,720
ADT/distance to major road (vehicles/day)/m	23.8	236.0	764.4
Diesel fraction (%)	1.4%	5.9%	14.5%
Average daily truck traffic (vehicles/day)	125	703	12,637
Average daily truck traffic / distance to major road (vehicles/day)/m	0.57	21.4	61.6

[a] Distributions are based on sampled sites ($N = 44$).

the regression model implies source importance. We calculated the percent of daytime hours with mean wind speed below 2 m/second and the harmonic mean mixing height for the sampling period, estimated from twice-daily measurements of mixing height obtained from the National Climatic Data Center. Snowfall during the sampling period, which could have physically obstructed the sampler, was also considered. Site characteristics examined as possible modifiers of the source–concentration relationship included the presence of an obstruction (generally a building) between the monitor and nearest road and nearest major road and monitor height (either the level of the ground or the floor at which the outdoor sampling occurred). Meteorological covariates and other potential modifiers of traffic–pollution relationships are summarized in Table 1.

Road networks and traffic data were obtained from the MHD. Because different aspects of traffic (including density, roadway configuration, and average vehicle speed) may affect emission rates, pollutant mix, and dispersion, and because different indicators have been predictive of concentrations or health outcomes in different studies, we opted to create a suite of traffic indicators (Table 2) capturing varying aspects of traffic. We built raster-based cumulative density scores (vehicle-meters/m^2/day) for average

daily traffic (ADT) counts within radii of 50 to 500 m from each home. Because roadway segments nearer to the home may have greater influence on concentrations, we also explored inverse-distance quadratic functions (kernel-weighted buffers) for the same radii, which capture standard Gaussian dispersion concepts. As traffic counts on smaller residential roads were sparse within the MHD data set for our modeling domain, we created cumulative density scores including only larger roads (above 8500 cars/day), summary measures of total roadway length within radii of 50 to 300 m from each home, and the product of roadway length and ADT counts within 200 m of each home. We considered the distance to various roadway types, including to the nearest larger road (greater than 8500 cars/day), major road (greater than 13,000 cars/day), highway (greater than 19,000 cars/day), and designated truck route. To explore the influence of major roads on nearby neighborhoods, we created indicators of the roads' ADT and diesel traffic (estimated using axle-length data collected by the TRAX I Plus) and weighted each by the home's distance to the road.

We developed and examined, in a subset of analyses, two meteorology-weighted covariates hypothesized to better capture the effect of local traffic sources:

1. Total roadway length within 100 m of the home, divided into eight directional wedges from the home (using the Aspect function in ArcGIS 9.1), weighted by the percent of time during each sampling period when the predominant wind was from each direction
2. A similar wind-direction-weighted term, using total traffic (roadway length × vehicles/day) within 100 m of the home, rather than roadway length

We considered other GIS covariates that may be associated with traffic or other local pollution sources. Measures of population and area (m^2) at the block-group level were used to estimate block-group population density and mean population density within 200 m of the home. These metrics can serve as proxies of traffic density or of residential combustion activities. NLCD-50 land-use categories and elevation data were downloaded from the U.S. Geological Survey's National Land Cover Dataset and the National Elevation Dataset, respectively (both at *http://gisdata.usgs.gov*). From field logs compiled during the sampling period, we constructed covariates for smoking or grilling near the outdoor monitor and for construction within 100 m of the home. We used U.S. Census 2000 data to determine the number of homes using residential fuel oil within 200 m of the sampled home, as a proxy for the intensity of local home-heating-oil use.

Finally, we gathered an array of housing characteristics through the City of Boston, Brookline, Cambridge, and Somerville property tax records. This information was meant to supplement questionnaire data and to potentially make it possible to develop indoor exposure models without having to administer questionnaires, which would streamline the model-building process. Covariates available included land-use type, year the property was built, living area, building style and materials, type of heating system, and presence or absence of air conditioning.

QUALITY CONTROL AND QUALITY ASSURANCE

Field blanks were collected totaling approximately 10% of the number of samples. Field blanks were transported and handled like regular samples, but the filters were not attached to pumps. For NO_2, $PM_{2.5}$, and EC, these samples were used to determine background contamination and for the calculation of method LODs. Sample concentrations for each pollutant were blank corrected using the mean field blank value, when this value was significantly different from zero at the 95% level. The method LODs were calculated to be three times the standard deviation of the field blanks. For the elemental concentrations determined by XRF and ICP–MS, the laboratories provided concentration uncertainties equaling one standard deviation of error estimates based on analytical precision, for each element and sample. As in previous studies reporting aerosol elemental concentration data, three times this measure of uncertainty was considered to be the LOD for each element and sample (Long and Sarnat 2004). Precision was determined using duplicate samples (with a target of 10% of samples) and calculated as the mean relative difference, or the average of the absolute difference of a pair of duplicates divided by its mean. Finally, questionnaire responses were evaluated for completeness and, where applicable, validated with data available through the City of Boston, Brookline, Cambridge, and Somerville property tax records.

ADDITIONAL AIR POLLUTION SAMPLING

One of the inherent limitations of our analysis (and any intensive multipollutant residential investigation) is that only a limited number of measurements could be collected concurrently, complicating the distinction between spatial and temporal variability. This issue is particularly significant in the development and interpretation of latent-source factors derived from particle-constituent data, as the correlations observed across constituents may be influenced by the sampling design.

To examine whether our factor analysis findings may be influenced by our study design, we analyzed samples collected as part of a toxicological study at the Harvard School of Public Health, which collected particle samples on multiple days but at a single site (near Huntington Avenue in Boston, a significant source of traffic emissions). Most prior studies using factor analysis to examine source apportionment have primarily leveraged temporal variability at fixed-site monitors, rather than spatial variability across urban locations. Use of the Huntington Avenue site data therefore allowed for a qualitative comparison of whether similar factors (e.g., sources) would have been identified in urban Boston using a traditional (single-site) factor analysis study design, as were identified in an approach using data from residences distributed across the city.

For this temporal analysis substudy, 5-hour, daily-integrated measures of $PM_{2.5}$ were collected between 10 am and 4 pm from January to February 2008, using the Harvard Ambient Fine Particle Concentrator (HAPC) system. This system uses three successive stages of virtual impactors with cut points in the range of 0.15 μm, which enrich the mass concentration by a factor of about 30 (Sioutas et al. 1995) without substantially changing its physicochemical characteristics (Sioutas et al. 1997). Ambient $PM_{2.5}$ samples were collected from a sampling manifold located downstream of the HAPC size-selective inlet, on 47-mm Teflon filters at a flow rate of 30 L/min. Concentrated ambient particle (CAPs) samples were collected from a post-concentrator sampling manifold located at the entry to an attached animal-exposure chamber, on 47-mm Teflon filters at a flow rate of 3 L/min. To evaluate EC and organic carbon (OC), samples were collected from the same manifold location but on quartz fiber filters and at a flow rate of 2 L/min.

Both ambient-particle and CAPs filters were analyzed following protocols identical to those described for the primary study, including XRF and ICP–MS analyses in sequence (see Laboratory Analysis, under Methods and Study Design). In principle, the ambient-particle samples from this substudy were most directly comparable to the outdoor (ambient) samples collected in the main study. However, because the sampling periods for the substudy were short (5-hour), leading to potential LOD issues for the ambient-particle samples, and because particle composition has been shown to be minimally affected by concentration (Sioutas et al. 1997), the CAPs filters provided a stronger basis for comparison. The quartz filters were analyzed for EC and OC concentrations using thermal-optical transmittance (Sunset Laboratory, Tigard, OR).

STATISTICAL METHODS AND DATA ANALYSIS

In this section, we describe our data analysis strategy, starting with an overview of the logical flow of these analyses. We first conducted exploratory indoor–outdoor modeling for individual pollutants (see Indoor–Outdoor Concentration Modeling in this section) to estimate effective penetration efficiencies, identify potential indoor sources, and characterize pollutants as dominated by either indoor or outdoor sources. For outdoor concentrations (see LUR Modeling — Outdoor NO_2, EC, and $PM_{2.5}$ in this section), we developed LUR models predicting concentrations of three commonly modeled pollutants ($PM_{2.5}$, EC, and NO_2). We subsequently extended these models for $PM_{2.5}$, EC, and NO_2 to indoor concentrations (see LUR Modeling — Indoor NO_2, EC, and $PM_{2.5}$ in this section), adding questionnaire data and ventilation proxies as predictors. We then moved beyond these three pollutants by applying constrained factor analyses to capture latent-source effects both outdoors and indoors (see Constrained Factor Analysis in this section). We developed LUR models for outdoor factor scores (see LUR Modeling — Outdoor Factor Scores in this section), and modeled indoor factor scores using outdoor-source terms as well as questionnaire data and ventilation proxies (see LUR Modeling — Indoor Factor Scores in this section). Finally, we developed a simulation approach to determine the implications of exposure misclassification on epidemiologic study findings (see Exposure Misclassification Analysis in this section).

INDOOR–OUTDOOR CONCENTRATION MODELING

In these analyses, we use the core principles of the mass-balance model to examine relationships among indoor and outdoor pollutant concentrations, indoor sources, and ventilation. Our initial analyses of indoor concentrations were compared against the paired residential outdoor measurements, to better understand the effects of ventilation characteristics and indoor-source contributions. Subsequent indoor-concentration models (see LUR Modeling — Indoor NO_2, EC, and $PM_{2.5}$, below) relied on central site monitoring data, GIS covariates, and questionnaire data to predict concentrations, given that measurements at each home (indoors or outdoors) would be unavailable for all participants in most large cohort studies. The first set of models fulfilled two key purposes: (1) it identified candidate source variables for subsequent models, and (2) it allowed characterization of pollutants as primarily influenced by either indoor or outdoor sources. As such, these analyses served as precursors to the indoor LUR and factor analysis models for indoor constituents.

The relationship between indoor and outdoor concentrations is theoretically characterized, under steady-state conditions, with a single-compartment mass-balance model, as in equation (1):

$$\begin{aligned} Cin_{ij} &= \frac{P_j a_i}{a_i + k_j} Cout_{ij} + \frac{Q_{ij}/V_i}{a_i + k_j} \\ &= F_{INFij} \times Cout_{ij} + \frac{Q_{ij}/V_i}{a_i + k_j} \end{aligned} \quad (1)$$

where Cin_{ij} is the indoor concentration of pollutant j for sampling session i; P_j is the penetration efficiency (dimensionless) for pollutant j; a_i is the air-exchange rate for sampling session i (h^{-1}); k_j is the decay rate for pollutant j (h^{-1}); $Cout_{ij}$ is the outdoor concentration; Q_{ij} represents the various individual indoor-source terms; and V_i is the house volume (m^3). Note that, for some indoor analyses, we averaged both samples within each sampling session rather than all individual samples, because the questionnaire data capture source activities for the entire week. Within the formulation of equation (1), the subscript i captures both participant and time period. The quantity $P_j a_i / (a_i + k_j)$ is also called the "infiltration factor" (F_{INFij}) and represents the fraction of $Cout_{ij}$ that penetrates indoors in each home. Therefore $F_{INFij} \times Cout_{ij}$ describes the ambient contribution to the indoor concentration, and $(Q_{ij} / V_i) / (a_i + k_j)$ represents the contribution of indoor sources.

While our study measured indoor and outdoor concentrations and collected questionnaire data that could proxy for indoor-source terms, air exchange rate (AER) measurements were not logistically feasible. Thus, we needed to develop a reasonable proxy of ventilation. F_{INF} is directly associated with AER, especially if P and k are relatively less variable across homes, so a proxy variable for the infiltration factor provides similar information as a proxy variable for AER. This has been corroborated in previous Boston-area residential studies, which documented strong relationships between F_{INF} and AER (Long and Sarnat 2004).

Prior studies have used sulfur indoor–outdoor (I/O) ratios to represent the infiltration factor, as there are generally few residential indoor sources of sulfur (see, for example, Sarnat et al. 2002). Equation (2) below illustrates that, in the absence of indoor sources, the mass-balance model in equation (1) reduces to an expression linking indoor concentrations to outdoor concentrations via the infiltration factor.

$$\begin{aligned} Cin_{iS} &= F_{INFiS} \times Cout_{iS} + \frac{Q_{ij}/V_i}{a_i + k_S},\ Q_{ij} = 0 \\ &\Rightarrow F_{INFiS} = \frac{Cin_{iS}}{Cout_{iS}} \end{aligned} \quad (2)$$

where F_{INFiS}, Cin_{iS}, and $Cout_{iS}$ are the individual infiltration factor, and the indoor and outdoor concentrations of sulfur, respectively. We empirically tested the assumption that there were no significant indoor sources of sulfur by regressing indoor sulfur concentrations against outdoor sulfur concentrations, and by testing whether the intercept was significantly different from zero. In addition, we examined whether any questionnaire information with a plausible link to sulfur emissions significantly predicted indoor concentrations. For our models, we created categorical infiltration-factor indicators, rather than using the actual sulfur I/O ratio, because of the ratio's instability at higher values and uncertainty about P and k (both their mean values and variance across homes), and because effective penetration efficiencies vary by particle constituent.

With this information, we used a sequential model-building approach for our indoor–outdoor analyses. First, for all particle constituents and NO_2, we regressed Cin_{ij} on $Cout_{ij}$. For a subset of our indoor–outdoor analyses, our primary objective was to identify potential indoor sources of pollutants for subsequent regression analyses, so we next introduced potential indoor sources and source-related activities. For other indoor–outdoor analyses, our primary objective was to determine the fraction of indoor concentrations attributable to indoor sources, so for these analyses we instead regressed Cin_{ij} on the product of $Cout_{ij}$ and a categorical infiltration factor. Each model also contained an independent linear indoor-source term.

To minimize potentially spurious associations, we focused on a subset of questionnaire information that had a logical causal connection to the generation or resuspension of investigated pollutants, although we examined a broader array of correlations to determine the consistency and coherence of our data set. For example, gas stoves have been shown to be a source of NO_2 (Lee et al. 1998) and other combustion pollutants, with their effect influenced by the presence or absence of a pilot light as well as the amount of time spent cooking per day. Cooking activity, especially frying, grilling, and burning food, has been associated with $PM_{2.5}$ and numerous particle constituents (Ozkaynak et al. 1996). Humidifier use has been associated with elevated $PM_{2.5}$, especially for those elements characteristic of the "tap water fingerprint" (e.g., Ca, Cl, K, Si, and S) (Highsmith et al. 1992). Candle use is a potential source of EC (Wallace 2005). Some cleaning activity variables (e.g., vacuuming) could indicate either removal or sources of suspended particles, especially for crustal elements (Wallace 1996). We created an occupant-density variable (number of occupants/number of rooms) to capture resuspension related to foot traffic and greater overall source activity relative

to home volume, and we considered the number of occupants and the number of rooms as separate covariates.

After investigating indoor pollutant sources, we considered our categorical infiltration factor as a potential modifier of indoor-source terms, following the general structure of the mass-balance model shown in equation (1). The resulting structure of our pollutant-specific models follows equation (3) below:

$$Cin_{ij} = \beta_{0j} + \beta_{1j} \times Cout_{ij} \times Ventilation_i + \beta_{xj} \times Q_{ij} \times Ventilation_i \quad (3)$$

where $Ventilation_i$ is the categorical variable derived from the individual infiltration factors (F_{INFiS}). Due to the limited statistical power of our study, we could not incorporate multiple source terms with interactions, but we explored this model for insight on influential sources and the modifying influence of $Ventilation_i$ on indoor- and outdoor-source terms.

LUR MODELING AND CONSTRAINED FACTOR ANALYSIS

While the evaluation of indoor–outdoor relationships based on home-specific measurements is informative, characterizing exposures for a large epidemiologic study necessitates development of models of outdoor or indoor concentrations based on information that could be readily obtained for all cohort participants. Our model-building approach differed somewhat by setting (indoors versus outdoors), pollutant, and outcome (individual pollutants versus factor scores), but included many common features. Because we had many potential predictors and a relatively small sample size, in all analyses we introduced covariates in a predefined sequence to limit spurious associations. We selected candidate variables in an initial model-building step to restrict the number of covariates introduced into multivariate regressions. Finally, for all models, we examined the implications of alternative statistical approaches to selecting covariates. As described below, differences in the model-building approach across settings, pollutants, and outcomes were principally related to the nature of the dependent and independent variables of interest.

LUR Modeling — Outdoor NO_2, EC, and $PM_{2.5}$

We did not develop LUR models for each individual particle constituent, for parsimony and to minimize the likelihood of spurious associations. Instead, we focused initially on three pollutants previously examined in LUR models — NO_2, EC, and $PM_{2.5}$ — and subsequently built LUR models for latent-source factors derived using particle constituent data.

For our LUR models of outdoor NO_2, EC, and $PM_{2.5}$, we built models separately for each pollutant, allowing different aspects of traffic, meteorology, and site-specific factors to predict concentrations of different pollutants. We first regressed residential outdoor concentrations against central site monitoring data for the specific hours that each residential sample was collected, and we used the residuals from this regression to select candidate spatial covariates. We selected candidate traffic indicators and modifiers based on their nonparametric univariate correlations with the residuals from the regression, with a generous significance criterion ($P < 0.3$) to avoid eliminating covariates that might be significant in multivariate but not univariate models.

Because traffic indicators can be highly correlated, however, we also considered a secondary covariate selection method, using a form of classification and regression tree (CART) analysis. We used the tree command in the statistical computing software R to group observed concentrations by applying an impurity criterion, to minimize within-group variance while maximizing between-group differences. We compared concentration groups created using the traffic indicators as predictors, yielding the indicators that best distinguished high- and low-pollution locations. As we wished to treat the traffic indicators as continuous, rather than binary, variables in our final LUR models, we did not use the outputs from the classification analysis directly, but those indicators selected by both correlation and classification methods were considered the strongest candidates for inclusion.

We then followed a stepwise multivariate modeling approach in which we included the following, in order: central site monitoring data, candidate traffic indicators, meteorological and site-specific modifiers as interaction terms with traffic indicators, and additional outdoor sources (e.g., grilling or smoking noted near the outdoor monitor, land-use type). While some additional terms may capture traffic-related effects, given high correlations between traffic and land-use patterns, we opted to emphasize more directly interpretable terms, and therefore focused first on the traffic indicators in Table 2. Multivariate regression models followed the general form of equation (4), below, with a maximum P value of 0.1, to retain variables at each stage.

$$Cout_{ijt} = \beta_{0j} + \beta_{1j} \times Cambient_{jt} + \beta_{2j} \times Traffic_{ij} + \beta_{3j} \times Traffic_{ij} \times Modifier_{ijt} + \beta_{4j} \times Other\ sources_{ijt} \quad (4)$$

where $Cout_{ijt}$ is the measured concentration of pollutant j at location (home) i during time t; $Cambient_{jt}$ is the mean concentration of pollutant j at the central site during time t; $Traffic_{ij}$ represents the different traffic indicators listed in Table 2 at location i, selected separately for each pollutant; $Modifier_{ijt}$ is the value of meteorological or site characteristics (listed in Table 1) that can alter the association between traffic indicators and $Cout_{ijt}$; and $Other\ sources_{ijt}$ represents additional source terms at home i for pollutant j during time t. Unlike in the modeling described above and in equation (1), in this model, we used all individual observations rather than aggregating them by sampling session, given that we were not utilizing questionnaire information here and aimed to preserve statistical power for subsequent factor analyses and regression modeling.

For these models, we examined the quantile–quantile (Q–Q) plots and Anderson-Darling test statistics for all three pollutants, to determine whether to log-transform concentrations prior to selecting covariates. Although log-transformation could potentially complicate comparisons between the models or with the subsequent factor analysis models, we wanted to initially consider the most statistically robust formulation of the models, to allow for comparisons with models that may have greater physical interpretability. We tested the sensitivity of our findings to the effect of log-transformation.

Finally, as part of the model-building process, we examined scatterplots and distributions of key covariates and removed outliers and influential points that indicated either likely measurement error or observations that were not reflective of the preponderance of the data. We used the data set with influential points removed for our final models but tested the implications of their inclusion for covariate significance and directionality.

LUR Modeling — Indoor NO_2, EC, and $PM_{2.5}$

For our LUR modeling of indoor concentrations of NO_2, EC, and $PM_{2.5}$, we followed an approach that was slightly different from the one for outdoor concentrations or for our indoor–outdoor analyses. Our primary objective with this analysis was to explain variability in indoor concentrations of NO_2, EC, and $PM_{2.5}$ using terms that would likely be available for all participants in an epidemiologic study, since we then planned to use these models in the exposure misclassification analysis detailed under Statistical Methods and Data Analysis. Thus, we did not use concentrations measured outside of participant homes, but instead used publicly available traffic data and other source data, along with questionnaire responses. Because the questionnaire data spanned the entire sampling session, measurements were averaged within sampling sessions, and the theoretical functional form of these regression models was

$$\begin{aligned} Cin_{ij} = \beta_{0j} &+ \beta_{1j} \times Cambient_{ij} \times Ventproxy_{ij} \\ &+ \beta_{2j} \times Q_{ij} \times Ventproxy_{ij} \\ &+ \beta_{3j} \times Traffic_{ij} \times Ventproxy_{ij} \end{aligned} \quad (5)$$

where Cin_{ij} is the indoor concentration of pollutant j for sampling session i; $Cambient_{ij}$ is the mean concentration of pollutant j at the central site during sampling session i; Q_{ij} represents the various indoor-source terms, $Traffic_{ij}$ represents the different traffic indicators listed in Table 2 at location i, selected separately for each pollutant; and $Ventproxy_{ij}$ represents publicly available information or questionnaire responses found to be predictive of AERs. This model differs from the model represented by equation (3) in its use of central site monitoring data and GIS-based traffic covariates, rather than concentrations measured outside the home, and in its use of generally available data to represent ventilation, rather than measured indoor–outdoor sulfur ratios.

To implement this approach, we therefore have three categories of covariates that need to be developed from generally available data — traffic sources, indoor sources, and ventilation. For the traffic sources, we allowed the traffic terms to be selected separately within each model through a model-fitting process (as opposed to the ventilation and indoor sources, which were represented by terms selected in earlier models, as described in more detail below). While we could have restricted our models to the use of the traffic terms found to be optimal for our outdoor LUR model, in theory, traffic proxy variables that could predict concentrations at a specific site outside the home might be different from those that could predict concentrations at a specific site inside the home. As done previously, we chose the optimal traffic term based on the lowest P value.

Because we had many candidate traffic variables and a relatively small sample size, comparison of models on the basis of P values could have been problematic, both because multiple variables could have similar significance levels and because observed relationships could be due to chance. To provide a general indication of the weight of evidence for each traffic term and the amount of uncertainty in choosing the best model, we adopted a Bayesian approach to estimate the probability that a model using a given traffic covariate was the best model. Of note, while an identical approach could also be applied to the outdoor LUR models and the outdoor and indoor factor analysis LUR models, we focus here on the indoor models for the sake of brevity, as the conclusions are comparable. For the

same reason, we did not apply the classification analysis from the outdoor LUR models to the indoor LUR model.

We used equations (6) through (8) to calculate the posterior model probabilities for each pollutant (George and McCulloch 1997; Chipman et al. 2001).

$$P(M_k|Y) \propto l(Y|M_k) \times P(M_k) \tag{6}$$

where M_k is the model with traffic term k when all of the other variables (e.g., ambient concentrations, indoor sources) are in the model; Y is the observed indoor concentrations for one of the pollutants; $P(M_k \mid Y)$ is the posterior model probability of M_k given Y; $l(Y \mid M_k)$ is the marginal likelihood of Y given M_k; and $P(M_k)$ is the prior probability that M_k is the true model. We assumed the same prior probability, $P(M_k)$, for all traffic terms, which was equal to

$$l(Y|M_k) = \sqrt{\frac{1}{c+1}} \times \frac{1}{\left(\sum_{i=1}^{k} Y_i^2 - \frac{\left(\sum_{i=1}^{n} X_{ik} Y_i \right)^2}{\left(1+\frac{1}{c}\right) \times \sum_{i=1}^{n} X_{ik}^2} \right)^{\frac{n}{2}}} \tag{7}$$

where Y_i is the residual, from sampling session i, from regressing indoor concentrations on ambient concentrations and indoor-source terms; X_{ik} is the residual from regressing traffic term k on ambient concentrations and indoor-source terms; n is the number of observations; and c is a parameter that scales the prior variance of the regression coefficients of the traffic terms in the model, thereby reflecting prior uncertainty about these coefficients. The parameter c appears in the marginal likelihood (equation 7) when the coefficients are integrated over, and results can be sensitive to the choice of c. Based on a review of possible values for c in Chipman and colleagues (2001), we specified that $c = n$, so that c was large enough to acknowledge reasonable uncertainty in the effect estimates, while still attributing very low prior probabilities to very unlikely effect estimates. We also conducted sensitivity analyses by calculating the posterior probabilities across a range of prior uncertainties (the value of c ranged from 5 to 100) (Chipman et al. 2001).

The posterior probabilities then needed to be normalized as shown in equation (8) (and multiplied by 100 to calculate a percentage).

$$P(M_k|Y) = \frac{P(M_k|Y)}{\sum_{i=1}^{N} P(M_k|Y)_i} \times 100 \tag{8}$$

In a sensitivity analysis, we considered another model, in which M_0 (equation 9) is the model without a traffic term. We assumed $P(M_k)$ of $\frac{1}{2}$ and $\frac{1}{2N}$ for M_0 and M_k (models with the traffic term), respectively. This assumed an equal chance of traffic affecting indoor concentrations as not, a semiarbitrary assumption meant to evaluate the implications of a high prior probability of no significant traffic impact. Using the $\frac{1}{2N}$ weights in the model selection inherently penalizes for testing many traffic terms in a small data set. The posterior probabilities of M_0 for each pollutant were calculated as shown by equation (9) and normalized utilizing equation (8).

$$\begin{aligned} P(M_0|Y) &\propto l(Y|M_0) \times P(M_0) \\ &\propto \frac{1}{\left(\sum_{i=1}^{k} Y_i^2 \right)^{\frac{n}{2}}} \times P(M_0) \end{aligned} \tag{9}$$

To minimize possible spurious associations when considering indoor sources, we allowed into the model only those indoor-source terms that were previously predictive of indoor concentrations in models with outdoor concentrations of pollutants that had been measured directly outside sample homes (described above in Indoor–Outdoor Concentration Modeling under Statistical Methods and Data Analysis).

Finally, we incorporated ventilation effects. While indoor–outdoor sulfur ratios were used to indicate the infiltration factor in the indoor–outdoor analyses (described in Indoor–Outdoor Concentration Modeling), this information would not be available for all participants in a large epidemiologic study. To develop an alternative ventilation term, we first performed a sequential model-building process to explain variability in the infiltration factor we derived from indoor–outdoor sulfur ratios, using questionnaire data and property tax records. Only variables with a logical causal connection to the infiltration factor were considered (i.e., season, floor level, multiunit versus single-family dwellings, age of home, home air-conditioning use, opening of windows). We considered the infiltration factor both as a continuous and categorical variable, to identify the covariate or combination of covariates available across epidemiologic study participants that would best predict infiltration factors.

For the indoor–outdoor modeling, concentrations were aggregated by sampling session. Although many homes were visited in both the heating and cooling seasons, seasons were broadly defined and spanned up to 6 months, with questionnaire information and occupant activities varying across seasons. Given these factors and our limited statistical power, we opted to treat multiple observations

for the same household as independent measurements. As described above for the outdoor LUR models (LUR Modeling — Outdoor NO_2, EC, and $PM_{2.5}$, under Statistical Methods and Data Analysis), for the indoor NO_2, EC, and $PM_{2.5}$ LUR models, we removed a small number of outliers and influential points, and we evaluated the implications of their inclusion for our final models.

Constrained Factor Analysis

To develop the factor scores for both indoor and outdoor concentrations, we used a constrained factor-analytic approach that produces only positive factor loadings, as in nonnegative matrix factorization (NNMF) (Hoyer 2004) or positive matrix factorization (PMF). We selected this approach for our urban setting, because chemical mass-balance models would require more precise knowledge of the number of sources and their profiles (signatures) than is currently available (Bruinen de Bruin et al. 2006), and traditional factor analysis has the familiar rotation problem in which different rotation methods lead to different interpretations. We originally considered a Bayesian approach, with informative priors regarding the constituents that should be found together for a given source type, but found a general lack of convergence in fitting these models, in part due to a mismatch between the priors and the underlying data. Thus, the computational complexities of the Bayesian approach offered little advantage over more traditional frequentist approaches.

NNMF factors a data matrix into a product of two matrices having nonnegative entries, which is similar to what a singular value decomposition does, but it lacks an underlying probability model that would justify standard inference procedures and the appropriate handling of missing data. However, the most important aspect of NNMF, nonnegative loadings (i.e., sources never subtract from concentrations), is maintained simply by fitting a factor analytic model while constraining the source loading matrix to nonnegative entries. This leads to a locally unique solution, thus solving the multiple-rotation problem of traditional factor analysis, while guaranteeing numerical stability and appropriate asymptotic statistical properties. Using maximum likelihood estimation, we fit the following factor analysis model (with F factors):

$$\mathbf{z} \sim \mathrm{MVN}(0, \boldsymbol{\Lambda}'\boldsymbol{\Lambda} + \boldsymbol{\Psi}) \tag{10}$$

where $\mathbf{z}$ is a vector of z-scores corresponding to J constituent concentrations, $\boldsymbol{\Lambda}$ is a $J \times F$ matrix of positive real numbers, and $\boldsymbol{\Psi}$ is a diagonal $J \times J$ matrix with positive entries on the diagonal. To facilitate finding a global maximum, the likelihood function that corresponds to the model above was maximized using the simulated annealing (SANN) algorithm described by Belisle (1992) and implemented using the R function **optim.** Missing observations in $\mathbf{z}$ were addressed by subsetting the $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}$ matrices in a manner consistent with the observed data. That is, the submatrices of $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}$ corresponding to the observed data were used to construct the multivariate normal variance-covariance matrix corresponding to the vector of observed data. This approach entails an assumption that the missing data mechanism does not depend on the observed data.

This constrained factor analysis model uses a multivariate normal likelihood function, which raises the question of whether our data adhere to this assumption. Houseman and colleagues (2006) describe a method of checking the normality of a multivariate random vector having arbitrary correlation structure. Briefly, if $\mathbf{Y}$ is a multivariate vector having mean $\boldsymbol{\mu}(\boldsymbol{\theta})$ and parameterized covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ (where $\boldsymbol{\theta}$ represents a vector of regression and variance component parameters), then for a broad class of orthogonal projections $\mathbf{P}$ (including the identity matrix), Q–Q plots of the elements of $\mathbf{P}\sum^{-1/2}\left(\hat{\boldsymbol{\theta}}\right)\left[\mathbf{Y} - \boldsymbol{\mu}\left(\hat{\boldsymbol{\theta}}\right)\right]$ ("Cholesky residuals") will adequately capture departures from normality. $\mathbf{P}$ is chosen to diagnose specific types of nonnormality (e.g., for best linear unbiased predictors of random coefficients in mixed effects models), but in this case we can take $\mathbf{P}$ to be the identity for an omnibus normality check. Results from our normality checks are presented in Appendix F (available on the HEI Web site) and emphasize that, while our data do not adhere to multivariate normality, other transformations of the data adhere even less to a normal distribution, and our factor analyses can be expected to be unbiased, although inefficient without transformation.

For both indoor and outdoor concentrations, we restricted the factor analysis to a subset of pollutants. We considered the LOD for particle constituent data to be three times the individual sample SEs reported by the laboratory, as done previously (Long and Sarnat 2004), and excluded constituents for which fewer than 60% of the samples were above the LOD. We also excluded constituents not previously associated with one of our hypothesized outdoor-source factors of interest, which included coal combustion, fuel oil, and traffic characteristics. As a result, some contributors to $PM_{2.5}$ concentrations known to be present in Boston (e.g., sodium and chlorine related to sea salt) were not included, producing a factor analysis more focused on specific interpretable factors of interest and reducing statistical demands on our relatively small sample size.

To select particle constituents for factor analysis given our source factors of interest, we first reviewed the outdoor-source apportionment literature, focusing on a subset

of studies with measurements taken in a similar geographic area (i.e., the eastern United States) or in a laboratory or similarly controlled setting, within the past 10 years. Table 3 lists the constituents included in our factor analysis of outdoor concentrations, their hypothesized sources, and the studies that reference them. Of note, Cd, Ce, and Tl are not included in Table 3, as these were not source-specific tracers in the studies listed, but were included in our factor analyses because they have been significant in other factor analyses (Zhao et al. 2006, 2007; Yakoleva et al. 1999; Hopke et al. 2003) and had enough measurements above the analytic LOD.

We applied a similar constrained factor analysis approach to the filters collected as part of a toxicological study (see Additional Air Pollution Sampling, under Methods and Study Design), to assess whether similar correlation structures and factor interpretations (e.g., sources) would likely be observed if repeated outdoor samples were collected over time at one urban Boston location, during one season. For comparability, we used the same constituents used in our primary outdoor-concentration factor analysis, although NO_2 data were not available and EC concentrations were based on thermal-optical methods rather than reflectance.

For the indoor-concentration factor analysis, we examined the same list of constituents. This approach may have omitted some important indoor constituents, but it improved the comparability of the indoor and outdoor factor analysis results. Also, as few investigators have conducted factor analyses of indoor pollutants, we had little information to support the inclusion of more constituents, and we concluded that a shorter candidate constituent list should

Table 3. Hypothesized Sources of Particle Constituents Included in Factor Analyses of Outdoor Concentration[a]

	Traffic				Local / Regional	Long-Range Transport		
	Motor Vehicle	Brake and Tire	Soil and Road Dust	Diesel	Fuel Oil	Coal / Secondary	Steel Making	Wood/ Vegetative
EC	1, 12			**2, 3, 4, 10**				3
Al			**2, 3, 4, 5, 6**	2				
As			2					
Ba	12	**7, 11**	11					
Ca	**1, 10, 12**		**1, 3, 5, 6, 9, 11, 12**	**3, 10**				
Cr		11	11	2				
Cu		**7, 11**	2, 11	3				
Fe	**1, 8, 12**	11	**1, 2, 3, 5, 6, 11, 12**	4			**4, 8, 9**	
K	1		5, 6					**3, 4, 6**
La					8			
Mg		11						
Mn		11	2, 11, 12	4			**4, 8, 9**	
Mo		11						
Ni					**2, 3, 5, 8, 12**	8		
P	10			10				
Pb	8		11				4, 8	
S	8, 12			10		**2, 8**		
Sb		**7, 11**						
Se						**2, 6, 8**		
Si			**1, 3, 4, 6**					
Sr		14						
V	12				**3, 5, 8, 12**			
Zn	**1, 5, 8**	**7, 11**	11	3			**4, 8, 9**	

[a] Bold text indicates more frequent references. Numbers correspond to references as follows: 1 (Zhao et al. 2006); 2 (Ogulei et al. 2006); 3 (Qin et al. 2006); 4 (Rizzo and Scheff 2007); 5 (Li et al. 2004); 6 (Lee et al. 2008); 7 (Iijima et al. 2008); 8 (Hammond et al. 2008); 9 (Pekney et al. 2006); 10 (Spencer et al. 2006); 11 (Schauer et al. 2006); 12 (Lall and Thurston 2006).

Table 4. Hypothesized Sources of NO_2 and Particle Constituents Included in Factor Analyses of Indoor Concentration[a]

	Indoor Combustion				Other Indoor Sources		
	Smoking	Cooking	Gas Stove	Candles	Resuspension / Soil	Tap Water	Personal Activities / Personal Care Products / Cleaning Products
NO_2	7		**7, 9**	15			
EC	**3, 7, 14**	14	7	16			
Al		12			8		**2, 5**
Ca	3, 12	**2, 3, 12**			**1, 3, 8, 10**	11	**3, 6**
Cd	13						5
Cu					3		**5, 8, 14**
Fe	3, 14	**4, 12**			**1, 3, 10**		
K	**2, 3, 10, 14**				**1, 3, 10**	11	**3, 6**
Mn		4					
Ni							5
P							**5, 8**
Pb	3			16	3		5
S	14	3			3	11	5
Se					3		
Si	3	**2, 3, 6**			**1, 3, 8**	11	**5, 6**
V							5
Zn	3, 14	3		16	3		**3, 5**

[a] Bold text indicates more frequent references. Numbers correspond to references as follows: 1 (Yli-Tuomi et al. 2008); 2 (Zhao et al. 2007); 3 (Zhao et al. 2006); 4 (Larson et al. 2004); 5 (Hopke et al. 2003); 6 (Yakoleva et al. 1999); 7 (Lai et al. 2006); 8 (Koistinen et al. 2004); 9 (Lee et al. 1998); 10 (Wallace 1996); 11 (Highsmith et al. 1992); 12 (Ozkaynak et al. 1996); 13 (Clayton et al. 1993); 14 (Brunekreef et al. 2005); 15 (Lee and Wang 2006); 16 (US Environmental Protection Agency 2001).

reduce the likelihood of spurious correlations and uninterpretable factors. Potential indoor sources for these constituents are listed in Table 4. Note that indoor concentrations are also influenced by the outdoor concentrations of pollutants (Table 3). Whereas the studies represented in Table 3 were selected from the source apportionment or factor analysis literature, there were fewer such publications related to indoor concentrations. To select studies, we conducted a literature survey using the keywords "indoor," "source apportionment," and "$PM_{2.5}$" and included both studies directly identified through this approach and studies cited by these publications, some of which were not strictly source apportionment studies. As a result, NO_2 is included in Table 4 but not Table 3, as previous outdoor-concentration factor analyses did not include NO_2, but multiple indoor monitoring studies did evaluate sources of NO_2. Those constituents included in Table 3 but not in Table 4 were not clearly linked to any identifiable indoor sources in our literature search. It should also be noted that the multiple-source categories (e.g., resuspension and soil) in Table 4 are somewhat vaguely defined. The indoor source apportionment studies reported fewer distinctive source profiles than did the studies of outdoor sources, and some studies found combinations of constituents indoors that the investigators were unable to attribute to specific sources (Meng et al. 2007; Yli-Tuomi et al. 2008), reinforcing the difficulty of source attribution in the indoor residential environment.

For the constrained factor analyses, models allowing two to nine factors were produced, and factor loadings were examined to produce an optimal factor analysis model. In constrained factor analysis, there is no unambiguous approach to determine *F*, the appropriate number of factors. Formal criteria include the Kaiser–Guttman (K–G) criterion used in scree plots (Yeomans and Golder 1982) (with number of eigenvalues of the correlation matrix greater than one), cross-validation estimates of a cost or loss function (Hastie et al. 2001), and well-known measures of model parsimony such as the Akaike information criterion and the Bayesian information criterion (BIC), which approximate cross-validation estimates. As the SANN algorithm was computationally demanding, we focused on the K–G criterion and the BIC to guide selection

of *F*, but validated the choice using cross validation on traditional factor analysis. To conduct this cross validation, we used 20-fold cross validation, 40-fold cross validation, and n-fold cross validation (jackknife) to estimate the average negative log-likelihood (cost), with $\mathbf{\Lambda}$ and $\mathbf{\Psi}$ obtained from traditional factor analysis on each iterated training set. While we used these various statistical criteria to inform selection of *F*, we also used our judgment regarding the physical interpretability of the factors, given differences in *F* determined using the BIC and K–G criteria.

LUR Modeling — Outdoor Factor Scores

For the application of LUR models to outdoor factors (outdoor factor LUR models), as for the outdoor-concentration LUR models for NO_2, EC, and $PM_{2.5}$ (see LUR Modeling — Outdoor NO_2, EC, and $PM_{2.5}$, under Statistical Methods and Data Analysis), we aimed to develop parsimonious and physically interpretable models. However, in this case, we also hypothesized that significant LUR predictors would corroborate initial factor interpretations (which included sources other than traffic). Given this aim and lacking central site data for $PM_{2.5}$ constituents, we introduced covariates individually into the model, following an approach that differs slightly from that used for NO_2, EC, and $PM_{2.5}$.

For all factors, using a stepwise forward regression process, we first included central site data, maintaining the best fit of temporally varying $PM_{2.5}$, EC, or NO_2 concentrations. When modeling factors rather than individual pollutants, it is less obvious which temporally varying central site pollutant (if any) would be temporally related to each factor, but inclusion of a central site concentration and a binary season variable allowed for a reasonable assessment of temporal variability.

For factors hypothesized to be directly related to traffic, we considered traffic indicators followed by other source and population terms, and for those factors hypothesized to be related to sources other than traffic, other source and population terms were tested first. This sequencing reflected concerns about spatial covariance between predictors, a desire to maximize the physical interpretability of our regression models, and concerns related to our small sample size. As this sequencing could potentially bias our selection of LUR-model terms to explain each factor, we considered other model-building approaches in our sensitivity analyses, described below. For parsimony, we allowed only one covariate of each type (e.g., traffic indicators, population characteristics) and one interaction term per model.

Candidate variables for the LUR models were selected using univariate regression models predicting raw factor scores from covariates in four categories (central site concentrations, season, traffic indicators, and local source/population terms). For source categories in which no covariates attained statistical significance (using a generous significance criterion of $P < 0.25$ in this initial screening step), we considered the best candidate available as determined by *P* value and univariate R^2 value.

For the final model, we used the general form of equation (11) below, with a maximum *P* value of 0.25 to retain variables prior to testing interactions. Thereafter, *P* values below 0.1 were required for a covariate to be retained in the final models, with the exception of insignificant main source terms if they were significant in interaction terms.

$$\begin{aligned} Factor_{ijt} = \beta_{0j} &+ \beta_{1j} \times Cambient_{jt} + \beta_{2j} \times Season_{jt} \\ &+ \beta_{3j} \times Traffic_{ij} \\ &+ \beta_{4j} \times Population/Local\ sources_{ijt} \\ &+ \beta_{5j} \times Source_{ijt} \times Modifier_{ijt} \end{aligned} \quad (11)$$

where $Factor_{ijt}$ is the loading on factor *j* at home *i* during time *t*; $Cambient_{jt}$ is the mean concentration of $PM_{2.5}$, EC, or NO_2, measured at the central site during time *t*, using the best-fit pollutant for each factor *j*; $Season_{jt}$ is a dummy variable for the heating or cooling season, considered separately for each factor; $Traffic_{ij}$ represents the different traffic indicators listed in Table 2 at location *i*, selected separately for each pollutant; $Population/Local\ sources_{ijt}$ includes population density terms, fuel oil use, construction, and grilling or smoking near the outdoor monitor; and $Source_{ijt} \times Modifier_{ijt}$ represents possible interactions between any significant source terms (traffic or other local sources) and meteorological or site characteristics (listed in Table 1) that alter the association between these source terms and factor scores.

We conducted multiple sensitivity analyses on the final multivariate models. As in our prior LUR applications, we were concerned about the robustness of covariate selection. We tested this robustness in two ways: (1) by replacing selected traffic and source terms in each model with other candidate covariates in the same category, and (2) by performing random forest covariate selection using the randomForest package in R. This latter approach is a well-known enhancement of the CART procedure used via the tree algorithm (described in LUR Modeling — Outdoor NO_2, EC, and $PM_{2.5}$, under Statistical Methods and Data Analysis) and represents another statistical approach for covariate selection. As previously, due to the similarity in findings and the desire for parsimony, we did not apply all approaches to all models, but examine here the information value of the random forest algorithm.

The random forest algorithm (Breiman 2001) uses a bootstrap approach to fitting regression trees while simultaneously estimating prediction error (measured as the square of observed minus predicted). It thus generates predictions that account for nonlinearities and interactions between covariates. Covariate importance is assessed by permuting each covariate with respect to the outcome variable and to the remaining covariates and comparing the resulting prediction errors. That is, the covariate whose importance is to be assessed is shuffled, holding the other data fixed. Then the regression tree is traversed with the resulting data set to obtain a different outcome prediction with a different prediction error. If the covariate were an important partial predictor, the loss of information resulting from the permutations would produce a larger prediction error. We applied the random forest algorithm using 10,000 bootstrap iterations with six covariates (the approximate square root of 33 total covariates) selected per bootstrap. The random forest algorithm does not produce interpretable models, but ranks candidate covariates by strength of association with the outcome, and thus can corroborate univariate covariate selection.

In addition, we examined the robustness of model fit and covariate selection across seasons using interaction terms. As in prior LUR models, we removed outliers and tested model robustness to influential points, considering regression models both with and without these observations. Finally, due to high correlations across covariates and possible differences in covariate selection depending on the model-building process, we performed backwards elimination, to ensure that selected terms retained significance independent of other covariates in the model.

LUR Modeling — Indoor Factor Scores

For the application of the LUR-style models to the indoor factor scores, we followed a model-building approach that paralleled our approach to the LUR models for the outdoor factor scores (see the previous section, LUR Modeling — Outdoor Factor Scores) and incorporated aspects of the indoor–outdoor regression analyses (see Indoor–Outdoor Concentration Modeling, under Statistical Methods and Data Analysis, above). For indoor concentrations, one of the difficulties in interpreting factors is that concentrations reflect a combination of outdoor sources, indoor sources, and ventilation, such that constituents may be correlated with one another without being produced from a common source (e.g., constituents from different indoor activities may be elevated under low-ventilation conditions). Given these complexities, we conducted our analysis in two stages:

1. For our primary analysis, we performed the constrained factor analysis using indoor concentrations of NO_2 plus all particle constituents included in the outdoor-concentration factor analysis, and we constructed GIS-based regression models to explain variability in these factors. While this approach should in principle capture both indoor and outdoor sources, it relies on concentrations that are a combination of indoor and outdoor sources and may result in factors and models that are difficult to interpret.
2. To better isolate indoor-source contributions, we regressed the indoor pollutant concentrations individually against the corresponding outdoor concentrations, modified by a categorical marker for ventilation, the infiltration factor. The residuals from these regressions were interpreted as the portions of pollutant concentrations attributable to indoor sources, which we examined in a separate constrained factor analysis and subsequent regression models.

Note that the residual factor analyses referred to in the second point above could not be directly applied in an epidemiologic context (as they depend on residential outdoor measurements, which would not be available for all members of a cohort). However, they may inform interpretation of indoor-concentration factor analyses and provide insight about indoor sources. More generally, these analyses represent multiple layers of models that rely on a relatively small data set. Given this fact, the number of candidate indoor and outdoor sources, and the limited number of studies determining the signatures of indoor sources, we consider these analyses to be exploratory in nature.

For each analysis, the constrained factor analysis approach was as described in Statistical Methods and Data Analysis; in each stage, factor loadings were examined using the K–G criterion, cross validation, BIC measures of model parsimony, and factor interpretability to arrive at the appropriate numbers of factors. However, preliminary examination of the concentration data for indoor constituents indicated a number of highly influential values, which impaired the interpretation of our factor analyses. Some of these observations may have reflected laboratory errors, while others may have reflected true but highly influential points that would mask the overarching associations in our data. In this analysis, we had the advantage of having indoor- and outdoor-concentration data that were collected simultaneously, which could help us identify observations that that were not consistent with basic indoor-concentration modeling principles (equation 1). We therefore regressed indoor concentrations on outdoor concentrations and removed influential points based on Cook's

distance > $4/(n - P)$. The updated dataset was used for the constrained factor analyses for both constituent concentrations and indoor-attributable concentration residuals.

Similarly, as was done for the factor analysis LUR models of outdoor concentrations, the factor analysis LUR-style models of indoor concentrations were constructed using a forward stepwise model-building process. To minimize spurious associations, we constrained our candidate outdoor-source terms to those found to be significant for at least one factor in outdoor factor analysis LUR models. This strategy is slightly different from what we used for our indoor LUR models of EC, NO_2, and $PM_{2.5}$, in which we constrained the indoor-source indicators and allowed the outdoor-source terms to be selected in the model. We used this approach because we wanted to maintain consistency with our outdoor factor analysis LUR models, to facilitate joint interpretation of factors, and to explore the predictability of factors representing indoor sources.

We included central site monitoring data to adjust for temporal variability. Questionnaire-based indoor-source covariates were then incorporated into the model. They included indicators for indoor combustion (i.e., cooking or gas stove use for more than one hour per day, stovetop frying or grilling, presence of a pilot light, burning of candles or incense, and burning of food), occupancy as a broad indicator of indoor-source activity (including occupant density, number of people, and number of rooms), indoor resuspension (i.e., sweeping, vacuuming, and dusting, presence of carpeting, and wearing shoes indoors), and indoor cleaning activities. Note that we considered a somewhat larger set of questionnaire-based indoor-source covariates in this analysis than in the indoor–outdoor analysis of EC, NO_2, and $PM_{2.5}$, given the somewhat exploratory nature of the indoor factor analysis and our desire to examine constituent source signatures.

Candidate covariates were selected as the best-fit terms from each source category, using Pearson correlations of period-specific factor scores against outdoor-source terms in four categories (long-range transport/meteorology, traffic-related sources, population characteristics, and other outdoor sources) and indoor-source terms in four categories (indoor combustion, occupancy, resuspension, and cleaning activities). Covariates with univariate associations significant at $P < 0.3$ were considered as candidates for multivariate models.

For the multivariate models, we applied a manual, forward stepwise model-building process. For parsimony, and in light of our small sample size, we allowed only one covariate per source category and tested the modification of both indoor and outdoor sources by the previously determined proxy for ventilation (as described in LUR Modeling — Indoor NO_2, EC, and $PM_{2.5}$, under Statistical Methods and Data Analysis). We used a maximum P value of 0.2 to retain covariates prior to testing for interactions. After testing for interactions, a P value below 0.1 was required to retain each term, except for nonsignificant main effects for covariates with significant interaction terms.

We examined the sensitivity of model results to various permutations in our analytic methods. First, we tested whether hypothesized source indicators explained additional variability in final models. Second, we tested the robustness of the covariate selection approach by comparing its results to the covariates selected using random forest methods, as was done for our outdoor factor analysis LUR models. Third, we tested each regression model for robustness to repeated measures using random effects by household. Fourth, due to high correlations across covariates with a small sample size, we sequentially removed each term from the final models, to ensure that all terms retained significance independent of other covariates. Similarly, we reversed the order of terms introduced into the model, to ensure that the terms independently met the retention criteria. Fifth, we compared the final models resulting from this forward stepwise process to those produced using backwards elimination in a stepwise process beginning with all candidate covariates with $P < 0.25$, applying the same retention criteria at each step. Finally, we reexamined the relationships between selected covariates and factor scores, to ensure that results were robust to influential points.

EXPOSURE MISCLASSIFICATION ANALYSIS

Subsequent to the completion of our LUR models for residential indoor NO_2, $PM_{2.5}$, and EC (see LUR Modeling — Indoor NO_2, EC, and $PM_{2.5}$, under Statistical Methods and Data Analysis), we wanted to quantify the implications, in terms of epidemiologic results, of using exposure-surrogate models with varying measurement errors. We included the LUR models we developed — which can be considered as a validation study evaluating various exposure surrogates in a multivariate context — as well as models with less explanatory power that might be applied in the absence of validation studies, based on indicators of traffic exposure or indoor sources. We therefore considered exposure models within three different categories:

1. Our indoor-concentration LUR models that included terms for ambient concentrations, GIS-based traffic indicators, indoor-source terms, and ventilation characteristics (referred to below as “LUR models”).
2. Exposure models using a single indoor-source term or traffic indicator that was used in our LUR models (referred to below as “good exposure surrogates”).

3. Models using traffic indicators that were poor predictors of measured indoor concentrations in our study, but that have been used previously in the literature. This category represents those variables that could potentially be selected in the absence of validation data but that do not strongly predict exposure (referred to below as "poor exposure surrogates").

To conduct this analysis, we constructed a hypothetical epidemiologic study linking indoor pollutant concentrations and recurrent wheeze in the first year of life. We then calculated the estimated bias and quantified the estimated inflation of the SEs caused by the use of the various exposure models and determined the power of the study to detect statistically significant associations between exposure and health.

Data-Generation Step

First, we constructed a simulated epidemiologic data set, including exposure and health outcomes data, which was the "data-generation" step. In this step, we used "true" indoor concentrations to estimate health outcomes derived from "true" odds ratios (ORs). We then constructed an epidemiologic data set that included variables that could be used to predict indoor concentrations at all homes in the absence of measured concentrations. Because neither the true indoor concentrations nor the true ORs could be ascertained from empirical data, a simulation approach was necessary.

We assumed that the structure of this simulated epidemiologic study was similar to that of the ACCESS cohort study, with air pollution measurements collected for 1 week for a subset of participants in order to develop exposure models for all participants. The simulated study, therefore, followed a main study/validation study design, in which the main study included health outcome, questionnaire information, and GIS information for all participants, and the validation study included a subset of those participants whose homes were measured for concentrations of indoor pollutants.

To develop this simulated epidemiologic data set, we first sampled with replacement from the 193 ACCESS homes with complete data at the time of our analysis, to create 4500 simulated data sets consisting of 1000 homes each (the target sample size for the ACCESS cohort study). The participant data included all terms used in potential exposure models. For each of these homes, we applied our LUR models to each week of the baby's first year of life, to estimate indoor concentrations of pollutants, accounting for changing ambient concentrations, seasonal differences in ventilation characteristics, and other key covariates. As our LUR models only explained a portion of between-home variability in measured concentrations, we added normally distributed random variation with mean zero and standard deviation equal to the square root of the residual variance from our LUR models. These modeled concentrations were considered to be the "true" exposures in these homes — the gold standard — that our various exposure-surrogate models were intended to estimate.

Within each data set, we also used our LUR models to generate 1-week average indoor pollutant concentrations for a subset of 53 participants, representing the simulated validation studies. This structure was necessary for subsequent investigation of alternative exposure models, and conformed to the actual sampling duration and number of air pollution sampling sessions with complete data in our study. We generated simulated concentrations (i.e., we performed model-based resampling) to conform to the assumption that the validation study was a subset of the main study. Normally distributed random variation was added to these simulated concentrations as described above.

We then needed to simulate health outcomes for each participant and data set, given simulated indoor concentrations. We focused on recurrent wheeze in the first year of life, an outcome of interest for the ACCESS cohort study and other similar investigations. The probability of recurrent wheeze can be characterized by equation (12).

$$P_i = \frac{e^{(\beta_0 + \beta_X C_{in\,year_i})}}{1 + e^{(\beta_0 + \beta_X C_{in\,year_i})}} \quad (12)$$

where P_i is the probability of recurrent wheeze in the first year of life for subject i; β_0 is the baseline risk of recurrent wheeze, which we initially assumed to be 11.5% (Marbury et al. 1996); β_x is the coefficient corresponding to reported ORs for the pollutant in question; and $C_{in\,year_i}$ is the average indoor concentration of NO_2, $PM_{2.5}$, or EC during the first year of life. We considered ORs of 1.05, 1.50, and 2.00 per interquartile range (75th–25th percentiles) increase in indoor NO_2, $PM_{2.5}$, or EC concentrations, consistent with the literature (Brauer et al. 2002; Belanger et al. 2003; Ryan et al. 2005).

Using the gold standard indoor concentrations in the first year of life and equation (12), we randomly assigned a binary outcome to each participant. This approach assumed that there were no other factors affecting the health outcome. We made this assumption for simplicity, to avoid having to posit distributions for other variables and associations of those variables with health outcomes, as well as to maintain our focus on the impact of the various the air pollution exposure estimates.

Simulation of Epidemiologic Findings

The "data-generation" phase of the analysis resulted in 4500 simulated data sets consisting of 1000 children each. We assigned each child a health outcome (wheeze yes or no) and an exposure from each of the different surrogate exposure models (including the gold standard). Although the health outcomes were generated using "true" exposures, these were simulated and, in reality, this information would not be available to investigators. Actual studies would need to rely on a variety of different surrogate exposure models — exemplified by our LUR models, good exposure surrogates, and poor exposure surrogates. Note that, rather than using individual exposure surrogates directly, we produced predicted levels of specific pollutants in each case, so that the health estimates derived from good or poor exposure surrogates were on a comparable scale to those estimated with our LUR models.

Thus, the exposure misclassification analysis involved the following: (1) using the generated 1-week average indoor-pollutant concentrations in the simulated validation studies to fit the different surrogate exposure models; (2) using each fitted-model relationship to predict indoor concentrations in the first year of life for all study participants; and (3) regressing the predicted indoor concentrations derived from each exposure model on the binary health outcomes. In step 1, we fit the exposure model using the simulated validation data. For the LUR models, we used the regression model structures defined in LUR Modeling — Indoor NO_2, EC, and $PM_{2.5}$, under Statistical Methods and Data Analysis. The exposure models for good exposure surrogates and poor exposure surrogates were fit as

$$C_{in_{wi}} = \gamma_0 + \gamma_1 \times C_{ambient_{wi}} + \gamma_2 \times S_i + \epsilon \tag{13}$$

where, for each pollutant, $C_{in_{wi}}$ and $C_{ambient_{wi}}$ are the 1-week average indoor and ambient concentrations, respectively, from the validation study; γ_0, γ_1, and γ_2 are the regression coefficients; S_i is one of the exposure surrogates; and the ϵ term accounts for random error.

In step 2, we used these exposure models to predict the indoor pollutant concentrations for each subject in his or her first year of life. Carrying out this procedure for each surrogate exposure model gives us a different predicted indoor concentration for each subject from each model. Finally, we used logistic regression to associate the binary health outcomes with the predicted indoor concentrations, providing an estimated health effect coefficient and SE for each OR and data set.

To assess the simulation results, we quantified the overall bias in the health effects estimates as the mean of the estimated health effect coefficients (across the 4500 simulated data sets) minus the true health effect coefficient (corresponding to ORs of 1.05, 1.50, and 2.00). We estimated the SE for each of the 4500 individual bias estimates to determine whether the estimated bias is significantly different from zero in each case. In addition, we used the estimated median SE (as opposed to the mean, given skewed distributions) to compare the SE of each surrogate exposure model to that of the gold standard. The estimated root mean square error (RMSE) was also calculated as shown in equation (14), to quantify the accuracy of the estimated coefficients:

$$RMSE = \sqrt{\frac{1}{n}\sum_{s=1}^{n}\left(\beta_X - \hat{\beta}_{Xs}\right)^2} \tag{14}$$

where n = the number of simulated data sets; $\hat{\beta}_{Xs}$ = the estimated health effect coefficient for simulation s; and $\hat{\beta}_X$ = the true health effect coefficient (corresponding to an OR of 1.05, 1.50, or 2.00). A 95% confidence interval was then calculated for each $\hat{\beta}_{Xs}$, and the proportion of those simulations with significant positive associations was tabulated. Finally, to examine the robustness of our simulations, the models' coverage probabilities — the proportion of simulations for which the 95% confidence interval contained β_X — were calculated.

We tested the sensitivity of our conclusions by rerunning the analyses with baseline risks of twice and half the original assumption of 11.5% (i.e., for wheeze in the first year of life). In addition, we conducted a sensitivity analysis on the additional residual variability we originally incorporated when generating the indoor pollutant concentrations in the first year of life. This variability, defined with a standard deviation equal to the square root of the residual variance from our LUR models, may have been overestimated in that it was based on the model residuals from the weekly concentrations, which theoretically included both temporal and spatial variability. However, the temporal variability should have been reduced when the validation study models were applied to predict yearly, rather than weekly, concentrations. Thus, we ran the analyses with this variability reduced by half. We also ran the simulations using a true OR of 1.00, to determine whether the proportion of significant positive associations was approximately 2.5%, as would be anticipated under a two-sided test with a significance level of 5%.

As a final step, we derived the approximate relationship between the statistical power of our simulated study using the true exposure (i.e., the gold standard, X) and its power using a surrogate measure of exposure (W) from an alternative

model. Using a one-sided test, the power to detect a significant positive association based on a surrogate model given the power of the true exposure model is

$$Power_W \approx \Phi(\sqrt{R^2}\left[\Phi^{-1}(Power_X)+c\right]-c) \quad (15)$$

where Φ is the normal cumulative distribution function; R^2 is the coefficient of determination from the fitted exposure model; and c is the critical value under the normal distribution function (1.64 for a one-sided test with a significance level of 5%, and 1.96 for 2.5%). The derivation of this equation is in Appendix G (available on the HEI Web site).

RESULTS

PARTICIPANT CHARACTERISTICS

In total, sampling occurred at 44 sites. One site provided only outdoor measurements, 5 sites provided only indoor measurements, and the remaining 38 sites provided both indoor and outdoor measurements (Figure 2). Logistical considerations (e.g., lack of access to electricity) precluded the collection of outdoor measurements at 5 sites; and the 1 site lacking indoor measurements was included to increase outdoor spatial coverage of the study and the availability of duplicate samples. We collected questionnaire data at all 43 sites for which we had indoor measurements. Note that all subsequent discussion about participant characteristics in this report focuses on these 43 homes, which were distributed among 39 households; participants in 4 of the households moved and allowed us to sample in their new home.

Across the sampling sites, a total of 67 sampling sessions were conducted, with 23 homes monitored in both seasons, 16 in the cooling season only, and 5 in the heating season only. Within these 67 sampling sessions, a total of 98 measurements were collected (Figure 2), with 37 sessions including only one 3-to-4-day measurement, 29 sessions including two 3-to-4-day measurements, and 1 session including three 3-to-4-day measurements. Because of our multiday sampling strategy, none of the 98 measurements were weekend only, but 39 were weekday only.

As indicated in Figure 3, our approach for selecting participants successfully yielded both variability in proximity to roadways (measured by kernel-weighted traffic volume within 100 m of the home) and coverage across neighborhoods. Although the study included a disproportionate representation of cohort members in some neighborhoods, such as East Boston (see Figure 1), variability in exposure and spatial coverage was enhanced by the inclusion of a convenience sample of noncohort participants. In general, the traffic exposure was at the lower end of values across the surface generated for the entire Boston area, although the higher values were generally found near major highways where relatively few residences are found (Figure 3).

The distributions of basic home characteristics for all sites, for cohort and noncohort members, are summarized in Table 5. Of the 43 homes with indoor measurements, 25 (58%) were homes of cohort members, and 18 (42%) were not. The age and type of home are similar for both groups, with predominantly older multifamily housing. The noncohort members have slightly larger apartments, but the mean number of rooms (fewer than 5 rooms) remains relatively small.

SUMMARY STATISTICS — OUTDOORS

In total, across 44 sampling locations, we conducted 67 sampling sessions consisting of 98 multiday segments, of which 90 contained at least 1 outdoor measurement (Figure 2). Because of laboratory errors and filter losses during the analysis phase, our maximum sample size for data on the constituents of outdoor $PM_{2.5}$ was 79.

For these outdoor samples, there were 33 pollutants with both hypothesized links to sources of interest and at

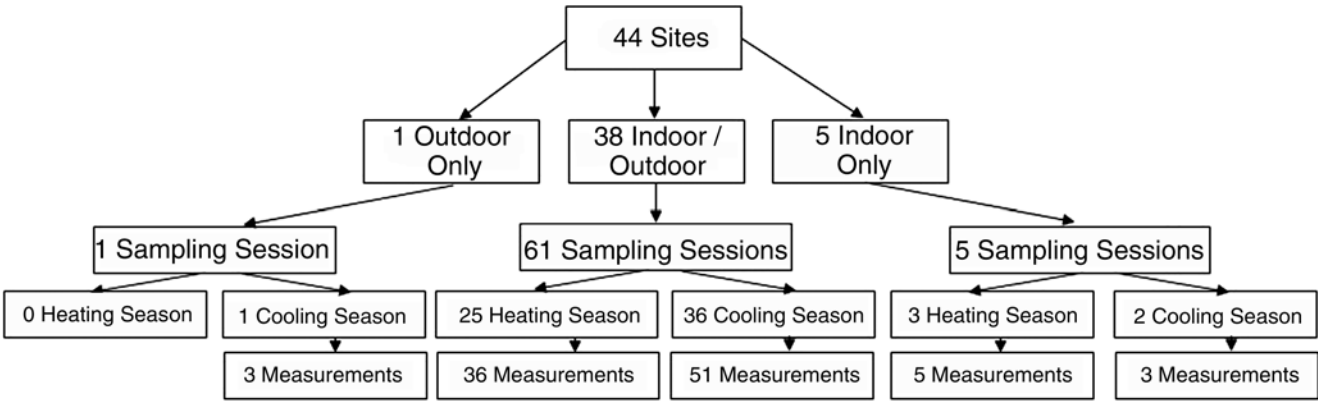

Figure 2. Distribution of samples by site, session, and season.

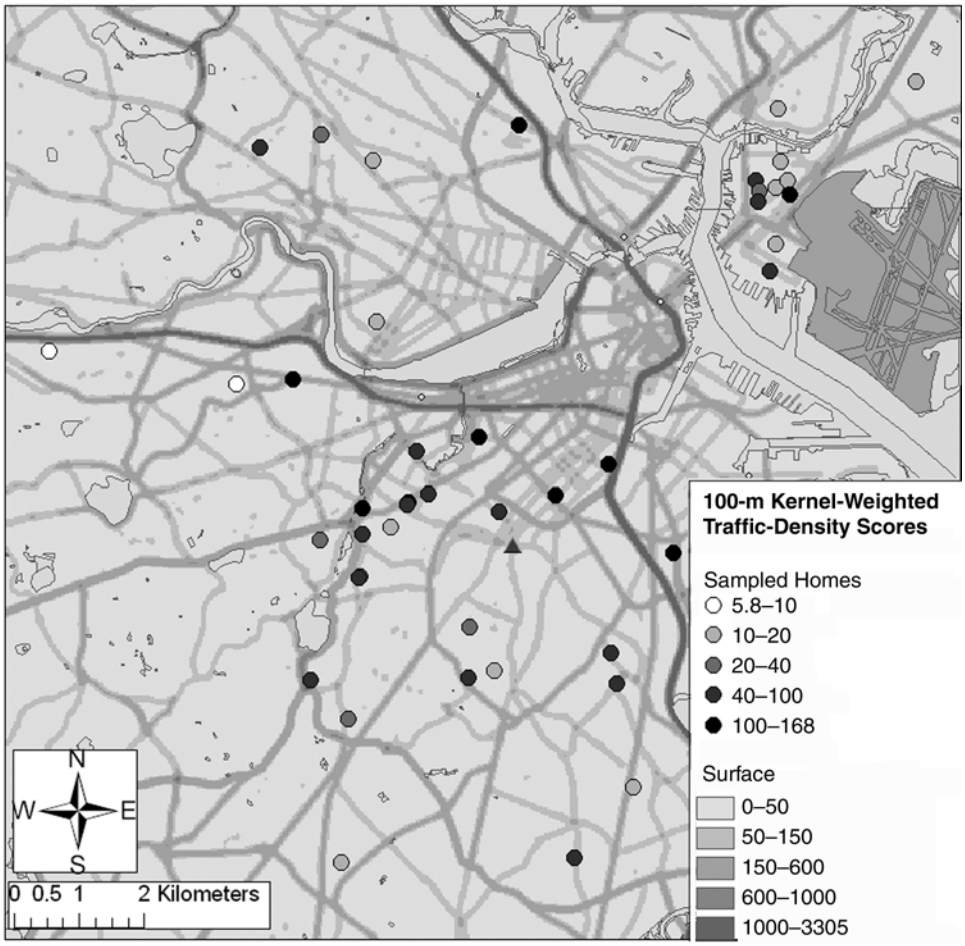

Figure 3. 100-m kernel-weighted traffic scores for sampled homes and the Boston urban area (vehicle-meters/m²/day). Darker shading indicates higher traffic density. The triangle represents the central site monitor in Roxbury, Massachusetts.

Table 5. Distributions of Basic Home Characteristics for All Participants, Cohort Members, and Noncohort Members

	Total (N = 43)	Cohort Members (n = 25)	Noncohort Members (n = 18)
Categorical variables (%)			
Type of housing			
Single-family home	5	4	6
Multifamily home	56	60	50
Apartment building	39	36	44
Year built			
Before 1900	20	16	27
1900–1949	56	52	61
1950–1969	12	16	6
1970–later	12	16	6
Continuous variable (mean [SD])			
Number of rooms	4.2 (1.7)	3.7 (1.2)	4.8 (2.1)

least 60% of observed values above the method LOD for this study (at least three times the analytical SE); these were used in our subsequent factor analyses. These 33 pollutants include NO_2 and EC (measured by reflectance), 23 particle constituents measured by ICP–MS analysis, and 8 elements measured by XRF (all but 2 of which were also measured by ICP–MS). This implies 27 distinct pollutants, though ICP–MS and XRF may not yield identical information, so all 33 are retained. The concentrations of the constituents included in factor analyses, along with total $PM_{2.5}$, are summarized in Table 6. Significant variability was observed across all included constituents.

It should be noted that the sample size differed slightly by constituent. For the ICP–MS analysis, portions of two sample batches were affected by laboratory errors, and a subset of elements from those batches was unreportable.

Table 6. Summary of Outdoor Concentrations, with Percent Above Analytic LOD[a] for Constituent Measures

	n	Mean	Median	SD	Minimum	Maximum	% > LOD
$PM_{2.5}$ (μg/m^3)	83	13.76	12.44	5.27	6.07	31.25	100
EC ($m^{-1} \times 10^{-5}$)	79	0.63	0.58	0.438	0.084	3.81	100
NO_2 (ppb)	61	17.59	15.82	7.34	5.21	39.93	100
Constituents Measured by ICP–MS (μg/m^3)							
Al	69	0.011	5.00×10^{-3}	0.014	-7.47×10^{-4}	0.069	60
As	78	3.17×10^{-4}	2.56×10^{-4}	2.49×10^{-4}	1.46×10^{-5}	1.22×10^{-3}	70
Ba	63	0.024	1.69×10^{-3}	0.049	8.15×10^{-5}	0.196	89
Ca	79	0.067	0.0294	0.162	−0.061	1.32	62
Cd	79	5.64×10^{-5}	3.92×10^{-5}	5.34×10^{-5}	5.34×10^{-5}	2.80×10^{-6}	95
Ce	79	1.31×10^{-5}	7.32×10^{-6}	3.16×10^{-5}	7.0×10^{-8}	2.81×10^{-4}	96
Cr	78	1.56×10^{-4}	1.06×10^{-4}	3.65×10^{-4}	3.88×10^{-6}	3.29×10^{-3}	93
Cu	79	9.97×10^{-4}	8.35×10^{-4}	9.43×10^{-4}	9.09×10^{-6}	7.61×10^{-3}	96
Fe	79	0.011	9.05×10^{-3}	7.56×10^{-3}	4.22×10^{-4}	0.037	99
K	73	0.061	0.039	0.069	3.14×10^{-3}	0.411	84
La	79	1.15×10^{-5}	7.76×10^{-6}	1.73×10^{-5}	4.0×10^{-8}	1.52×10^{-4}	96
Mg	79	8.13×10^{-3}	6.43×10^{-3}	6.42×10^{-3}	1.02×10^{-3}	0.042	92
Mn	79	6.86×10^{-4}	5.43×10^{-4}	8.06×10^{-4}	5.44×10^{-5}	6.60×10^{-3}	99
Mo	79	1.13×10^{-4}	8.98×10^{-5}	1.02×10^{-4}	4.07×10^{-6}	6.77×10^{-4}	88
Ni	79	1.16×10^{-3}	9.81×10^{-4}	8.73×10^{-4}	1.81×10^{-5}	4.66×10^{-3}	96
P	79	3.82×10^{-3}	3.12×10^{-3}	2.64×10^{-3}	-1.76×10^{-4}	0.015	98
Pb	79	7.63×10^{-4}	5.80×10^{-4}	6.15×10^{-4}	1.81×10^{-5}	2.99×10^{-3}	98
S	79	0.994	0.782	0.811	0.047	5.49	99
Sb	79	2.39×10^{-4}	1.98×10^{-4}	1.64×10^{-4}	1.37×10^{-5}	8.17×10^{-4}	100
Se	79	3.73×10^{-4}	3.05×10^{-4}	3.58×10^{-4}	-7.00×10^{-6}	2.12×10^{-3}	73
Sr	67	6.36×10^{-4}	2.23×10^{-4}	1.04×10^{-3}	-1.75×10^{-4}	6.87×10^{-3}	64
Tl	79	6.50×10^{-6}	5.09×10^{-6}	8.30×10^{-6}	-3.0×10^{-7}	6.96×10^{-5}	99
V	79	2.72×10^{-3}	2.10×10^{-3}	2.06×10^{-3}	1.21×10^{-4}	0.011	100
Constituents Measured by XRF (μg/m^3):							
Ca	79	0.029	0.025	0.016	0.010	0.113	99
Fe	79	0.066	0.065	0.024	0.015	0.123	98
K	79	0.055	0.046	0.051	8.48×10^{-3}	0.446	100
P	79	0.065	0.054	0.048	1.50×10^{-3}	0.397	99
S	79	1.46	1.26	1.30	0.063	11.26	100
Si	78	0.051	0.036	0.065	8.89×10^{-3}	0.549	74
V	79	5.09×10^{-3}	3.60×10^{-3}	5.04×10^{-3}	1.08×10^{-3}	0.038	98
Zn	79	0.014	0.011	0.017	4.08×10^{-3}	0.152	98

[a] 3 times the laboratory-based concentration uncertainties.

Among the constituents included in our factor analyses, this problem most significantly influenced samples sizes for Al, Ba, Sr, and K. For the XRF analysis, one implausibly large value was removed for Si, but no other data were unreportable. One batch of NO_2 samples did not meet our data quality criteria, and the results for that batch were not reported.

As a measure of quality assurance for the constituent measurements included in Table 6, we examined the mean relative difference for duplicate samples. For this analysis, we included both indoor and outdoor samples, to increase the sample size. Because of budgetary constraints, the number of duplicate samples was smaller for the ICP–MS analysis than for the XRF analysis, so the mean relative differences for those constituents analyzed by ICP–MS are correspondingly more uncertain. As indicated in Table 7, the mean relative difference varied by constituent but was generally on the order of 10% to 30%, indicating reasonable quality assurance.

Strong correlations were observed across a number of outdoor constituents, suggesting possible common sources (Table 8). For example, EC was most strongly correlated with S, P, and Ca, all of which are characteristic of lubricating oil from diesel vehicles (Spencer et al. 2006). Sulfur was also highly correlated with Se, generally indicative of a coal signature (Table 3). Additional strong correlations included Al and Ba, generally related to road dust and other traffic contributions; Ni and V, a fuel oil signal; and K, Sr, and Cu.

For elements analyzed by both XRF and ICP–MS, correlations between the two measures varied. Ca and Fe were relatively weakly correlated ($r = 0.08$ and 0.49, respectively) (Table 8). For K, S, and V, the correlations between XRF and ICP–MS were higher ($r = 0.73$, 0.92, and 0.71, respectively), possibly because these elements are typically found as water-soluble species.

LUR MODELS FOR OUTDOOR EC, $PM_{2.5}$, AND NO_2

We found, upon examining Anderson-Darling statistics, that $PM_{2.5}$ and EC concentrations were lognormally distributed and were therefore transformed prior to covariate selection ($P = 0.11$ and 0.25, respectively, after log-transformation, compared to 0.005 and 0.04 prior to log-transformation). In contrast, NO_2 values were normally distributed and not transformed ($P > 0.25$ without transformation).

Outdoor $PM_{2.5}$ was highly correlated with central site $PM_{2.5}$ ($R^2 = 0.68$ in log–log regression), as illustrated in Appendix A, indicating a predominance of temporal variability and relative spatial homogeneity in $PM_{2.5}$ across the urban area. In multivariate regressions that included

Table 7. Mean Relative Difference for Duplicate Samples of Total $PM_{2.5}$, NO_2, and Constituents Included in Factor Analyses

	Mean Relative Difference (%)
$PM_{2.5}$	13
EC	24
NO_2	22
Al	55
As	18
Ba	31
Ca	40
Cd	20
Ce	16
Cr	38
Cu	41
Fe	28
K	34
La	20
Mg	28
Mn	16
Mo	34
Ni	30
P	28
Pb	29
S	17
Sb	24
Se	32
Sr	31
Tl	11
V	24
Ca-XRF	26
Fe-XRF	28
K-XRF	12
P-XRF	14
S-XRF	4
Si-XRF	24
V-XRF	27
Zn-XRF	17

Table 8. Pearson Correlations Across All Measures of Outdoor Constituents[a]

	EC	NO_2	Al	As	Ba	Ca	Cd	Ce	Cr	Cu	Fe	K	La	Mg	Mn	Mo	Ni	P	Pb	S	Sb	Se	Sr	Tl	V	Ca-XRF	Fe-XRF	K-XRF	P-XRF	S-XRF	Si-XRF	V-XRF	Zn-XRF
EC	**1.00**																																
NO_2	0.03	**1.00**																															
Al	0.14	0.09	**1.00**																														
As	−0.09	0.27	−0.02	**1.00**																													
Ba	−0.20	0.15	**0.85**	−0.04	**1.00**																												
Ca	**0.73**	0.04	0.43	0.03	0.11	**1.00**																											
Cd	0.23	0.21	0.09	0.37	0.01	0.16	**1.00**																										
Ce	−0.03	−0.22	−0.01	0.15	−0.05	0.08	0.36	**1.00**																									
Cr	−0.06	−0.11	0.00	0.11	0.01	0.02	−0.04	−0.01	**1.00**																								
Cu	0.06	0.19	0.19	0.54	0.12	0.14	0.32	0.01	0.03	**1.00**																							
Fe	0.03	0.38	0.12	0.65	0.07	0.19	0.51	0.35	0.13	0.51	**1.00**																						
K	−0.11	0.15	0.68	0.23	**0.72**	0.11	0.29	0.05	0.00	0.67	0.25	**1.00**																					
La	0.03	−0.22	0.01	0.15	−0.04	0.13	0.37	**0.98**	−0.02	0.05	0.40	0.07	**1.00**																				
Mg	0.16	−0.08	0.04	0.35	−0.10	0.19	0.31	0.07	0.10	0.64	0.36	0.41	0.13	**1.00**																			
Mn	0.33	0.19	0.03	0.26	−0.14	0.21	**0.81**	0.16	0.01	0.18	0.46	0.05	0.21	0.26	**1.00**																		
Mo	0.03	0.12	−0.01	0.33	−0.02	0.04	0.24	0.00	**0.72**	0.31	0.51	0.07	0.04	0.24	0.33	**1.00**																	
Ni	0.20	0.04	0.00	−0.02	−0.03	0.00	0.18	−0.02	0.19	0.10	0.26	0.01	0.07	0.22	0.22	0.44	**1.00**																
P	0.50	0.06	0.13	0.15	−0.09	0.55	0.39	0.18	−0.04	0.26	0.52	0.03	0.27	0.34	0.51	0.35	0.30	**1.00**															
Pb	0.20	0.32	−0.01	0.64	−0.14	0.26	0.67	0.31	−0.04	0.61	**0.76**	0.32	0.33	0.48	0.53	0.35	0.18	0.52	**1.00**														
S	0.59	0.16	0.21	0.40	−0.08	0.65	0.49	0.28	−0.01	0.34	0.66	0.09	0.33	0.37	0.51	0.32	0.15	0.68	0.68	**1.00**													
Sb	0.10	0.29	0.16	**0.72**	0.10	0.21	0.47	0.09	0.00	**0.78**	**0.77**	0.46	0.15	0.49	0.34	0.40	0.18	0.43	**0.76**	0.59	**1.00**												
Se	0.41	0.23	0.26	0.48	0.08	0.52	0.51	0.19	0.05	0.35	0.58	0.17	0.21	0.18	0.46	0.41	0.10	0.53	0.65	**0.84**	0.63	**1.00**											
Sr	0.04	0.11	0.56	0.30	0.51	0.28	0.25	−0.02	0.02	**0.82**	0.26	**0.92**	−0.01	0.56	0.04	0.11	0.00	0.12	0.44	0.22	0.58	0.27	**1.00**										
Tl	0.35	0.09	0.06	0.27	−0.14	0.20	**0.81**	0.21	−0.04	0.13	0.30	0.08	0.22	0.17	**0.91**	0.17	0.04	0.38	0.50	0.48	0.28	0.46	0.07	**1.00**									
V	0.22	0.12	0.00	−0.05	0.00	−0.05	0.09	−0.05	−0.06	0.13	0.21	0.04	0.07	0.15	0.17	0.28	0.66	0.32	0.13	0.11	0.17	0.06	−0.02	0.00	**1.00**								
Ca-XRF	0.36	0.04	0.05	−0.02	−0.08	0.08	0.49	0.03	0.02	0.03	0.19	0.02	0.11	0.21	0.69	0.23	0.33	0.38	0.21	0.29	0.09	0.15	−0.04	0.59	0.45	**1.00**							
Fe-XRF	0.17	0.17	−0.04	0.34	−0.07	0.02	0.36	0.07	0.07	0.30	0.49	0.07	0.12	0.16	0.46	0.35	0.41	0.28	0.34	0.30	0.41	0.29	0.05	0.34	0.27	0.50	**1.00**						
K-XRF	0.01	0.11	0.17	0.35	0.10	0.04	0.39	0.08	0.01	**0.86**	0.28	**0.73**	0.08	0.63	0.20	0.14	0.02	0.12	0.56	0.19	0.55	0.17	**0.83**	0.23	0.04	0.11	0.11	**1.00**					
P-XRF	**0.71**	0.09	0.28	0.28	−0.04	**0.76**	0.41	0.20	−0.03	0.27	0.45	0.10	0.25	0.31	0.44	0.17	0.01	0.62	0.53	**0.94**	0.45	**0.75**	0.24	0.46	0.00	0.28	0.22	0.18	**1.00**				
S-XRF	**0.76**	0.05	0.31	0.22	−0.03	**0.81**	0.34	0.17	0.00	0.24	0.39	0.08	0.21	0.29	0.37	0.16	0.01	0.59	0.47	**0.92**	0.40	**0.74**	0.24	0.40	−0.01	0.24	0.17	0.14	**0.99**	**1.00**			
Si-XRF	0.03	−0.09	0.18	−0.01	0.13	0.08	0.11	0.00	0.07	−0.07	−0.01	0.06	−0.02	−0.02	0.15	0.00	0.42	−0.03	−0.01	0.00	−0.06	0.00	0.05	0.15	−0.18	0.10	0.39	−0.06	−0.04	−0.04	**1.00**		
V-XRF	0.30	0.05	0.00	−0.23	−0.04	0.06	−0.06	−0.09	−0.08	−0.05	−0.12	−0.04	−0.02	0.01	0.01	0.03	0.36	0.10	−0.07	0.02	−0.07	−0.03	−0.05	−0.06	**0.71**	0.44	0.09	0.01	0.07	0.08	−0.12	**1.00**	
Zn-XRF	0.12	0.03	−0.10	−0.14	−0.08	0.02	0.02	−0.06	−0.08	−0.07	−0.12	−0.05	−0.06	−0.07	0.01	−0.04	0.01	0.03	−0.01	−0.01	−0.06	0.01	−0.04	0.01	0.03	0.07	−0.02	0.07	0.03	0.03	−0.12	0.09	**1.00**

[a] Correlation coefficients greater than 0.7 are presented in bold.

central site data, the best traffic indicator was total roadway length within 100 m of the home, which retained statistical significance in multivariate models (Table 9). Final multivariate model results indicate that the traffic–$PM_{2.5}$ relationship was not significantly altered by any of our candidate modifiers. Other combustion sources (smoking or grilling) and population density significantly contributed to concentrations (overall $R^2 = 0.76$).

EC showed relatively poor associations with central site data overall (Appendix A), though this was partly attributable to seasonal differences in the relationship, with varying slopes and stronger correlations during the cooling season (Spearman $r = 0.66$) than the heating season ($r = 0.37$). In the final multivariate model ($R^2 = 0.52$), EC was best predicted by total roadway length within 200 m of the home, and the association between EC and traffic was increased under low-wind-speed conditions (Table 9). During the cooling season, residential EC concentrations were somewhat lower and displayed stronger associations with central site data. Approximately 30% of the variability in EC was explained by temporal terms, and 14% by the traffic term (spatial component). The interaction of traffic with hours of low wind speed, which incorporated both spatial and temporal variance, accounted for an additional 8%.

NO_2 was weakly associated with central site concentrations ($R^2 = 0.21$) (Table 9), suggesting significant spatial heterogeneity within urban residential areas (Appendix A). The final multivariate model ($R^2 = 0.56$) included total roadway length within 50 m of the home, significantly attenuated by an obstruction (e.g., a building) between the monitor and nearest major road (Table 9). Season had a significant effect, beyond the effects captured by central site monitoring data, and concentrations were positively associated with population density (Table 9). Spatial terms (traffic, obstruction between the monitor and nearest major road, and population density) together accounted for approximately 23% of NO_2 variability. Temporal terms (central site concentration, season) accounted for about 33% of NO_2 variability.

Table 9. Multivariate LUR Model Results for Outdoor $PM_{2.5}$, EC, and NO_2

	ln($PM_{2.5}$) (μg/m^3)			ln(EC) ($m^{-1} \times 10^{-5}$)			NO_2 (ppb)		
Predictor Type	Model	β (*P* Value)	Sequential R^2	Model	β (*P* Value)	Sequential R^2	Model	β (*P* Value)	Sequential R^2
Intercept		0.205 (0.32)	—		−0.907 (< 0.0001)	—		−12.50 (0.009)	—
Central site concentration	ln(Central Site [$PM_{2.5}$])	0.776 (< 0.0001)	0.68	ln(Central site [EC])	0.103 (0.59)	0.03	Central site (NO_2)	1.06 (< 0.0001)	0.21
				ln(Central site [EC]) × cooling season	0.82 (0.004)	0.26			
Traffic indicator	Roadway length (100 m)	1.48×10^{-4} (0.02)	0.70	Roadway length (200 m)	1.10×10^{-4} (0.01)	0.40	Roadway length (50 m)	0.0144 (0.002)	0.22
Traffic indicator × modifier	N/A	N/A	N/A	Roadway length (200 m) × % hours with still winds	4.38×10^{-4} (0.02)	0.48	Roadway length (50 m) × monitor obstructed from nearest major road	−0.0094 (0.005)	0.31
Other sources / land use	Smoking or grilling	0.156 (0.01)	0.73	Cooling season	−0.268 (0.057)	0.52	Cooling season	4.93 (0.001)	0.44
	Population density	9.24×10^{-6} (0.01)	0.76				Population density	4.01×10^{-4} (0.001)	0.56

A one-at-a-time exclusion cross validation was performed to assess the internal consistency of model results. The Spearman correlation between estimated and measured log $PM_{2.5}$ was 0.80, 0.66 for log EC, and 0.66 for NO_2 ($P < 0.0001$ in all cases), indicating acceptable internal validity.

Because we had many candidate covariates and a relatively small sample size, we conducted a range of sensitivity analyses in which we considered the following factors: alternative traffic indicators, central site monitoring data, site-specific monitors, and other model decisions. The results of these sensitivity analyses are described below.

Selection of Traffic Indicators

Sensitivity analyses (Appendix B) indicated that other traffic indicators could not be substituted to create a comparable model for $PM_{2.5}$. For EC, diesel-based measures derived from our traffic counter could explain slightly more variability, with R^2 values of approximately 0.54, but were available for only a subset of locations ($n = 34$) and were therefore not considered for the primary model. For the full sample, no indicator was exchangeable with roadway length within 200 m of the home. In addition, the interaction term of traffic modified by low wind speeds remained significant in several cases in which the main effect of traffic did not maintain significance. For NO_2, sensitivity tests supported the finding that shorter buffer lengths were most effective at predicting variability in NO_2. The use of longer buffer lengths did not result in a comparable model, but kernel-weighted traffic density within 50 m of the home could be substituted effectively for total roadway length within 50 m of the home, as could unweighted cumulative traffic density within 100 m of the home. Similar to EC, a diesel-based term provided slightly greater explanatory power for NO_2, but it was not available for the full sample.

Accuracy of Traffic Data

To validate raster-based traffic indicators, we considered base cell sizes from 10-to-50-m square. We found that base cells in this size range, as compared to the default 25-m square cell size, had little influence on traffic indicator values. Because of concerns about data quality, where possible we verified MHD traffic counts against traffic data obtained from the Massachusetts Executive Office of Transportation, the ESRI Business Analyst (ESRI; Redlands, CA), and the traffic counts we collected on the largest road within 100 m of each home (using the JAMAR TRAX I device). Correlations across traffic metrics derived from the different sources were generally above 0.7.

Selection of the Central Site Monitor

We considered several alternatives to the use of the Roxbury central site monitor concentrations for temporal correction, including using other available urban monitors individually, the average concentration from all urban monitors available for each sampling period, and the mean concentration at a background monitor 10 miles south of the Roxbury central site monitor (available for summer months only). Given the size of our domain and the small number of central site monitors with hourly concentration data located within the domain, a kriged surface was not feasible. No alternative to the Roxbury central site sampling period mean explained greater variability in concentrations or significantly altered traffic–pollution relationships in multivariate models. As indicated in Appendix A, correlations among urban central site monitors were generally high, with the exception of monitors placed in nonrepresentative locations.

Selection of Meteorological and Site-Specific Modifiers for EC and NO_2

All EC models showed a significant, positive effect of low wind speeds on the traffic–pollution relationship. Sensitivity analyses indicated that other wind variables (mean daytime wind speed, percent of day downwind from road) were significant and could be substituted for percent of low wind speed hours, losing only marginal explanatory power ($R^2 = 0.52$ and 0.49, respectively). These similar findings for wind speed and direction could be expected, as wind speed and direction at our central site were highly correlated, with higher wind speeds from the west (data not shown).

For NO_2, no modifier could replace the term of obstruction between the monitor and nearest major road in the final model. Because the presence of an obstruction could theoretically proxy for distance to the nearest major road, we replaced the term with distance to the nearest major road, but we found highly nonsignificant results, indicating that obstruction does not likely serve as a proxy for distance in this data set.

Exploration of Log-Transformation of $PM_{2.5}$ and EC Data

The selection of the 100-m roadway length term and other predictors for the $PM_{2.5}$ model was not dependent on log-transformation. Using untransformed $PM_{2.5}$ data, we achieved an R^2 value of 0.73, and all predictors retained significance. For untransformed EC data, the traffic term of total roadway length within 200 m of the home and all other predictors retained significance, with an R^2 value of 0.51.

Inclusion of a Categorical Variable for Season

Because the season term might have been extraneous in models including central site monitoring data, thereby producing models that might have been difficult to interpret, we explored the effect of removing season from the final EC and NO_2 models. For EC, removing the season term caused the central site monitor coefficient to drop by half and to no longer be significant, while the effect of low wind speed increased by almost 50%, and overall model fit declined. Because overall explanatory power was reduced when the season term was removed, and because of known seasonal differences in the relationship between our reflectance measures and central site aethalometer data, we opted to retain both the season covariate and season-specific coefficients in the final model.

For NO_2, dropping the season term decreased the effect of the central site monitor by approximately 50%, but did not affect overall model fit or other parameters. Thus, although NO_2 was higher at the study residences during the cooling season, the effect was captured in part by the central site monitor; because the term did not significantly alter other parameters, we opted to leave it in the final model. Finally, we tested the addition of a season term to the $PM_{2.5}$ model and found no effect on the central site monitor term or overall fit, although the influence of other combustion sources (i.e., smoking or grilling) was increased by approximately 35%, indicating possible seasonal differences in these source activities. Because traffic was the source of primary interest, however, and because the season term did not improve overall fit or alter the observed influence of traffic, we opted to maintain the original, more parsimonious $PM_{2.5}$ model.

Robustness to Within-Site Autocorrelation

Because many homes were monitored during two seasons, we examined the effect of within-site autocorrelation using random effects by household. Autocorrelation by site did not influence model fit or parameter estimates for any model for all pollutants.

Robustness to Outlier Inclusion or Exclusion

For $PM_{2.5}$, our examination of scatterplots and distributions led us to remove three observations that reflected likely measurement error or nonrepresentative data. In one of these cases, the pump was unplugged within 30 minutes of monitor deployment. In the other two cases, ADT was quite high given close proximity to a major highway. These observations were also removed for EC, along with one very high EC outlier and one observation with extremely high wind speeds (given the potential influence of wind speed on local source contributions). The two very-high-traffic sites and the high-wind-speed observation were also removed for NO_2. In all cases, inclusion of these outliers and influential points did not have a material influence on our findings, with all covariates retaining significance and directionality.

CONSTRAINED FACTOR ANALYSIS FOR OUTDOOR CONCENTRATIONS

Comparing across models with varying numbers of factors (F), the BIC was minimized with $F = 2$ (with a local minimum at $F = 5$), while $F = 8$ was optimal by the K–G criterion. The cross-validation cost function was minimized at $F = 3$. We selected a five-factor model as a reasonable compromise among these various approaches, offering reasonably good physical interpretability of the factors. We evaluated interpretability by examining the consistency between factor interpretations and the hypothesized sources in Table 3, and we examined Pearson correlations among factors, measured concentrations, and source indicators. The five-factor model yielded the loading matrix shown in Figure 4, with darker shading representing higher factor loadings. Constituents with the "XRF" suffix are from XRF analyses; others are from ICP–MS.

As depicted in Figure 4, more constituents loaded strongly onto Factor 1 than onto any other factors, including S and Se, which are associated elsewhere with power plant emissions (Table 3). Other constituents with high loadings on Factor 1 are not directly associated with power plant emissions, but reflect industrial sources not found within urban Boston (e.g., iron and lead as potential markers of steel manufacturing) as well as traffic. These observations would seem to indicate a factor that corresponds with higher total $PM_{2.5}$, potentially driven by meteorological conditions or long-range transport. Given the large number of constituents, including constituents with few known sources close to our monitoring region, we interpreted Factor 1 as *long-range transport.*

Factor 2 was dominated by Cu, Sr, K, and Sb, with small but positive NO_2 loading. As indicated in Table 3, many of these constituents are associated with vehicular brake or tire wear, while NO_2 may indicate local combustion sources, including motor vehicles. Thus, we hypothesized that this factor was *brake wear and local traffic*.

Factor 3 shows heaviest loadings for EC, Ca, S, and P. As mentioned earlier, this combination is commonly associated with lubricating oil from diesel exhaust, while EC itself has been more broadly associated with diesel combustion. We therefore interpreted this factor as *diesel exhaust.*

Ni and V, common *fuel oil combustion* markers, load heavily on Factor 4. Finally, Al, Ca, Cd, Mn, and Tl load

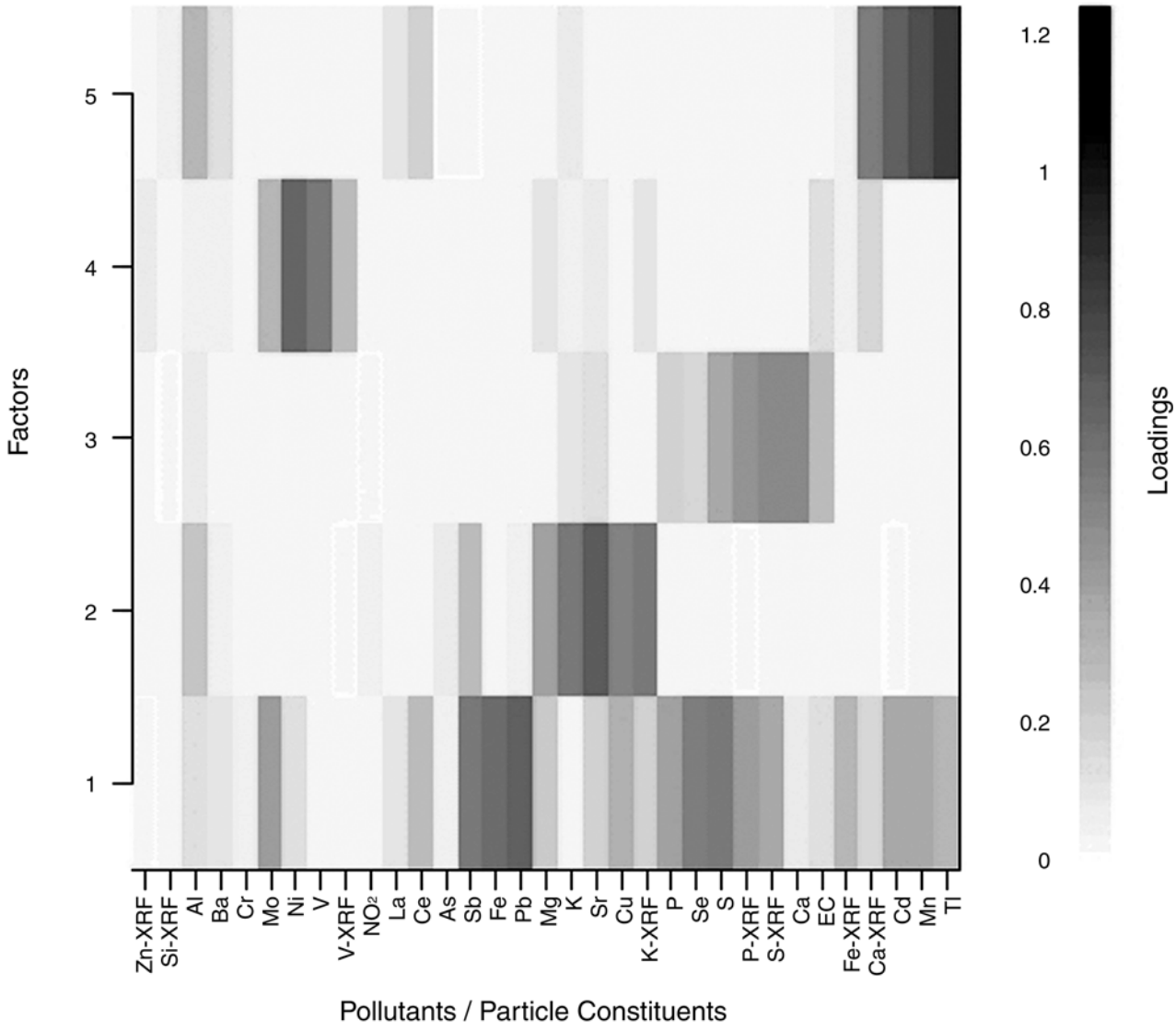

Figure 4. Factor loading matrix for outdoor concentrations using a five-factor model. Pollutants and particle constituents are sorted by their position in a hierarchical clustering dendrogram, to optimize visual interpretation by placing correlated constituents together.

strongly on Factor 5. Many of these constituents have been associated with *soil and road dust resuspension* (Table 3).

Mean factor scores by location are presented in Figure 5. These maps emphasize that spatial patterns differ somewhat across the factors, although the values in these figures are not adjusted a priori for meteorology and other temporally varying terms.

To evaluate the robustness of our factor interpretations, in light of the complexities associated with samples that vary over both space and time, we conducted an identical constrained factor analysis on a separate set of $PM_{2.5}$ filters collected over time at one location within a single season (see Additional Air Pollution Sampling, under Methods and Study Design). Results from this analysis generally corroborated the structure and interpretation of our primary factor analysis, while pointing out uncertainties associated with the interpretation of our brake wear and road dust terms (Appendix C).

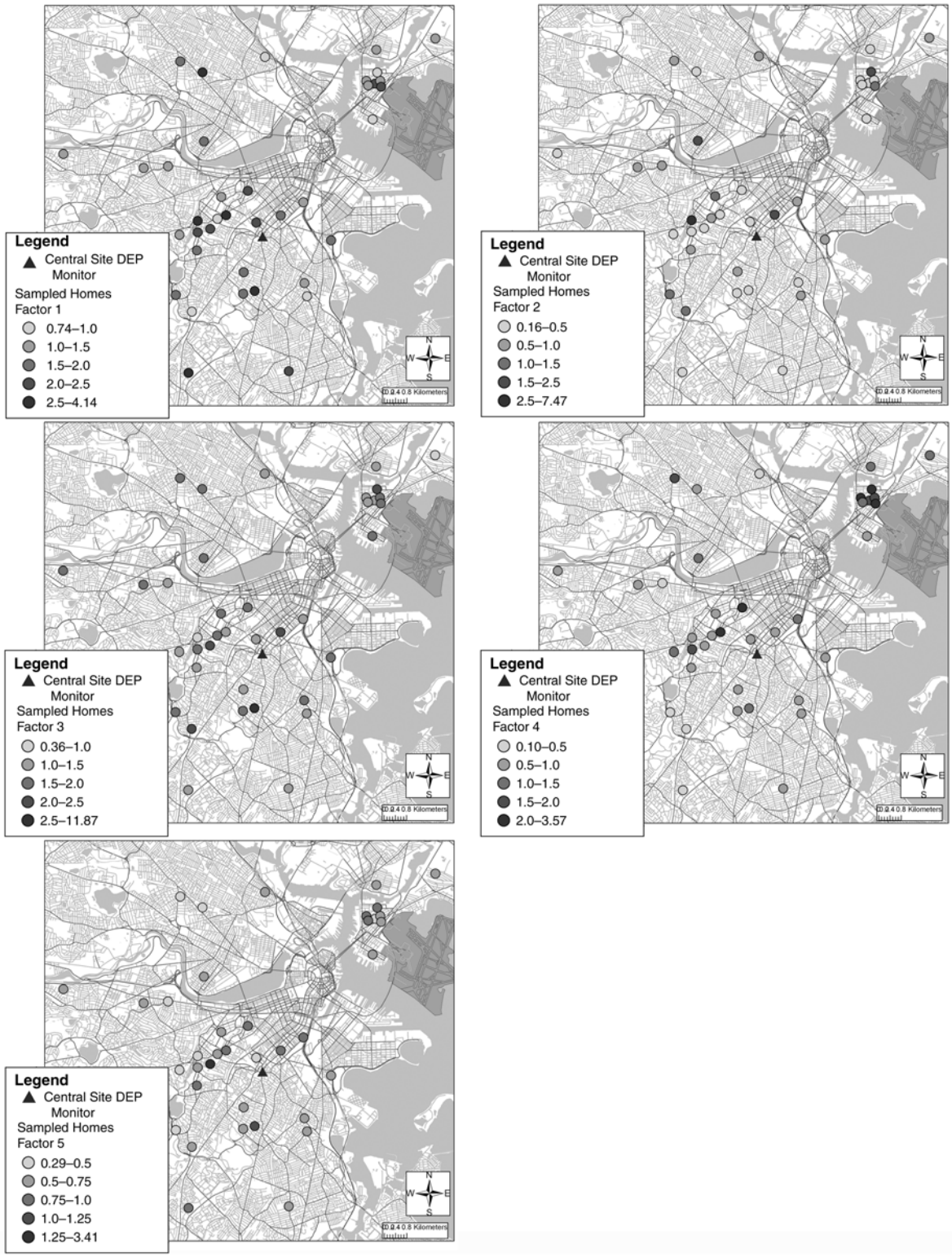

Figure 5. Distribution across sampled sites of site-averaged outdoor-concentration factor scores.

LUR MODELS FOR OUTDOOR LATENT VARIABLES

Univariate associations between raw factor loadings for outdoor concentrations and candidate source terms in four categories are presented in Table 10. Factors 1 and 4 (interpreted as long-range transport and fuel oil combustion) demonstrated the most significant associations with central site monitors and season of measurement, with Factor 1 showing higher values in the cooling season and Factor 4 showing higher values in the heating season. Factor 2 was significantly, but negatively, associated with central site NO_2; given that we could not interpret this term, we did not consider it going forward.

Different candidate traffic terms were optimal across the five factors, with generally modest associations in univariate comparisons. Multiple additional local sources or population terms had univariate P values of less than 0.25 and were considered as candidate variables in our multivariate models, with the recognition that some significant univariate terms may serve as proxies for other sources or represent spurious correlations if they are not adjusted for other terms.

Results of our forward stepwise modeling process are summarized in Table 11, with sequential R^2 values from each step indicating the marginal explanatory power of each new term included. The multivariate LUR models largely supported our hypothesized factor interpretations. Factor 1 (long-range transport) was strongly predicted by central site $PM_{2.5}$ and season, producing the best fit of any of our models ($R^2 = 0.69$) and also the strongest relationship with central site monitoring data. Other candidate terms did not retain significance or improve model fit, suggesting little local-source impact.

In contrast, Factor 2 (brake wear and local traffic) retained no central site monitoring terms in the multivariate model. Local roadway length weighted by wind direction was moderately predictive and was more influential during the cooling season. In addition, an inverse relationship was observed with high-density residential land use.

Factor 3 (diesel exhaust) was positively associated with central site $PM_{2.5}$ and was higher during the heating season, and additional variability was explained by the percentage of diesel traffic on the nearest major road, an effect modified by monitor height (the building floor from which the outdoor sampler was mounted).

Factor 4 (fuel oil combustion) was predicted by central site NO_2 (a strong marker for heating season and temperature) and population density within 200 m of the home, with population density exerting a greater effect under low-wind conditions, suggesting influential local sources. Residential fuel oil use was not significant in the final model, but was highly correlated with population density; therefore, we would not expect to retain both in multivariate models.

Factor 5 (road dust and resuspension) was positively associated with the heating season and with high-density residential land use, an effect modified by an obstruction (e.g., a building) between the monitor and nearest road.

Sensitivity Analyses

The robustness of final covariate selection for the outdoor factor analysis LUR models in Table 11 was examined by replacing traffic and source covariates in each model with candidate terms from each category in Table 10. In no case did other indicators improve significance or overall model fit.

For all factors, the random forest algorithm indicated that the covariates previously selected to predict each factor were informative. Most deviations in our models, relative to factor rankings from the random forest algorithm, were due to our model-building approach, which was deliberately constrained to preserve power and interpretability. For example, for the first factor (long-range transport), central site $PM_{2.5}$ and EC were the two most informative predictors indicated by the random forest algorithm, though our model-building criteria led us to include only one best central site term ($PM_{2.5}$) in our model. The random forest algorithm is insensitive to multicollinearity, so both $PM_{2.5}$ and EC (which were significantly positively correlated with one another) were selected by the random forest but would not both be included in an interpretable regression model. Similarly, for all other factors, all significant model terms were among the 10 most important terms indicated by the random forest algorithm. Note that while the random forest algorithm selects the "optimal" predictors, providing a robust indication of variable importance, our simpler LUR approach yields more interpretable models.

Prior to constructing our primary regression models (Table 11), we removed a small number of influential points (one observation was removed from the models for Factors 2, 3, and 5; no observations were removed for Factors 1 and 4). All model coefficients were statistically significant with inclusion of these influential points. In addition, after reviewing plots of factor scores against selected covariates, we tested the effect of removing additional outliers and potential influential points. For Factors 1, 3, and 5, no clear outliers were observed. Factor 4 had one outlier, but removing it did not affect model fit. Factor 2 (brake wear and local traffic) contained a cluster of high concentrations in one neighborhood; because these scores may have been driven by a unique local source not captured in our data, we tested the effect of their removal. The traffic term weighted by wind direction retained significance, though its interaction by season did not, likely because the six high observations that were removed were all cooling season values, thus reducing between-season variability.

Table 10. Candidate Variables for Outdoor Factor Analysis LUR Models, by Univariate Association with Raw Factor Scores, Adjusted for Repeated Measures by Household[a]

Factor: Interpretation / Covariate Type	Variable(s)	Univariate Coefficient	Univariate (P Value)	Univariate r^2
Factor 1: Long Range Transport				
Central site	$PM_{2.5}$ (μg/m^3)	0.130	(< 0.0001)	0.57
	NO_2 (ppb)	29.1	(0.21)	0.02
	EC (μg/m^3)	1.53	(< 0.0001)	0.33
Season	Cooling season*	0.951	(< 0.0001)	0.25
Traffic terms	Distance to designated truck route (m)	1.15×10^{-4}	(0.37)	0.01
Local sources / population	Smoking/ grilling near outdoor monitor*	0.440	(0.09)	0.04
Factor 2: Brake Wear and Local Traffic				
Central site	NO_2 (ppb)	−27.5	(0.06)	0.25
Season	Cooling season*	0.213	(0.11)	0.03
Traffic terms	Roadway length within 100 m weighted by wind direction (m)	3.06×10^{-3}	(0.24)	0.02
	Roadway length within 200 m (m)	-6.4×10^{-5}	(0.25)	0.02
Local sources / population	Fuel oil use within 200 m (per 10,000 users)	1.66×10^{-8}	(0.05)	0.05
	Construction within 100 m*	0.238	(0.18)	0.03
	Population density within 200 m (pers/km^2)	-1.91×10^{-5}	(0.10)	0.04
	High density residential land use*	−0.241	(0.07)	0.04
Factor 3: Diesel Exhaust				
Central site	$PM_{2.5}$ (μg/m^3)	0.024	(0.01)	0.08
Season	Cooling season*	−0.103	(0.35)	0.01
Traffic terms	Diesel fraction on nearest major road (%)	2.88	(0.14)	0.05
	Distance to nearest larger road (m)	-6.60×10^{-4}	(0.25)	0.02
Local sources / population	Construction within 100 m*	0.116	(0.44)	0.009
Factor 4: Fuel Oil Combustion				
Central site	NO_2 (ppb)	68.5	(0.0009)	0.14
	EC (μg/m^3)	0.36	(0.19)	0.02
Season	Cooling season*	−0.509	(0.009)	0.09
Traffic terms	Roadway length within 100 m (m)	4.09×10^{-4}	(0.11)	0.03
	Roadway length within 200 m (m)	2.34×10^{-4}	(0.004)	0.11
	Roadway length within 300 m (m)	1.25×10^{-4}	(0.0002)	0.16
	Distance to nearest truck route (m)	-2.03×10^{-4}	(0.14)	0.03
Local sources / population	Fuel oil use within 200 m (10,000 users)	3.4×10^{-8}	(0.006)	0.09
	Construction within 100 m*	−0.433	(0.10)	0.04
	Block group population density (pers/km^2)	5.06×10^{-5}	(0.001)	0.13
	Population density within 200 m (pers/km^2)	6.95×10^{-5}	(< 0.0001)	0.22
	High density residential land use*	0.496	(0.0096)	0.08
Factor 5: Road Dust and Resuspension				
Central site	$PM_{2.5}$ (μg/m^3)	−0.0070	(0.16)	0.03
	NO_2 (ppb)	11.3	(0.05)	0.05
	EC (μg/m^3)	−0.127	(0.09)	0.04
Season	Cooling season*	−0.133	(0.01)	0.08
Traffic terms	Roadway length within 100 m (m)	9.09×10^{-5}	(0.19)	0.02
	Distance to highway (m)	-1.05×10^{-4}	(0.23)	0.02
Local sources / population	High density residential land use*	0.138	(0.009)	0.09
	Construction within 100 m*	0.079	(0.25)	0.02

[a] Candidate covariates have univariate $P < 0.25$; for source categories with no significant covariate, we considered the best candidate (by P value, univariate r^2). Sources marked by an asterisk (*) are binary indicators.

Table 11. Multivariate LUR Model Results for Outdoor Concentration Factor Scores, Adjusted for Repeated Measures by Household[a]

Factor: Interpretation / Covariate Type	Variable(s)	Univariate Coefficient	Univariate (P Value)	Seq R^2 [b]
Factor 1: Long-Range Transport				
Central site	$PM_{2.5}$ (μg/m^3)	0.129	(< 0.0001)	0.63
Season	Cooling season*	0.527	(< 0.0001)	**0.69**
Traffic terms	—			
Local sources / population	—			
Modified traffic and local sources	—			
Factor 2: Brake Wear and Local Traffic				
Central site	—			
Season	—			
Traffic terms	Roadway length within 100 m weighted by wind direction (m)	0.00369	(0.16)	0.02
Local sources / population	High density residential land use*	−0.362	(0.006)	0.11
Modified traffic and local sources	Roadway length within 100 m weighted by wind direction × cooling season (m)	0.00329	(0.04)	**0.16**
Factor 3: Diesel Exhaust				
Central site	$PM_{2.5}$ (μg/m^3)	0.050	(0.001)	0.08
Season	Cooling season*	−0.268	(0.06)	0.13
Traffic terms	Diesel fraction on nearest major road (%)	6.07	(0.03)	0.26
Local sources / population	—			
Modified traffic and local sources	Diesel fraction on nearest major road × floor of building (%)	−1.72	(0.06)	**0.32**
Factor 4: Fuel Oil Combustion				
Central site	NO_2 (ppb)	59.7	(0.002)	0.14
Season	—			
Traffic terms	—			
Local sources / population	Population density within 200 m (pers/km^2)	5.3×10^{-5}	(0.0003)	0.31
Modified traffic and local sources	Population density within 200 m × percent hours of wind < 2 m/sec (pers/km^2)	2.1×10^{-4}	(0.02)	**0.41**
Factor 5: Road Dust and Resuspension				
Central site	—			
Season	Cooling season*	−0.144	(0.008)	0.08
Traffic terms	—			
Local sources / population	High density residential land use*	0.240	(0.002)	0.16
Modified traffic and local sources	High density residential land use × monitor obstructed from nearest road*	−0.145	(0.08)	**0.20**

[a] Sources marked by an asterisk (*) are binary indicators.

[b] Values in bold are final multivariate R^2 values for each model.

Finally, to test for correlation across variables and implications for covariate significance, we performed backwards elimination on final models to ensure that covariates retained significance independent of other predictors in the model. In all cases, terms retained significance and the original models were retained.

SUMMARY STATISTICS — INDOORS

Across the 43 sampling locations where indoor measurements were collected, 92 of the multiday sampling events contained at least one indoor particulate matter sample. Because of laboratory errors and other filter losses during the analysis phase, our maximum sample size for indoor $PM_{2.5}$ constituent data was 88. Summary statistics for indoor concentrations of the 33 constituents included in our factor analyses, along with total $PM_{2.5}$, are presented in Table 12.

Table 12. Summary of Indoor Concentrations, with Percentage Above Analytic LOD[a] for Constituent Measures

	N	Mean	SD	Median	Minimum	Maximum	% > LOD
$PM_{2.5}$ (μg/m^3)	92	20.1	12.0	16.8	3.6	74.9	100
EC ($m^{-1} \times 10^{-5}$)	88	0.58	0.46	0.50	0.0029	3.2	100
NO_2 (ppb)	74	19.1	10.7	16.6	4.9	61.1	100
Constituents measured by ICP–MS (μg/m^3):							
Al	78	0.011	0.027	4.2×10^{-3}	-3.1×10^{-3}	0.20	68
As	84	4.6×10^{-4}	1.3×10^{-3}	1.8×10^{-4}	2.7×10^{-6}	0.010	65
Ba	77	0.019	0.051	1.1×10^{-3}	-3.8×10^{-6}	0.33	86
Ca	84	0.060	0.081	0.033	−0.026	0.56	67
Cd	85	9.3×10^{-5}	2.1×10^{-4}	4.8×10^{-5}	-1.1×10^{-6}	1.5×10^{-3}	98
Ce	85	1.4×10^{-4}	4.0×10^{-4}	1.3×10^{-5}	-3.2×10^{-7}	2.2×10^{-3}	97
Cr	85	7.9×10^{-4}	4.1×10^{-3}	1.1×10^{-4}	-1.4×10^{-6}	0.037	98
Cu	85	8.4×10^{-4}	9.5×10^{-4}	5.9×10^{-4}	5.8×10^{-7}	7.0×10^{-3}	98
Fe	85	7.9×10^{-3}	7.4×10^{-3}	6.3×10^{-3}	-4.8×10^{-5}	0.044	97
K	81	0.076	0.081	0.041	-1.7×10^{-4}	0.38	91
La	85	9.2×10^{-5}	2.5×10^{-4}	1.1×10^{-5}	-1.0×10^{-7}	1.3×10^{-3}	96
Mg	85	0.011	0.013	7.5×10^{-3}	-2.4×10^{-4}	0.11	97
Mn	85	5.7×10^{-4}	4.1×10^{-4}	4.9×10^{-4}	-3.3×10^{-6}	2.3×10^{-3}	98
Mo	85	6.4×10^{-5}	5.2×10^{-5}	6.0×10^{-5}	-4.0×10^{-5}	2.9×10^{-4}	68
Ni	85	9.5×10^{-4}	8.8×10^{-4}	7.0×10^{-4}	-1.7×10^{-5}	6.6×10^{-3}	96
P	85	0.023	0.086	4.0×10^{-3}	-1.8×10^{-4}	0.68	99
Pb	85	1.5×10^{-3}	6.1×10^{-3}	5.9×10^{-4}	-2.3×10^{-6}	0.056	99
S	85	0.80	0.59	0.63	-6.3×10^{-4}	4.0	100
Sb	85	3.3×10^{-4}	1.5×10^{-3}	1.3×10^{-4}	9.3×10^{-7}	0.014	99
Se	85	2.4×10^{-4}	3.7×10^{-4}	1.2×10^{-4}	-2.0×10^{-4}	1.9×10^{-3}	62
Sr	77	5.4×10^{-4}	1.1×10^{-3}	2.4×10^{-4}	-5.3×10^{-4}	7.6×10^{-3}	69
Tl	85	8.0×10^{-6}	1.3×10^{-5}	5.1×10^{-6}	-3.7×10^{-8}	8.5×10^{-5}	100
V	85	2.0×10^{-3}	1.7×10^{-3}	1.5×10^{-3}	-3.1×10^{-9}	0.010	100
Constituents measured by XRF (μg/m^3):							
Ca	88	0.046	0.074	0.033	1.8×10^{-3}	0.68	99
Fe	88	0.047	0.022	0.045	3.5×10^{-3}	0.14	92
K	88	0.078	0.079	0.053	4.8×10^{-5}	0.48	100
P	88	0.070	0.093	0.044	1.7×10^{-3}	0.75	98
S	88	1.13	0.76	0.89	6.0×10^{-4}	5.0	99
Si	88	0.090	0.28	0.043	4.0×10^{-3}	2.6	73
V	88	3.7×10^{-3}	3.5×10^{-3}	2.8×10^{-3}	3.0×10^{-4}	0.027	91
Zn	88	0.014	0.016	9.3×10^{-3}	7.2×10^{-4}	0.11	96

[a] 3 times the laboratory-based concentration uncertainties.

Table 13. Pearson Correlations Across All Measures of Indoor Constituents[a]

	EC	NO_2	Al	As	Ba	Ca	Cd	Ce	Cr	Cu	Fe	K	La	Mg	Mn	Mo	Ni	P	Pb	S	Sb	Se	Sr	Tl	V	Ca-XRF	Fe-XRF	K-XRF	P-XRF	S-XRF	Si-XRF	V-XRF	Zn-XRF
EC	**1.00**																																
NO_2	0.01	**1.00**																															
Al	−0.09	−0.10	**1.00**																														
As	−0.12	−0.03	0.01	**1.00**																													
Ba	−0.08	−0.09	**0.95**	−0.03	**1.00**																												
Ca	−0.02	0.18	0.26	0.21	0.24	**1.00**																											
Cd	0.09	0.13	−0.05	−0.02	−0.07	0.03	**1.00**																										
Ce	−0.04	0.26	−0.05	−0.05	−0.05	−0.06	0.28	**1.00**																									
Cr	−0.07	−0.04	0.08	0.00	0.05	0.00	−0.04	−0.04	**1.00**																								
Cu	−0.01	0.01	0.09	0.21	0.08	0.13	−0.04	−0.17	−0.02	**1.00**																							
Fe	−0.04	0.15	0.45	0.21	0.38	0.25	−0.01	−0.09	0.05	0.41	**1.00**																						
K	−0.02	0.20	0.60	−0.03	0.57	0.27	0.37	0.27	−0.01	0.33	0.30	**1.00**																					
La	−0.03	0.27	−0.06	−0.05	−0.06	−0.05	0.33	**1.00**	−0.05	−0.18	−0.10	0.29	**1.00**																				
Mg	0.10	0.22	0.01	0.15	−0.01	**0.62**	−0.03	0.02	0.00	0.27	0.09	0.24	0.02	**1.00**																			
Mn	0.16	0.16	−0.05	0.09	−0.10	0.14	0.27	0.02	0.03	0.25	0.58	0.12	0.03	0.15	**1.00**																		
Mo	0.12	−0.01	0.33	0.07	0.31	0.24	0.04	−0.23	0.09	**0.63**	**0.69**	0.29	−0.22	0.26	0.56	**1.00**																	
Ni	0.21	0.06	0.02	−0.05	0.03	0.09	0.03	−0.15	−0.02	0.20	0.17	0.06	−0.14	0.15	0.33	0.45	**1.00**																
P	−0.10	0.26	−0.05	−0.04	−0.08	−0.04	−0.01	−0.05	−0.04	−0.05	0.09	−0.09	−0.05	−0.04	0.05	−0.08	−0.06	**1.00**															
Pb	0.13	0.39	−0.04	0.01	−0.05	−0.01	0.23	0.51	−0.02	−0.01	−0.02	0.46	0.54	0.01	0.13	−0.06	−0.06	0.00	**1.00**														
S	−0.02	0.14	0.00	0.15	−0.03	0.18	0.17	0.01	−0.02	0.36	**0.76**	0.09	0.01	0.16	**0.67**	0.57	0.21	0.09	0.04	**1.00**													
Sb	**0.61**	0.30	−0.04	−0.01	−0.04	−0.06	0.05	−0.01	0.00	0.06	0.06	−0.01	−0.01	0.05	0.01	0.00	0.03	−0.03	0.07	−0.03	**1.00**												
Se	−0.12	0.10	0.05	0.01	0.01	0.08	0.15	−0.03	−0.07	0.24	0.50	0.13	−0.03	−0.01	0.33	0.37	0.02	0.40	0.18	0.60	0.00	**1.00**											
Sr	−0.10	0.03	0.47	0.05	0.51	0.15	0.01	−0.09	−0.01	**0.79**	0.36	**0.65**	−0.09	0.30	0.09	0.55	0.15	−0.06	0.02	0.18	0.02	0.15	**1.00**										
Tl	0.06	0.14	0.01	−0.03	−0.03	0.06	**0.95**	0.31	−0.03	−0.06	0.04	0.41	0.36	−0.04	0.36	0.04	0.02	−0.02	0.29	0.22	−0.03	0.18	0.00	**1.00**									
V	0.18	0.05	0.17	−0.09	0.18	0.06	0.01	−0.16	0.01	0.16	0.17	0.13	−0.15	0.19	0.23	0.47	**0.63**	−0.08	−0.10	0.12	−0.03	−0.03	0.18	0.01	**1.00**								
Ca-XRF	0.08	0.17	−0.04	0.04	−0.05	**0.67**	−0.01	−0.03	−0.02	−0.01	−0.04	0.09	−0.03	**0.85**	0.09	0.09	0.10	−0.03	0.00	0.10	−0.03	−0.04	0.00	0.00	0.13	**1.00**							
Fe-XRF	0.07	0.14	−0.01	−0.01	−0.04	0.32	0.23	−0.02	0.00	0.36	0.50	0.09	−0.01	0.11	**0.61**	0.45	0.24	−0.08	0.02	0.60	−0.05	0.29	0.14	0.29	0.19	0.14	**1.00**						
K-XRF	0.06	0.30	0.00	−0.03	−0.01	0.18	0.60	0.38	−0.05	0.33	0.03	**0.72**	0.42	0.25	0.20	0.13	−0.01	−0.08	0.60	0.16	0.01	0.16	0.43	**0.62**	−0.03	0.14	0.23	**1.00**					
P-XRF	−0.12	0.27	−0.04	−0.01	−0.08	0.02	0.05	−0.06	−0.05	0.05	0.31	−0.07	−0.06	−0.04	0.25	0.08	−0.03	**0.94**	0.00	0.40	−0.05	0.58	−0.02	0.07	−0.08	−0.01	0.15	−0.02	**1.00**				
S-XRF	−0.03	0.12	0.03	0.10	−0.01	0.24	0.23	0.00	−0.05	0.31	**0.68**	0.09	0.00	0.11	**0.63**	0.51	0.14	0.03	0.03	**0.95**	−0.05	0.59	0.15	0.31	0.07	0.11	**0.66**	0.21	0.36	**1.00**			
Si-XRF	0.27	0.19	−0.03	0.00	−0.05	0.02	0.00	−0.01	−0.02	−0.05	0.02	−0.04	−0.01	0.26	0.10	0.10	0.11	−0.02	0.02	0.01	0.12	−0.02	−0.03	0.01	0.09	0.10	0.02	−0.03	−0.04	−0.02	**1.00**		
V-XRF	0.15	0.05	0.11	−0.12	0.14	0.03	0.10	−0.06	−0.05	0.00	−0.04	0.09	−0.04	0.12	0.06	0.26	0.49	−0.10	−0.06	−0.03	−0.05	−0.09	0.07	0.10	**0.90**	0.14	0.13	0.03	−0.11	−0.02	0.04	**1.00**	
Zn-XRF	0.27	0.22	−0.07	−0.07	−0.07	0.11	0.25	0.35	−0.06	0.02	−0.08	0.27	0.37	0.02	0.15	0.03	0.09	−0.03	**0.62**	0.02	0.09	0.06	−0.06	0.30	0.07	0.01	0.23	0.42	−0.03	0.04	0.03	0.11	**1.00**

[a] Correlation coefficients above 0.6 are presented in bold.

The indoor samples were influenced by analytical issues similar to those described for the outdoor samples (see Summary Statistics — Outdoors, under Results), resulting in lower sample sizes for NO_2 and for selected elements from the ICP–MS analysis.

In general, the correlations among indoor concentrations (Table 13) were slightly weaker than among outdoor concentrations (Table 8). Some of the strong correlations found within the outdoor measurements held for the indoor measurements, especially for constituents hypothesized to be predominantly of outdoor origin (e.g., S and Se, Ni and V). However, other correlations that were significant for outdoor measurements did not hold for indoor measurements (EC was not significantly correlated with S, P, or Ca), while some new correlations emerged for indoor measurements that were not seen for outdoor measurements (such as the strong positive correlations between Ca and Mg, Pb and Zn, and Fe and S). This provided a general indication that indoor sources and the differential effects of ventilation were altering the correlation structures among particle constituents so that the correlations related to the indoor environments were different from those related to the outdoor environments.

Finally, we considered the distribution of questionnaire responses related to indoor sources and other aspects of the indoor environment, to determine whether sufficient heterogeneity existed in key activities to observe the influence of indoor sources. As indicated in Table 14, for many of the hypothesized sources of interest, such as cooking and cleaning activities, there was adequate heterogeneity within the study population in the frequency of these activities. In addition, although basic home characteristics were similar for cohort and noncohort members (Table 5), multiple activity patterns differed significantly (Table 14). Cohort members used air conditioning more frequently and opened windows less, had greater occupant density, and tended to clean and cook more, although noncohort members were more likely to have gas stoves. Thus, a comparison of these groups can provide another means to evaluate the influence of occupant activity patterns on I/O ratios, and the inclusion of both groups may make our findings more generalizable.

INDOOR–OUTDOOR RELATIONSHIPS AND PREDICTORS

To develop a better understanding of which pollutants may have their indoor concentrations most influenced by infiltration from the outdoors, which pollutants may be most influenced by indoor sources, and which are influenced by a combination of factors, we considered the relationship between measured indoor and measured outdoor concentrations. We first examined the distribution of I/O

Table 14. Distributions of Selected Household Activities for All Sampling Sessions with Complete Questionnaire Data for All Participants, Cohort Members, and Noncohort Members

	Total (N = 61)	Cohort Members (n = 38)	Noncohort Members (n = 23)
Categorical variables (%)			
Cleaning activities[a] (% 2+ activities/wk)	67	82	43
Humidifier use (% Yes)	18	24	9
Candle use (% Yes)	26	24	30
Cooking time (> 1 hr/day)	34	47	13
Gas stove prevalence[b]	69	59	86
Gas stove use (> 1 hr/day)[b]	25	32	14
Opening of windows (% Yes)	62	53	78
Air conditioning use (% Yes)	25	37	13
Continuous variables (mean [SD])			
Number of occupants spending > 4 hrs in the home	3.7 (2.0)	4.6 (1.8)	2.2 (1.2)
Occupant density (people/room)	0.99 (0.58)	1.3 (0.46)	0.43 (0.14)

[a] Dusting, cleaning, and vacuuming.

[b] Data missing for 2 sampling sessions.

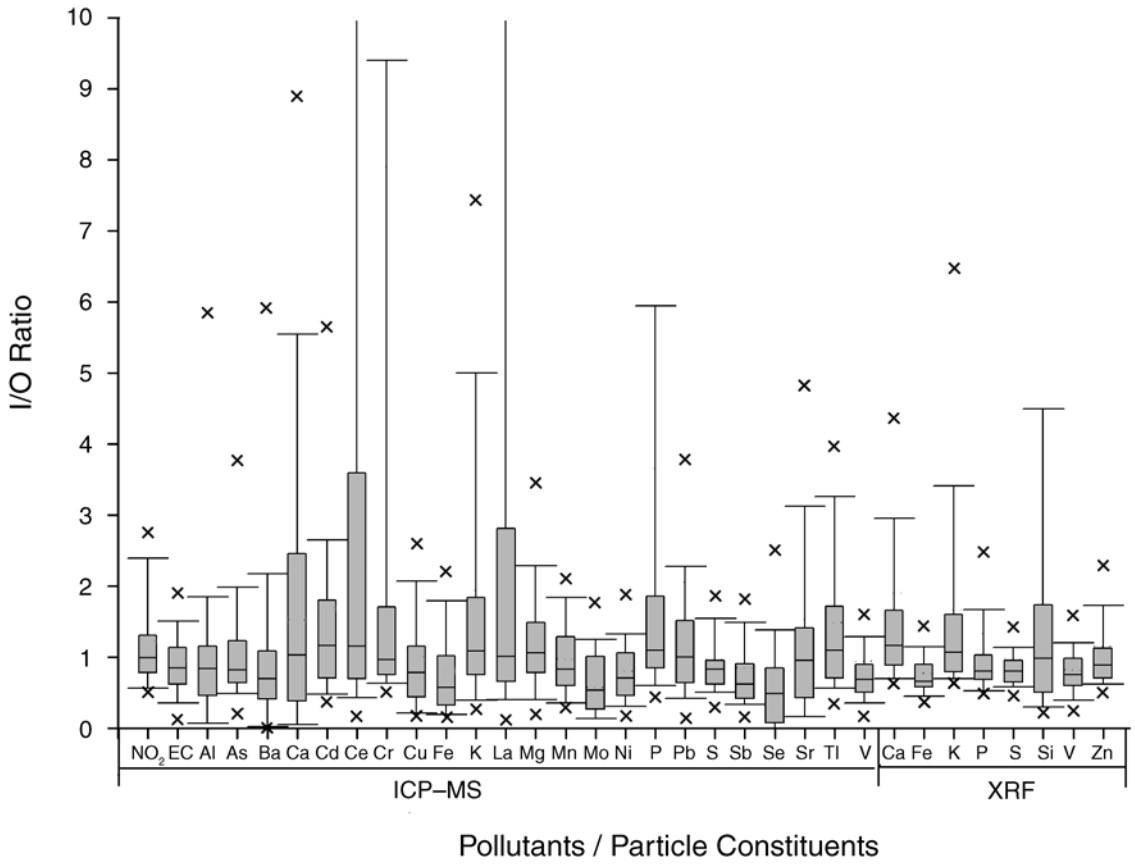

Figure 6. Distribution of I/O ratios for particle constituents and NO_2. Solid lines indicate the median, boxes indicate the interquartile range, whiskers indicate the 10th and 90th percentiles, and X's indicate the 5th and 95th percentiles.

ratios for particle constituents and NO_2. As indicated in Figure 6, lower median I/O ratios were exhibited for pollutants such as V, S, and Se, all of which would be anticipated to be dominated by outdoor sources. In contrast, median I/O ratios greater than 1 were exhibited for constituents such as Ca, K, and Cd, all of which have been associated previously with indoor sources (Table 4). Figure 6 also illustrates the presence of a number of extreme values, some of which may reflect analytical noise in the pollutant concentrations, while others may be indicative of significant indoor sources.

To more formally evaluate the relationship between indoor and outdoor concentrations in a mass-balance context, we regressed outdoor concentrations against indoor concentrations and examined the strength of association. We developed both univariate models and models that incorporated ventilation (using the I/O sulfur ratio to create a categorical variable for estimating the infiltration factor, as described in Indoor–Outdoor Concentration Modeling, under Statistical Methods and Data Analysis), and we built models with the inclusion and exclusion of influential points. As indicated in Table 15, the majority of variability in indoor concentrations of pollutants such as V, S, and Se was explained by outdoor concentrations with a simple ventilation term, providing further indication that these were outdoor-dominated constituents. Conversely, for constituents such as Ca and Si, the R^2 value was extremely low, and the influence of indoor sources was likely to be substantial. Interestingly, for both P and K, the indoor–outdoor R^2 value was substantially greater for XRF outputs than for ICP–MS outputs. This observation suggests that some indoor sources of P and K may differ in their solubility relative to outdoor sources (i.e., in the outdoor environment, P and K are generally found in highly water-soluble compounds).

When we added individual indoor-source terms to the indoor–outdoor regression models described in Table 15, we gained additional insight about the degree to which questionnaire data on source activities explained concentrations of individual constituents (data not shown). These results potentially help in the interpretation of our indoor

factor analysis models. Because of the large number of constituent–source combinations and the resultant likelihood of spurious associations, we do not consider these models as definitive, but they do provide some initial insight. For example, occupant density was most strongly associated with Cr and NO_2, with positive associations also seen with Fe, Mg, P, and Tl (data not shown). While not all of these pollutants have a connection with resuspension and other direct effects of higher occupant density, one might expect a higher level of many indoor activities to be associated with greater indoor occupant density. Interestingly, occupant density was negatively associated with multiple outdoor-dominated constituents (e.g., V, S), potentially indicating that this covariate also captures some residual ventilation effects. A number of constituents were positively associated with candle burning, with the greatest significance for Al and Ba, although there was no association with EC and a negative association with NO_2, complicating the interpretation of these findings. Frying or grilling of food was positively associated with La, K, and NO_2, but negatively associated with many other constituents, again including multiple outdoor-dominated constituents. Multiple pollutants were significantly positively associated with gas cooking or the presence of a pilot light, although NO_2 was not (however, the association with NO_2 was in the anticipated direction in both situations). In general, some of these findings do not strongly agree with the hypothesized sources in Table 4, indicating some of the challenges in building indoor factor analysis LUR models.

Note that the findings described in the previous paragraph would be slightly different if we had utilized session-average concentrations of pollutants as opposed to individual observations. We used the latter approach above to maximize our sample size and to provide inputs consistent with our factor analyses, but the former approach is more closely aligned with the time horizons of the questionnaire data and is more directly applicable to the exposure misclassification analysis. We focused on the three pollutants used in the exposure misclassification analysis (EC, NO_2, and $PM_{2.5}$). When using session-average concentrations, our study found no indoor-source terms significantly predicting indoor EC concentrations. For NO_2, gas stove usage demonstrated near significance ($P < 0.2$) and was carried forward as a candidate variable in regression models for indoor NO_2. For $PM_{2.5}$, both cooking activity and occupant density were positively associated with indoor concentrations, after controlling for outdoor concentrations, and were considered for subsequent LUR models.

One way to develop additional insight about the influence of indoor-source activities is to compare indoor–outdoor pollutant relationships for cohort versus noncohort participants. As illustrated in Tables 5 and 14, activity patterns differed substantially between these groups in spite of

Table 15. Regression Analysis of Indoor Concentrations on Outdoor Concentrations of Particle Constituents and NO_2, with and Without a Categorical Ventilation Term[a]

Pollutant	Indoor–Outdoor Model R^2 (n)	Indoor–Outdoor Model with Ventilation Term R^2 (n)
V-XRF	0.72 (67)	0.89 (67)
S-XRF	0.83 (71)	0.87 (71)
Sr	0.56 (68)	0.80 (52)
V	0.75 (68)	0.79 (68)
P-XRF	0.72 (70)	0.73 (70)
Pb	0.72 (71)	0.73 (71)
S	0.70 (68)	0.71 (68)
Sb	0.70 (67)	0.70 (67)
Ni	0.61 (66)	0.62 (66)
As	0.59 (66)	0.63 (66)
Ba	0.52 (51)	0.59 (51)
Se	0.56 (68)	0.58 (68)
Mn	0.54 (69)	0.55 (69)
Fe	0.51 (72)	0.54 (72)
Fe-XRF	0.41 (70)	0.53 (70)
K-XRF	0.49 (72)	0.51 (72)
Zn-XRF	0.36 (70)	0.41 (70)
Cu	0.40 (67)	0.40 (67)
EC	0.20 (71)	0.36 (71)
Cd	0.22 (68)	0.30 (68)
Mo	0.29 (67)	0.30 (67)
Mg	0.21 (70)	0.27 (70)
Tl	0.12 (69)	0.20 (69)
La	0.01 (68)	0.17 (68)
P	0.03 (68)	0.17 (68)
Ce	0.00 (69)	0.16 (69)
NO_2	0.14 (56)	0.16 (56)
Ca	0.02 (69)	0.15 (69)
Si-XRF	0.04 (70)	0.13 (70)
K	0.07 (58)	0.09 (58)
Al	0.09 (59)	0.09 (59)
Ca-XRF	0.05 (73)	0.06 (73)
Cr	0.00 (70)	0.02 (70)

[a] Outliers have been removed using Cook's distance $> 4/(n - P)$. Pollutants are listed, in decending order, by their R^2 value in the regression model with a ventilation term.

similar structural characteristics of their homes. As a result, the median I/O ratios of $PM_{2.5}$ and NO_2 were greater in cohort than noncohort homes (Figure 7). This indicates a greater influence of indoor sources and lower AERs in the homes of cohort participants. Similarly, the homes of cohort participants had higher median I/O ratios for particle constituents with hypothesized indoor sources and with low R^2 values (e.g., Ca, Cr, K, Si) (Table 15), while the homes of noncohort participants had higher median I/O ratios for constituents likely to be dominated by outdoor sources (e.g., Ni, S, Se, V) (Table 7; other data not shown).

Finally, to inform the structure of our indoor-concentration LUR models, we developed models to predict the infiltration factor (using the I/O sulfur ratio as a proxy, as described in prior publications, and supported by the strong relationship between its indoor and outdoor concentrations, as shown in Table 15). We first used information from tax assessor databases and other publicly available data, to determine the degree to which F_{INF} could be predicted without contacting any study participants. We then added terms related to occupant activities during the sampling week from questionnaire data. As described earlier in this report, only variables with a logical causal connection to F_{INF} were considered: For publicly available data, these included season, floor level of the home, whether homes were multiunit or single-family dwellings, and the age of the home; for questionnaire data, these included home air conditioning use and the opening of windows. We developed regression models both with F_{INF} as a continuous variable and with F_{INF} dichotomized at the median, given the possibility that small measurement errors could significantly influence the I/O ratio and related F_{INF} values.

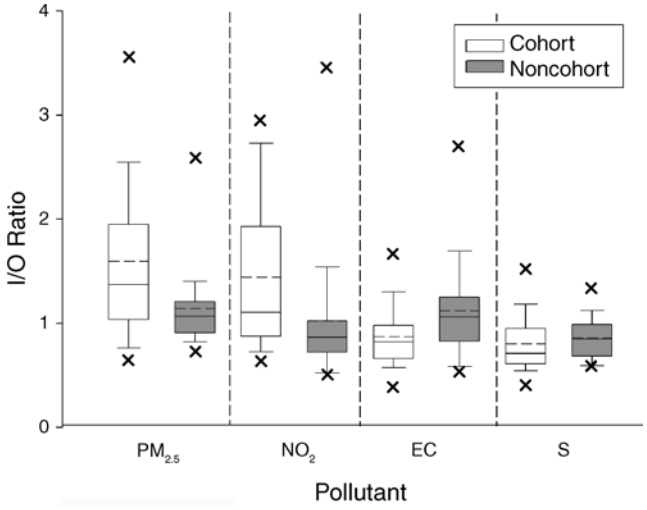

Figure 7. Distribution of I/O ratios for $PM_{2.5}$, NO_2, EC, and S stratified by cohort and noncohort homes. Solid lines indicate the median, dotted lines indicate the mean, boxes indicate the interquartile range, whiskers indicate the 10th and 90th percentiles, and X's indicate the 5th and 95th percentiles.

In this analysis, when F_{INF} was treated as a continuous variable for linear regressions (Table 16), season was the only significant publicly available term; F_{INF} was lower in the heating season than in the cooling season, as anticipated. The addition of questionnaire data increased the predictive power of the model, with opening of windows

Table 16. Predictors of F_{INF} (Proxied by the I/O Sulfur Ratio) Based on Publicly Available and Questionnaire Data

	Linear Regression (Continuous F_{INF})		Logistic Regression (Dichotomized F_{INF})	
	Publicly Available Data	Publicly Available Data and Questionnaire Data	Publicly Available Data	Publicly Available Data and Questionnaire Data
R^2	0.21	0.38		
Area Under Curve			0.82	0.86
Predictors	β (SE)	β (SE)	OR (95% CI)	OR (95% CI)
Season (reference: cooling)	−0.13 (0.04)[a]	−0.05 (0.07)	0.11 (0.03–0.43)[a]	0.17 (0.02–1.7)
Floor (reference: 1st level)	0.04 (0.06)	0.01 (0.06)	4.0 (0.72–22)	5.0 (0.71–35)
Housing type (reference: single family)	0.01 (0.07)	0.00 (0.06)	0.52 (0.08–3.6)	0.22 (0.02–2.1)
Year home built (reference: before 1950)	−0.05 (0.06)	−0.03 (0.06)	0.12 (0.02–0.86)[a]	0.15 (0.02–1.4)
Open windows (reference: no)	N/A	0.17 (0.06)[a]	N/A	7.2 (0.98–54)[a]
Air conditioning use (reference: no)	N/A	−0.07 (0.06)	N/A	0.19 (0.02–1.5)

[a] Indicates $P < 0.1$.

during the sample period as the only statistically significant term. In fact, when opening of windows was included, season lost statistical significance. Similarly, when F_{INF} was dichotomized for logistic regressions (Table 16), season and the year that the home was built were significant predictors in a model without questionnaire data, but when questionnaire data was included, only opening of windows retained statistical significance ($P < 0.1$). We therefore concluded that questionnaire information about open window status was the best available predictor of F_{INF} for our cohort and that, given a lack of significance for other terms and our use of categorical variables for ventilation, this covariate could be used directly for our regression modeling. Of note, the use of open windows as a ventilation proxy agrees with a similar study conducted in Boston (Brown 2006) that found higher AERs in homes with open windows. Moreover, this study concluded that an open-windows covariate may better estimate the exchange of indoor and outdoor air than measured AERs for multiunit buildings, such as those in the current study. This is because measured AERs cannot distinguish between make-up air from adjacent apartments and air from outdoors (Brown 2006).

LUR MODELS FOR INDOOR EC, $PM_{2.5}$, AND NO_2

For comparability to our outdoor-concentration model-building approach, and for applicability to our exposure misclassification analysis, we first constructed LUR models for indoor concentrations of EC, $PM_{2.5}$, and NO_2, using traffic and other GIS terms along with indoor source data, ambient monitoring data, and ventilation data. We initially considered models without ventilation terms, to focus on the most significant "main effects" of traffic, presented in Table 17. The unweighted cumulative traffic density within 50 m of the home was associated with higher indoor NO_2; gas stove usage and ambient NO_2 were nearly significant. A diesel-traffic indicator predicted indoor EC, with lower concentrations of EC associated with homes located further from a designated truck route. Ambient EC also predicted indoor EC, which appeared to have no significant indoor sources. No traffic variable significantly predicted indoor $PM_{2.5}$, which was best explained by ambient $PM_{2.5}$, cooking time, and occupant density.

To consider the robustness of the traffic-variable selection for each pollutant, the posterior probabilities of models using the different traffic variables were calculated and grouped based on the GIS algorithm used to create each one (Table 18). Models with posterior probabilities greater than three times the prior probability (4.2%) included unweighted cumulative traffic density within 50 m of the home, which yielded the highest probability (26.5%) for NO_2, and distance from a designated truck route (14.3%), for EC. ADT had the highest posterior probability in the $PM_{2.5}$ models (8.34%), but that probability was less than twice the prior probability, and other measures had comparable probabilities. We calculated these posterior probabilities using a range of c values (5–100) (see equation 7) with similar results (not shown).

In the Bayesian analysis results, shown in Table 18, all posterior probabilities were under 30%, emphasizing the difficulty in choosing the correct model with a small data set and many correlated predictors. For NO_2, models describing traffic closer to the home (buffers of 50–100 m) generally had the highest probabilities, agreeing with previous studies showing outdoor NO_2 levels decreasing with increasing distance from the road (Roorda-Knape et al. 1999; Gilbert et al. 2003). For EC, the highest-probability traffic terms were related to truck traffic, as anticipated

Table 17. Identification of Traffic Indicators Contributing to Indoor Concentrations After Adjustment for Ambient Concentrations and Indoor Source Terms[a]

Pollutant	R^2	Model	β (SE)	P value
NO_2 (ppb)	0.20	Ambient concentrations	0.66 (0.35)	0.06
		Gas stove usage	5.0 (3.0)	0.11
		Unweighted traffic density within 50-m buffer	0.06 (0.03)	0.02
$PM_{2.5}$ (μg/m^3)	0.36	Ambient concentrations	0.99 (0.25)	< 0.01
		Cooking time	5.1 (2.9)	0.08
		Occupant density	5.2 (2.2)	0.02
EC ($m^{-1} \times 10^{-5}$)	0.21	Ambient concentrations	0.26 (0.09)	< 0.01
		Distance to nearest designated truck route	-7.2×10^{-5} (4.2×10^{-5})	0.01

[a] Only covariates with $P < 0.2$ are shown.

Table 18. For Indoor LUR Models, GIS-Based Traffic Variables Grouped by Algorithm Used to Create Them and Their Posterior Probabilities[a]

	NO_2	$PM_{2.5}$	EC
Cumulative Traffic Density Scores (Vehicle-Meters/m²/Day)			
Density of larger roads			
Within 200 m	2.39	3.48	3.02
Unweighted traffic density			
Within 50-m buffer	**26.5**	3.08	2.97
Within 100-m buffer	2.15	2.90	2.95
Within 200-m buffer	2.23	4.07	3.18
Within 500-m buffer	2.46	5.33	3.82
Kernel-weighted density			
Within 50-m buffer	6.64	3.13	3.12
Within 100-m buffer	10.3	3.16	3.00
Within 200-m buffer	1.93	3.02	3.44
Within 300-m buffer	2.25	4.30	3.75
Within 500-m buffer	3.25	5.40	3.39
Summary Measures (m)			
Total roadway length			
Within 50 m	5.76	3.48	3.36
Within 100 m	2.31	4.40	3.41
Within 200 m	2.30	5.18	2.95
Within 300 m	2.42	5.78	4.33
Distance-Based Measures (m)			
Distance to nearest:			
Larger road (> 8500 cars/day)	3.90	5.43	3.57
Major road (> 13,000 cars/day)	3.93	6.28	3.65
Highway (> 19,000 cars/day)	2.01	2.97	3.72
Designated truck route	2.16	4.37	**14.3**
Characteristics of Nearest Major Road			
ADT	2.04	8.34	5.04
ADT/distance to major road	2.27	3.00	3.45
Diesel fraction	2.09	2.84	3.77
Trucks per day	2.45	2.87	8.63
Trucks per day/distance to major road	4.06	3.16	2.96
Population Density (for Census Block Containing Sampling Site)			
Population density	2.18	4.06	4.19

[a] Covariates with posterior probabilities 3 times (12.6%) greater than the prior probability (4.2%) are presented in bold.

(Table 18). In contrast, the traffic model with the highest probability (ADT) was not significant for indoor $PM_{2.5}$. No $PM_{2.5}$ model yielded probabilities over 10%, suggesting little differential information value across covariates. This finding was not unexpected, as $PM_{2.5}$ exhibits less spatial heterogeneity than do EC and NO_2, and because outdoor $PM_{2.5}$ was largely explained by temporal terms (Table 9).

To account for multiple testing, sensitivity analyses were conducted to compare the posterior probabilities for models with and without a traffic term (M_k and M_0, respectively), assuming an equal chance of traffic affecting indoor concentrations as not (data not shown). For all pollutants, the models without a traffic term had high probabilities (77.3% for NO_2, 84.3% for $PM_{2.5}$, 84.6% for EC), reflecting both the presumed prior probabilities and the relatively small amount of variability explained by traffic terms. The highest probabilities for those models with the traffic term were 6.02% (using unweighted cumulative traffic density within 50 m of the home) for NO_2, 1.31% (using ADT) for $PM_{2.5}$, and 2.21% (using distance from a designated truck route) for EC. This observation indicates the difficulty of relating traffic variables to indoor pollutant concentrations in homes within urban areas with less spatial variation than metropolitan regions that include urban, suburban, and rural sites. It also indicates the strong effects of indoor sources and ventilation. Small sample size and multiple testing also contribute to the difficulty of definitively identifying which (if any) traffic terms should be considered.

As described earlier, open windows appears to serve as the best ventilation proxy where AERs and I/O concentration ratios cannot be measured. Although this term would theoretically modify both indoor- and outdoor-source contributions, we apply it here only to outdoor-source contributions, due to our limited statistical power, reduced sample size with session-averaged concentrations, and statistical instability resulting from the modification of indoor sources (related to our use of many categorical indoor-source terms). The final indoor LUR models, including only the interaction terms with $P < 0.2$, are shown in Table 19. For NO_2 and EC, traffic variables were significantly positively modified by open windows. For $PM_{2.5}$, the effect of ambient concentrations was also significantly greater in homes with open windows. Including a ventilation proxy increased the R^2 value from 0.20 to 0.25 for NO_2, from 0.36 to 0.40 for $PM_{2.5}$, and from 0.21 to 0.32 for EC. (The table only shows the R^2 values after the inclusion of the ventilation proxy.)

CONSTRAINED FACTOR ANALYSES FOR INDOOR CONCENTRATIONS

As described in LUR Modeling — Indoor Factor Scores, under Statistical Methods and Data Analysis, we first constructed constrained factor analyses based on the measured indoor concentrations of the particle constituents listed in Table 3 (along with Cd, Ce, and Tl), for which a

Table 19. Indoor LUR Models Accounting for Effect Modification by Ventilation as Proxied by Open Windows[a]

Pollutant	R^2	Model	β (SE)	*P* Value
NO_2 (ppb)	0.25	Ambient concentrations	0.79 (0.35)	0.03
		Gas stove usage	6.8 (3.1)	0.04
		Unweighted traffic density within 50-m buffer × Open windows = Yes	0.07 (0.03)	0.01
		Unweighted traffic density within 50-m buffer × Open windows = No	−0.03 (0.06)	0.62
$PM_{2.5}$ (μg/m^3)	0.40	Ambient concentrations × Open windows = Yes	0.98 (0.32)	< 0.01
		Ambient Concentrations × Open windows = No	0.64 (0.32)	0.05
		Cooking Time	6.2 (2.9)	0.04
		Occupant Density	6.5 (2.3)	0.01
EC ($m^{-1} \times 10^{-5}$)	0.32	Ambient concentrations	0.38 (0.09)	< 0.0001
		Distance to nearest designated truck route × Open windows = Yes	-9.2×10^{-5} (4.1×10^{-5})	0.03
		Distance to nearest designated truck route × Open windows = No	1.0×10^{-4} (5.9×10^{-5})	0.86

[a] Only interaction terms with $P < 0.2$ are shown.

six-factor model provided the most reasonable compromise among BIC, K–G criterion, and cross-validation approaches and had reasonable physical interpretability. The loading matrix is shown in Figure 8, with darker shading representing higher factor loadings.

More constituents loaded strongly onto Factor 1 than onto other factors; these constituents included S and Se, which are associated with coal combustion and with few hypothesized indoor sources. Other constituents with high loadings on Factor 1 (e.g., Pb, Mn) are not normally associated with coal combustion but reflect broader-scale outdoor combustion sources, including industries not present in our region. We therefore interpreted this factor to represent outdoor concentrations of sources that covary strongly over time, given the modifying influence of ventilation, and define it as *long-range transport.*

Factor 2 was dominated by Ni and V, common markers of fuel oil combustion, similar to Factor 4 in the outdoor-concentration factor analysis model (Figure 4). Factor 2 also included other outdoor-dominant pollutants with expected local sources (e.g., traffic, home heating), such as EC, Sb, and Mn, so we consider this factor as *fuel oil/diesel combustion*. Factor 3 showed heavy loadings on Al and Ba, which we previously proposed were associated with road dust and resuspension, with a potential influence of brake wear (Schauer et al. 2006). However, multiple indoor-source terms were significantly associated with both Al and Ba, and indoor–outdoor relationships were weak for some of the pollutants loading on this factor (Table 15). We defined Factor 3 as *road dust and resuspension*, but this may have reflected a combination of indoor and outdoor sources. Factor 4 included heavy loads on La and Ce and weaker associations with K, Ca, and Mg. Given relatively weak indoor–outdoor relationships for these pollutants and an indication that some of them are associated with gas cooking and candle burning, we defined Factor 4 as *indoor combustion*. Factor 5 was dominated by Cd, Tl, EC, and K, most of which have been associated with *indoor smoking* in previous studies (see Table 4). Factor 6 showed a heavy loading of P, which has been associated with *indoor cleaning* (Koistinen et al. 2004).

LUR MODELS FOR INDOOR LATENT VARIABLES

Using a stepwise model-building structure, we developed regression models for the indoor-concentration latent variables as a function of covariates representing indoor and outdoor sources as well as ventilation and time-varying factors. The initial step in this analysis was to develop candidate predictors in various source categories, focusing on the outdoor-source indicators previously found significant in the outdoor factor analysis LUR models (Table 11)

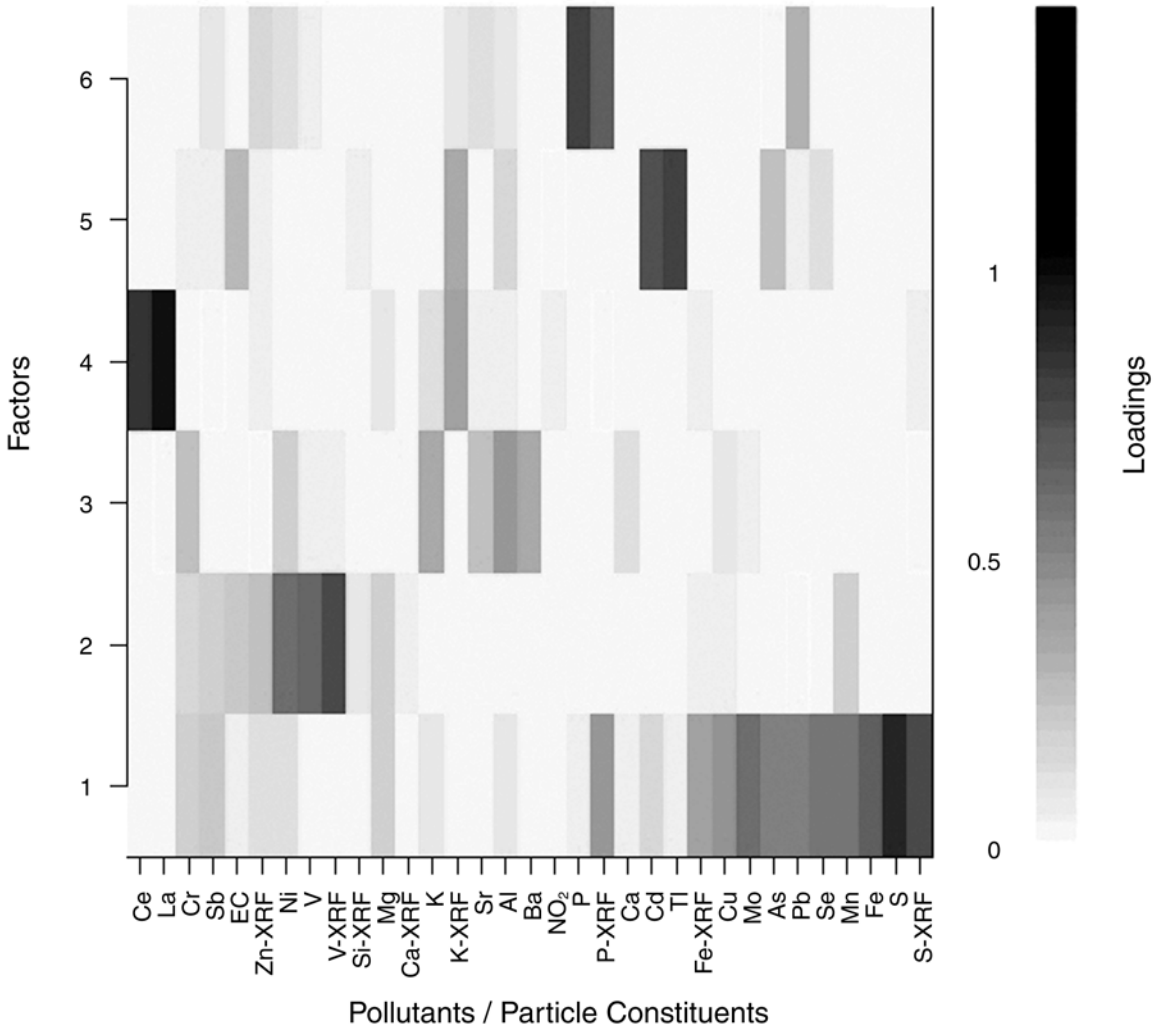

Figure 8. Factor loading matrix for indoor concentrations using a six-factor model. Pollutants and particle constituents are sorted by their position in a hierarchical clustering dendrogram, to optimize visual interpretation by placing correlated constituents together.

as well as central site monitoring data and questionnaire-based indoor-source covariates. As indicated in Table 20, the candidate variables agreed in some cases with initial factor interpretations, although generally only for hypothesized outdoor sources. For example, Factor 1 (long-range transport) was significantly correlated with central site $PM_{2.5}$, much more strongly so than were the other factors. Factor 2 (fuel oil/diesel combustion) was higher during the heating season and most strongly correlated with central site NO_2, roadway length, and high-density residential land use. Factor 4 (indoor combustion) was associated with number of people living in the home and several indoor combustion activities. The other factors did not correlate well with initial factor interpretations.

When these candidate variables were used in forward stepwise modeling, the results were generally consistent with univariate correlations (Table 21). Predictive power was greatest for Factor 1 (long-range transport), which was predicted only by temporally varying central site $PM_{2.5}$ and season (sequential $R^2 = 0.71$). Indoor-source activity or outdoor sources representing spatial variability did not improve model fit, reaffirming our original interpretation. Although population-density terms displayed univariate significance for Factor 2 (Table 20) and predicted outdoor concentrations of a fuel oil factor (Table 11), only central site NO_2 and roadway length within 300 m retained significance (Table 21). Frying and grilling retained significance as a predictor of Factor 4 (indoor combustion). As expected,

Table 20. Pearson Correlations Between Candidate Source Terms and Factors[a] Derived from Indoor Concentrations ($P < 0.3$)

Candidate Source Terms	Pearson Correlation (P Value)					
	Factor 1: Long-Range Transport	Factor 2: Fuel Oil / Diesel Combustion	Factor 3: Road Dust and Resuspension	Factor 4: Indoor Combustion	Factor 5: Indoor Smoking	Factor 6: Indoor Cleaning
Outdoor sources						
Long-range / meteorology						
Central site $PM_{2.5}$	0.83 (< 0.0001)	0.20 (0.08)			0.35 (0.02)	
Central site NO_2		0.29 (0.01)				
Cooling season	0.50 (< 0.0001)	−0.18 (0.09)		0.17 (0.12)	0.21 (0.05)	
Traffic-related sources						
Roadway length within 100 m		0.14 (0.22)		−0.16 (0.18)	0.22 (0.06)	
Roadway length within 200 m	−0.25 (0.03)	0.22 (0.06)		0.25 (0.03)		
Roadway length within 300 m	−0.18 (0.12)	0.34 (0.003)				
Diesel fraction	0.16 (0.28)			0.35 (0.02)		
Distance to nearest truck route		−0.20 (0.09)				
Population density						
High density residential area	−0.22 (0.06)	0.19 (0.10)				−0.21 (0.07)
Block group population density	−0.14 (0.22)	0.31 (0.006)		−0.22 (0.06)		
Population density in 200 m	−0.12 (0.29)	0.34 (0.003)		−0.18 (0.12)		
Fuel oil use within 200 m	−0.13 (0.28)	0.17 (0.14)	0.13 (0.26)	−0.20 (0.09)		
Other outdoor sources						
Construction		−0.21 (0.09)				
Outdoor grilling / smoking	0.17 (0.15)	−0.13 (0.26)		−0.14 (0.26)	0.17 (0.16)	
Floor (monitor height)			0.26 (0.02)	0.12 (0.18)		
Indoor Sources						
Occupancy						
People / rooms	−0.14 (0.19)	0.28 (0.01)	−0.16 (0.15)	0.15 (0.18)		
Number of rooms		0.22 (0.04)		0.24 (0.02)		
Number of people	0.13 (0.25)		−0.29 (0.008)	0.29 (0.007)		
Indoor combustion						
Candle burning	−0.14 (0.21)	0.20 (0.07)	0.15 (0.17)			
Pilot light				−0.18 (0.13)		
Frying / grilling		−0.13 (0.25)	−0.20 (0.07)	0.17 (0.13)		0.17 (0.12)
Gas cooking	−0.27 (0.03)	0.25 (0.03)		0.17 (0.16)	−0.16 (0.20)	
Burnt food		−0.14 (0.20)		−0.14 (0.19)	−0.13 (0.24)	
Indoor resuspension						
Carpeting						
Dusting	−0.17 (0.12)	0.16 (0.15)				
Sweeping		0.13 (0.23)	−0.22 (0.05)	0.14 (0.19)		0.12 (0.28)
Vacuuming		0.22 (0.04)		−0.11 (0.30)		−0.19 (0.08)
Wearing shoes				0.16 (0.14)		
Indoor chemical						
Cleaning	−0.16 (0.13)	0.19 (0.08)	−0.13 (0.24)			
Washing furniture, upholstery			−0.20 (0.07)	0.17 (0.11)		0.15 (0.18)

[a] Factor interpretations (e.g., long range transport, etc.) are based on constituent loadings.

Table 21. Regression Models for Effects of Source Terms on Factors[a] of Indoor Concentrations

	Factor 1: Long-Range Transport		Factor 2: Fuel Oil / Diesel Combustion		Factor 3: Road Dust and Resuspension		Factor 4: Indoor Combustion		Factor 5: Indoor Smoking		Factor 6: Indoor Cleaning	
Candidate Source Terms	β (*P* Value) (*n* = 74)	Seq R^2	β (*P* Value) (*n* = 74)	Seq R^2	β (*P* Value) (*n* = 74)	Seq R^2	β (*P* Value) (*n* = 71)	Seq R^2	β (*P* Value) (*n* = 74)	Seq R^2	β (*P* Value) (*n* = 71)	Seq R^2
Outdoor Sources												
Central site $PM_{2.5}$ × open windows	0.094 (< 0.0001) (no interaction)	0.68							7.83×10^{-3} (0.55) 5.53×10^{-3} (0.03)	0.12 0.18		
Central site NO_2			53.5 (0.02)	0.09								
Cooling season	0.324 (0.01)	0.71										
Roadway length within 100 m									2.50×10^{-3} (0.06)	0.24		
Roadway length within 300 m			7.91×10^{-5} (0.005)	0.18								
Diesel fraction on nearest road within 100 m							6.00 (0.02)	0.12				
High density residential area											−0.138 (0.03)	0.05
Indoor Sources												
Occupants					−0.130 (0.008)	0.08						
Frying / grilling							0.083 (0.02)	0.25				
Sweeping							0.352 (0.08)	0.31			0.123 (0.09)	0.08
Model R^2		0.71		0.18		0.08		0.31		0.24		0.08

[a] Factor interpretations (e.g., long range transport, etc.) are based on constituent loadings.

given poor univariate performance (Table 20), the remaining factors had indoor models with poor predictive power and poor physical interpretability.

Sensitivity Analyses

Given the generally poor performance of many of the regression models in Table 21, we conducted a series of sensitivity analyses to better understand whether this performance was related to our model-building process or other issues. First, because of our complex model-building process with many candidate variables, we tested the sensitivity of results to permutations in our analytic methods. We tested whether hypothesized source terms explained additional variability in final models, but that was not the case. In addition, we tested covariate selection by examining terms selected by the random forest algorithm, which indicated that covariates previously selected for each factor were informative relative to other covariates. For example, for Factor 2 (fuel oil/diesel combustion), population density was among the highest-rank covariates selected by the random forest algorithm, reinforcing its significance within our outdoor factor analysis model. For Factor 6 (indoor cleaning), both cleaning and vacuuming were among the highest-rank covariates selected by the random forest algorithm, in spite of their lack of significance in the final multivariate model shown in Table 21. The random forest algorithm identified few covariates that were predictive of Factors 3 through 6, reinforcing the weak model performances for indoor-dominated factors shown in Table 21.

Next, we tested all models for robustness to repeated measures using random effects by household, which eliminated a few covariates but produced only modest changes

to the final models. Because of high correlations among covariates and our small sample size, we sequentially removed each term from the final models, to ensure that each term retained significance independent of other covariates. Similarly, we reversed the order of terms introduced into the model, ensuring that all terms independently met the retention criteria. No covariates were eliminated based on these criteria. Next, we compared final models constructed using forward selection to those produced using backwards elimination, using candidate covariates with $P < 0.25$ and applying the same retention criteria at each step. We produced similar models by both methods, and our forward-selection models offered comparable or better interpretability in all cases. Finally, we again plotted the relationship between selected covariates and factor scores, to ensure that results were robust to influential points. Several remaining outliers were identified in this way, but their removal did not significantly alter results. Final models are reported with these outliers retained.

Our poor performance in predicting indoor-attributable factors may be associated with limitations in our questionnaire data, but could also be related to the fact that some of the factors may represent indoor and outdoor sources that are correlated because of ventilation characteristics, indoor activities that are correlated with home or occupant characteristics that also predict outdoor-source proximity, or correlations that are due to chance. We therefore constructed a second factor analysis using the residuals from the regression models in Table 15, in which indoor concentrations were regressed against the corresponding outdoor concentrations, modified by a categorical ventilation marker. The optimal model, based on BIC, K–G, and cross-validation criteria, included six factors, as presented in Figure 9.

In this model, Factor 1 includes many constituents of predominantly outdoor origin, such as S, Se, Ni, and V, although this modeling approach aimed to remove outdoor contributions to indoor concentrations. This factor likely represents *residual outdoor sources*, potentially due to the simple categorical ventilation term used. Factor 2 includes La and Ce, similar to Factor 4 from the indoor-concentration factor analysis (Figure 8), thus we also define this factor as *indoor combustion*. Factor 3 shows heavy loadings of Al and Ba, previously characterized as *road dust and resuspension*. Similar to Factor 1, this factor may include some residual contributions from outdoor sources, given our simple categorical ventilation term, or may represent indoor sources. Factor 4 is dominated by P, as is Factor 6 from the indoor-concentration factor analysis (Figure 8), which has been associated with *indoor cleaning*. Ca and Mg load heavily on Factor 5; the indoor-concentration factor analysis did not include a distinct factor for these crustal elements, which may be associated with *indoor resuspension*. Finally, Factor 6 shows heavy loadings of Tl and Cd, which were previously associated with *indoor smoking*. The general concordance between the groupings identified by the indoor-concentration factor analysis (Figure 8) and the indoor–outdoor residual factor analysis (Figure 9) suggests that we have reasonably identified indoor sources. The residual contribution of outdoor sources indicates the difficulty of isolating indoor and outdoor contributions.

We subsequently explored whether regression models for the indoor-attributable factors in Figure 9 provided more informative interpretations than the regression models for the indoor-concentration factors. Following a similar forward stepwise regression modeling procedure, we found modest improvements in interpretability for some factors, but in most cases, predictive power was poor and the physical interpretability remained limited (results not shown).

IMPLICATIONS OF EXPOSURE MISCLASSIFICATION

As described in Exposure Misclassification Analysis, under Statistical Methods and Data Analysis, our simulated epidemiologic study and corresponding exposure misclassification analysis were based on our indoor LUR models for EC, NO_2, and $PM_{2.5}$ (Table 19) along with alternative exposure models capturing indoor concentrations (both good exposure surrogates and poor exposure surrogates).

The distributions of the simulated ORs from the various LUR and exposure models and from the gold standard (the yearly indoor pollutant concentrations generated by our model) are shown in Figure 10 (NO_2), Figure 11 ($PM_{2.5}$), and Figure 12 (EC). For the exposure models using poor exposure surrogates, only results for the surrogate distance to the nearest major road are shown because they are representative of those observed across this category of models. For each of the three pollutants, the range of observed ORs obtained using the models with poor exposure surrogates was very large, with medians of approximately 1.00 regardless of the true OR. This finding was expected given the weak relationships between the traffic terms in these models and measured indoor concentrations. For exposure models using good exposure surrogates (e.g., unweighted traffic density within a 50-m buffer, gas stove usage), the ranges were smaller and the median estimated ORs were closer to the true health effect. The previously developed LUR models for NO_2 and $PM_{2.5}$ performed well in comparison to the gold standard, but somewhat more poorly for EC (Figure 12).

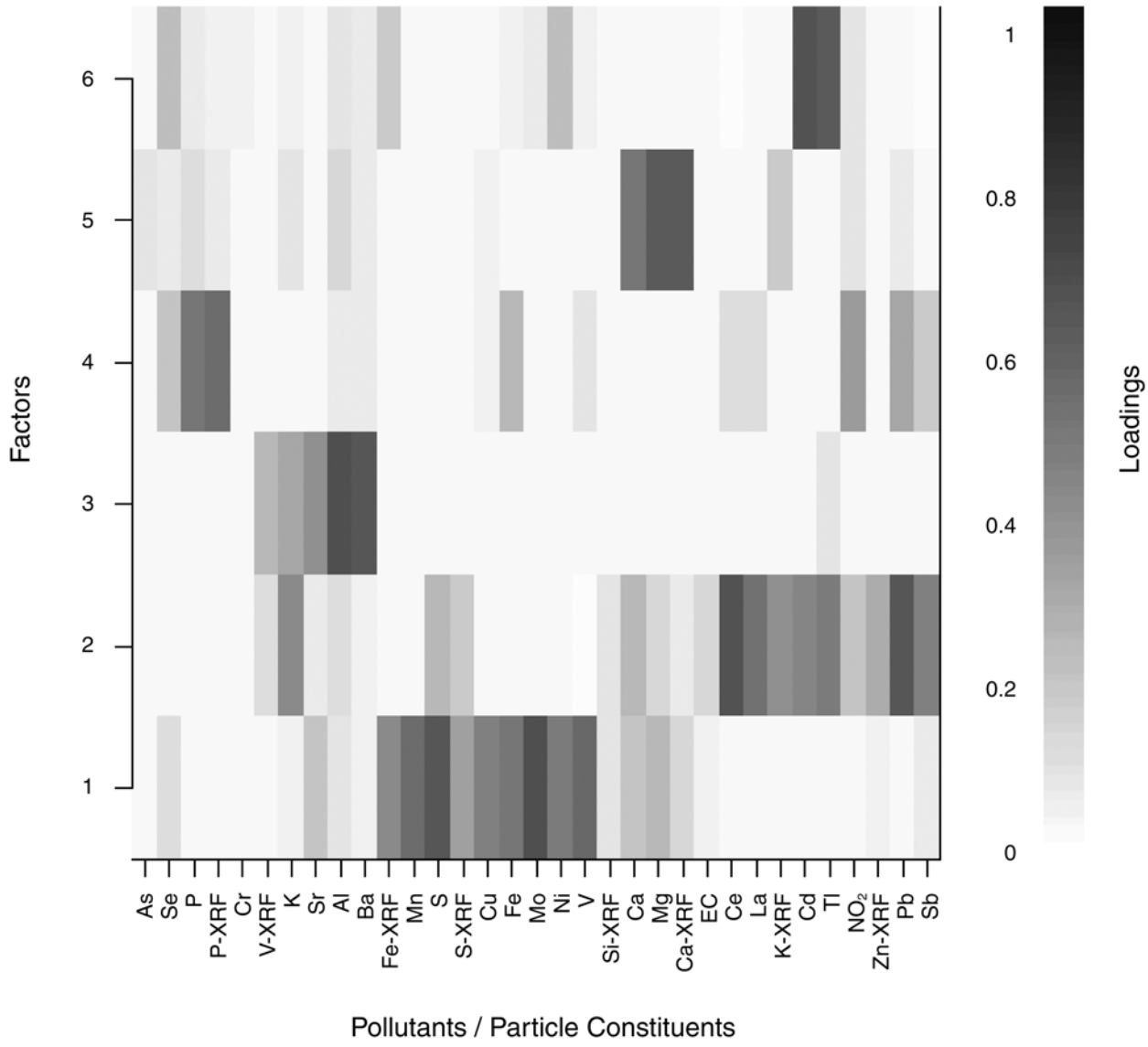

Figure 9. Factor loading matrix derived using residuals from a regression model of indoor concentrations on outdoor concentrations, modified by a categorical ventilation marker, using a six-factor model. Pollutants and particle constituents are sorted by their position in a hierarchical clustering dendrogram, to optimize visual interpretation by placing correlated constituents together.

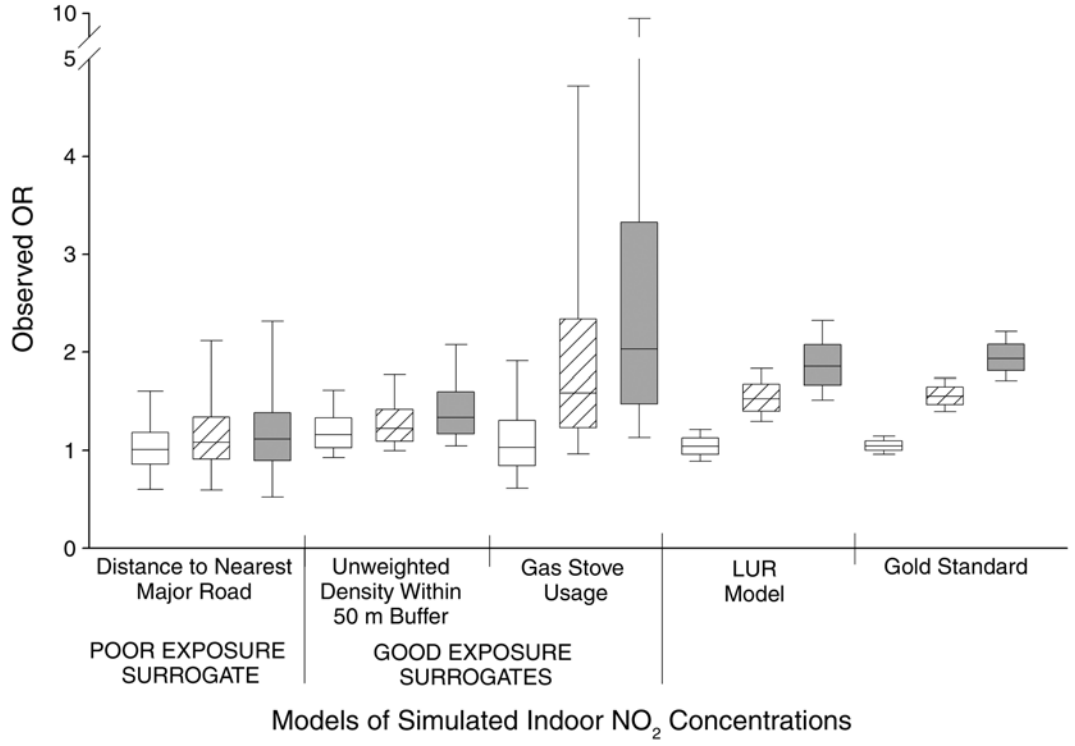

Figure 10. Distribution of estimated ORs per interquartile increase in concentration using various models of simulated indoor NO_2 concentrations given different true ORs. White boxes indicate a true OR of 1.05, crosshatched boxes indicate a true OR of 1.50, and grey boxes indicate a true OR of 2.00. Solid lines indicate the median, boxes indicate the interquartile range, and whiskers indicate the 10th and 90th percentiles. The gold standard is defined as the simulated average annual indoor concentrations at every home.

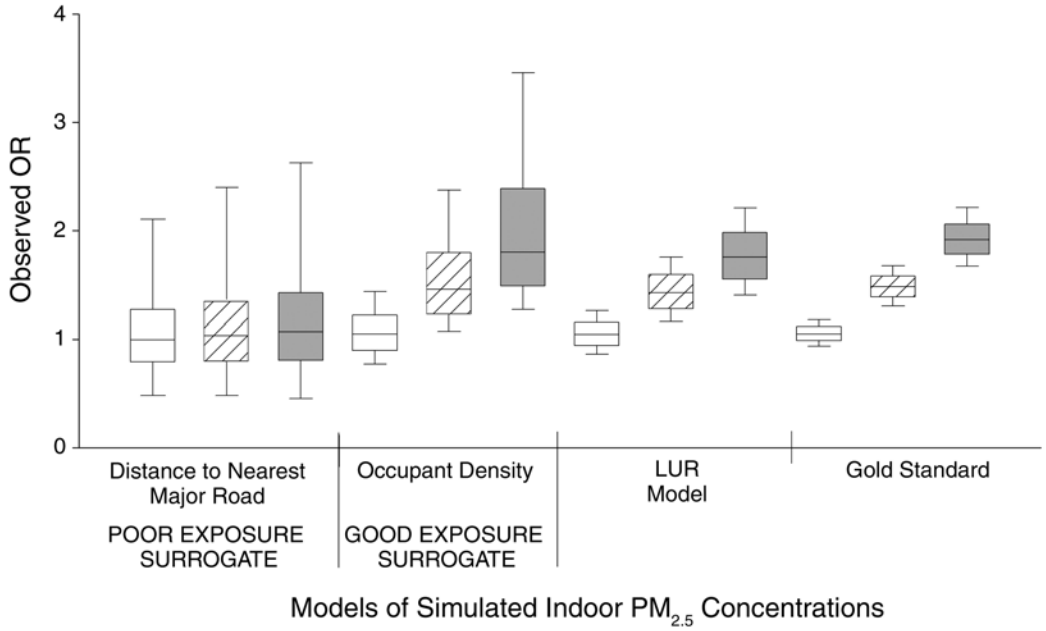

Figure 11. Distribution of estimated ORs per interquartile increase in concentration using various models of simulated indoor $PM_{2.5}$ concentrations given different true ORs. White boxes indicate a true OR of 1.05, crosshatched boxes indicate a true OR of 1.50, and grey boxes indicate a true OR of 2.00. Solid lines indicate the median, boxes indicate the interquartile range, and whiskers indicate the 10th and 90th percentiles. The gold standard is defined as the simulated average annual indoor concentrations at every home.

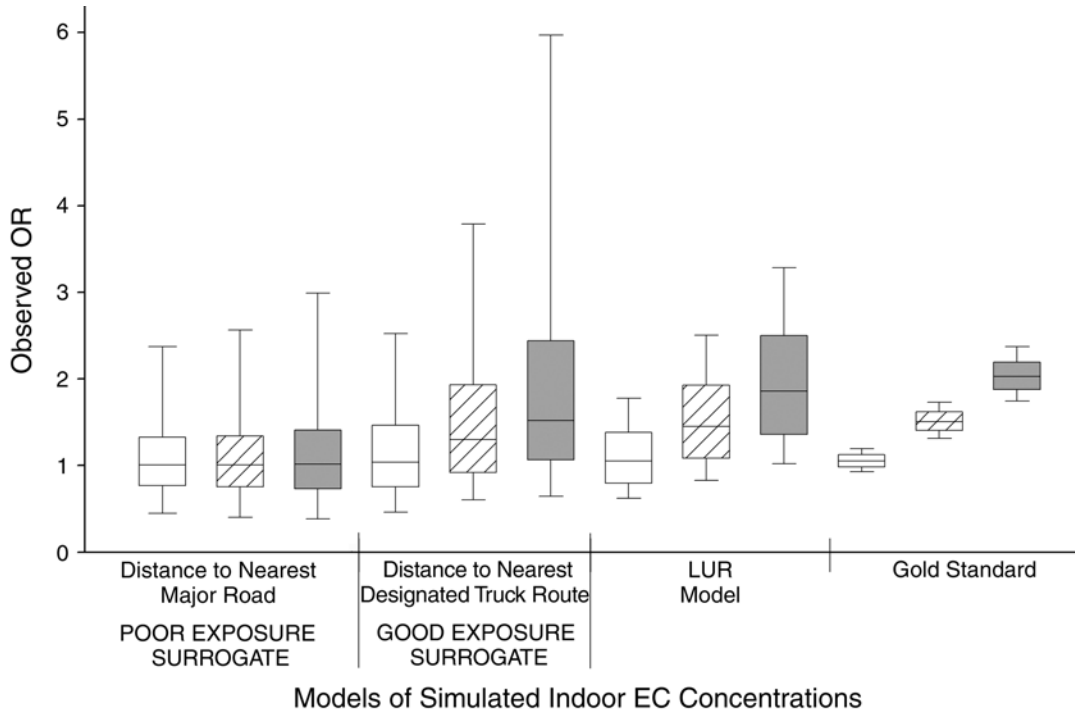

Figure 12. Distribution of estimated ORs per interquartile increase in concentration using various models of simulated indoor EC concentrations given different true ORs. White boxes indicate a true OR of 1.05, crosshatched boxes indicate a true OR of 1.50, and grey boxes indicate a true OR of 2.00. Solid lines indicate the median, boxes indicate the interquartile range, and whiskers indicate the 10th and 90th percentiles. The gold standard is defined as the simulated average annual indoor concentrations at every home.

Table 22. Estimated Bias, Median SE, and RMSE for Each Exposure Model of Indoor NO_2 Concentrations Given Different True Health Effect Estimates (β_X) and Corresponding ORs

	β_X (OR)								
	0.004 (1.05)			0.04 (1.50)			0.06 (2.00)		
Surrogate Model	Bias	Median SE	RMSE	Bias	Median SE	RMSE	Bias	Median SE	RMSE
Distance to nearest major road	−0.011	0.022	1.31	−0.030	0.024	1.37	−0.059[a]	0.026	1.37
Unweighted traffic density within 50-m buffer	−0.002[a]	0.013	0.283	−0.012[a]	0.015	0.107	−0.030[a]	0.017	0.389
Gas stove usage	0.007	0.028	0.344	0.009	0.032	0.497	0.016	0.035	0.717
LUR model	−0.001[a]	0.010	0.012	−0.001[a]	0.012	0.012	−0.003[a]	0.015	0.016
Gold standard[b]	0.000	0.006	0.006	0.000	0.008	0.008	0.000[a]	0.009	0.009

[a] Statistically significantly different from zero based on standard error of bias estimates ($P < 0.05$).

[b] Gold standard is defined as the simulated indoor concentrations at every home.

We can more formally evaluate model performance by examining estimated biases, median SEs, and RMSEs for the surrogate exposure models for each of the pollutants. These performance measures are shown in Table 22, for NO_2; Table 23, for $PM_{2.5}$; and Table 24, for EC. Results are shown for each hypothetical true health-effect estimate (β_x) associated with an interquartile increase in pollutant concentrations. The corresponding ORs (i.e., 1.05, 1.50, and 2.00) are included to provide interpollutant comparisons. The results from these tables suggest that, for all pollutants, the estimated bias is generally small for the LUR models, an expected finding given our simulation framework (using

these LUR models to generate the gold standard). However, the bias was statistically significantly different from zero based on the standard error of bias estimates ($P < 0.05$) in nearly all cases and was negative in all cases, albeit relatively small in magnitude.

For the other exposure models, the biases were somewhat greater, although the degree and statistical significance with respect to simulation variability depended on the pollutant, model, and OR. For NO_2 (Table 22), the gas stove usage model was positively biased but not significantly so. The unweighted traffic-density model was significantly negatively biased at all ORs. For $PM_{2.5}$ (Table 23), the bias was not statistically significant for the models using distance to nearest major road or occupant density; however, the mean biases were large for models using distance to the nearest major road for all ORs and for the occupant density model at an OR of 1.05. For EC (Table 24), the distance to the nearest designated truck route model was downwardly biased but only significantly so at an OR of 1.5.

Using the LUR models, the median SEs were approximately twice that of the gold standard median SEs for NO_2 and $PM_{2.5}$, but were about 4 times the gold standard median SEs for EC (Tables 22, 23, and 24). For the other exposure models, we observed SEs 2-to-5 times those of the gold standard for all three pollutants. Similarly, the RMSEs were approximately 1.5-to-2 times larger using the LUR models than they would be if the true exposures were known (i.e., the gold standard) for NO_2 and $PM_{2.5}$, but were 3-to-4 times larger for EC.

Table 23. Estimated Bias, Median SE, and RMSE for Each Exposure Model of Indoor $PM_{2.5}$ Concentrations Given Different True Health Effect Estimates (β_X) and Corresponding ORs

	β_X (OR)								
	0.004 (1.05)			0.03 (1.50)			0.05 (2.00)		
Surrogate Model	Bias	Median SE	RMSE	Bias	Median SE	RMSE	Bias	Median SE	RMSE
Distance to nearest major road	0.051	0.030	3.66	−0.049	0.029	6.01	−0.054	0.031	15.8
Occupant density	−0.005	0.017	0.161	−0.003	0.020	0.270	0.002	0.020	0.581
LUR model	−0.001[a]	0.012	0.012	−0.003[a]	0.012	0.013	−0.007[a]	0.013	0.015
Gold standard[b]	0.000	0.007	0.007	0.000[a]	0.007	0.007	0.000	0.008	0.008

[a] Statistically significantly different from zero based on standard error of bias estimates ($P < 0.05$).

[b] Gold standard is defined as the simulated indoor concentrations at every home.

Table 24. Estimated Bias, Median SE, and RMSE for Each Exposure Model of Indoor EC Concentrations Given Different True Health Effect Estimates (β_X) and Corresponding ORs

	β_X (OR)								
	0.17 (1.05)			1.37 (1.50)			2.35 (2.00)		
Surrogate Model	Bias	Median SE	RMSE	Bias	Median SE	RMSE	Bias	Median SE	RMSE
Distance to nearest major road	1.28	1.35	52.1	−1.99[a]	1.45	64.5	−0.960	1.61	55.8
Distance to nearest designated truck route	−0.006	1.57	11.8	−1.30[a]	1.67	40.1	−0.930	1.82	38.4
LUR model	−0.008	1.36	1.37	−0.140[a]	1.43	1.45	−0.320[a]	1.54	1.56
Gold standard[b]	0.004	0.321	0.324	0.000	0.352	0.354	0.010	0.404	0.400

[a] Statistically significantly different from zero based on standard error of bias estimates ($P < 0.05$).

[b] Gold standard is defined as the simulated indoor concentrations at every home.

Table 25. Percentage of Significant Positive Results for Each Exposure Model of Indoor NO_2, $PM_{2.5}$, and EC Concentrations Given Different True ORs[a]

Surrogate Model	NO_2 (True OR)				$PM_{2.5}$ (True OR)				EC (True OR)			
	1.00	1.05	1.50	2.00	1.00	1.05	1.50	2.00	1.00	1.05	1.50	2.00
Distance to nearest major road	2.73	2.60	9.36	9.36	2.58	2.67	3.76	4.80	2.60	2.82	2.73	2.49
Unweighted traffic density within 50-m buffer	2.69	3.58	22.9	33.1								
Distance to nearest designated truck route									2.24	3.22	8.13	13.7
Gas stove usage	2.44	3.80	29.1	48.8								
Occupant density					2.22	3.84	35.7	63.8				
LUR model	2.36	5.51	92.5	98.9	2.69	4.40	59.7	89.8	2.60	3.29	14.1	27.2
Gold standard[b]	2.56	10.3	100	100	3.31	7.93	99.0	100	2.00	7.89	97.5	100

[a] Blank spaces indicate an exposure model not considered as a potential surrogate for the pollutant.

[b] Gold standard is defined as the simulated indoor concentrations at every home.

The inflation of SEs as well as the bias in the health-effect coefficients had a noticeable influence on the power to detect significant associations (Table 25). At a true OR of 1.05, even the gold standard exposure measurements only produced significant effects in 8% to 10% of realizations, with even lower percentages for the various surrogate models. The performance of the LUR model for NO_2 was favorable compared to that of the gold standard, with 93% significant positive results at an OR of 1.5, and 99% at an OR of 2.0. This could be attributed in part to the fact that this model had an R^2 value of 0.42 when all simulated data sets were pooled. Recall that the R^2 values reported in this section refer to the models predicting pollutant concentrations in the first year of life and are therefore different from the values from models predicting shorter-term average concentrations presented in Table 19. For NO_2, there was a substantial drop-off in the percent of significant positive results for all of the surrogate models, with the gas stove usage model yielding the best results (29% and 49% significance for ORs of 1.5 and 2.0, respectively).

The LUR model for $PM_{2.5}$ (R^2 value of 0.28 across all realizations) had 60% significant positive results at an OR of 1.5, and 90% at an OR of 2.0 (Table 25). The occupant-density model performed reasonably, with 64% significant positive results at a true OR of 2.0. The LUR model for EC had an R^2 value of 0.05 for annual average concentrations across all realizations. This resulted in a worse performance relative to $PM_{2.5}$ and NO_2, with only 27% significant positive results even at an OR of 2.0.

Table 25 also shows that the proportion of significant positive associations under a true OR of 1.00 was close to 2.5% in all cases, indicating reasonable SEs for all models, using a two-sided test with a significance level of 5%. Our SE calculations did not account for overdispersion in the logistic regression model, potentially resulting in inaccurate SEs (Carroll et al. 1995). However, the robustness of the naïve SEs can also be illustrated by the coverage probabilities, which were approximately 0.95 at all three true ORs (greater than 1) for the LUR models and gold standard, and at a true OR of 1.05 for the individual surrogates (results not shown). Thus, the naïve SEs may be somewhat inflated but are reasonable using the LUR models in our simulations.* The coverage probability was low, however, at the larger true ORs for the individual surrogate models.

When we considered some key sensitivity analyses, we found the results were not sensitive to the baseline risk chosen during the data generation process (results not shown). The reduction by half in the error incorporated when generating $C_{in\,year_i}$ and $C_{in_{wi}}$ caused a narrower distribution in $\hat{\beta}_X$; however the general conclusions remained the same (results not shown).

* The reasonable results with regard to the naïve SEs can be explained in part as follows. When either β_x or the measurement error variance is small, overdispersion does not dramatically inflate the SEs. For example, assuming that only 20% of the original variance in indoor NO_2 levels is explained by the model, the largest of our ORs (2.0) produces an overdispersion estimate of 1.17, or an inflation in the SE of only 8%. This helps to explain why the coverage probabilities in our simulations are reasonable for small ORs and for the models with more explanatory power. Given our analytic results and the fact that in our simulations the coverage probabilities for the previously developed LUR models are close to 0.95, we did not adjust the SEs in our analyses.

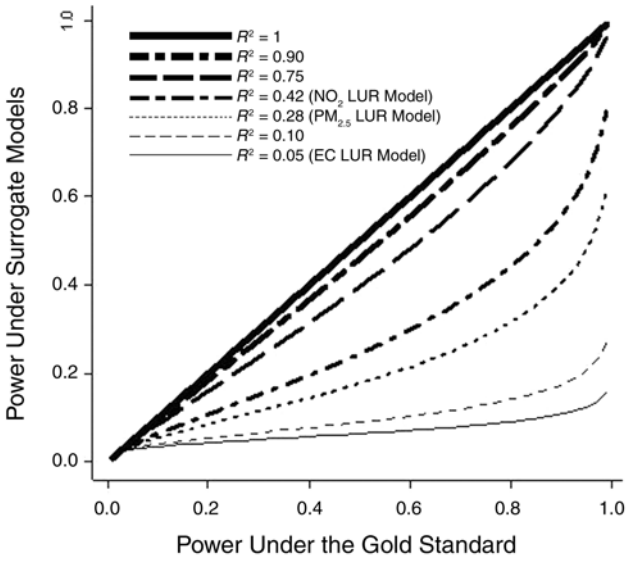

Figure 13. Power of a hypothetical validation study to detect positive associations under a surrogate model versus the gold standard, given different R^2 values.

To better understand the implications of validation study models with varying degrees of explanatory power, we plotted the approximate theoretical relationship between the power to detect a positive association using the gold standard and the power using surrogate exposure models (Figure 13). This relationship is based on the strength of association between the predicted exposures from surrogate models and true exposures (equation 15). We see that the empirical results reasonably follow the theoretical relationship, with the power based on a surrogate model increasing in a nonlinear fashion as a function of the power under the gold standard, and with low power for small R^2 values. More specifically, there was low power for EC ($R^2 = 0.05$ for the LUR model across realizations) under all OR scenarios and for $PM_{2.5}$ and NO_2 when the gold standard power was low (< 11% for true OR = 1.05). For $PM_{2.5}$ and NO_2, the gold standard empirical power was nearly 100% for ORs of 1.5 and 2.0, so the theoretical model indicated power on the order of 0.6, for $PM_{2.5}$ ($R^2 = 0.28$), and 0.8, for NO_2 ($R^2 = 0.42$), in general agreement with the simulation findings (Table 25).

DISCUSSION AND CONCLUSIONS

Our study aimed to provide insight about whether a combination of data from different sources (GIS-based covariates, a questionnaire, and central site monitoring), analyzed with a combination of LUR and factor analytic methods, could explain variability in indoor and outdoor concentrations of multiple air pollutants within urban residential areas. We were able to explain 52% to 76% of the variability in outdoor residential concentrations of EC, $PM_{2.5}$, and NO_2 with a combination of temporal and spatial terms, although more variability was explained for $PM_{2.5}$ (largely temporal) than for EC and NO_2. The combined factor analytic and LUR methods for outdoor concentrations helped to elucidate key sources, which largely conformed to expectations, supporting our initial factor interpretations. The explanatory power of the LUR models for outdoor factors was weaker than that of the LUR models for individual pollutants, as might be expected, but these models provided insight about spatially varying source contributions and site characteristics associated with that spatial variation.

Furthermore, as expected, indoor–outdoor relationships varied substantially across pollutants, which helped us better understand which were dominated by indoor sources and which were dominated by outdoor sources across urban settings. We were able to explain a majority of the variability in many individual indoor particle constituents using measured outdoor concentrations and a ventilation proxy. Our study indicates that characterizing indoor concentrations of ambient origin is viable, given our reasonably interpretable outdoor factors and our ability to explain variability in both outdoor concentrations and in effective penetration efficiencies. Further, our factor analyses of indoor concentrations were generally interpretable, if complicated by the simultaneous contributions of indoor and outdoor sources. However, our regression models predicting indoor concentrations or factors as a function of GIS covariates, questionnaire data, and meteorological data performed more poorly, with almost no ability to explain variability in factors attributed to indoor sources. That said, our epidemiologic simulations demonstrated that even a modest reduction in exposure misclassification associated with using indoor rather than outdoor concentrations could prove significant for epidemiologic purposes. We provide more detailed discussion, below, about each of the subanalyses, followed by a more general consideration of the implications of our study for future epidemiologic studies or exposure models.

LUR MODELS FOR OUTDOOR EC, $PM_{2.5}$, AND NO_2

We found significantly greater variability and stronger relationships with local traffic for EC and NO_2 than for $PM_{2.5}$. This was reflected in both covariate significance and the relative amount of variability explained by spatial rather than temporal terms. Our findings are consistent with prior literature (Martuzevicius et al. 2004; Wu et al. 2005), corroborating evidence that $PM_{2.5}$ patterns are

largely regional in nature for the eastern United States (Burton et al. 1996; Suh et al. 1997).

One implication of focusing on an urban residential setting is that traffic volumes may vary less there than across metropolitan areas. As illustrated in Figure 3, across the entire urban core, 100-m kernel-weighted traffic scores ranged from 0 to 3,305 vehicle-meters traveled per m^2 per day; at cohort homes, this measure ranged from 5.8 to 168 vehicle-meters traveled per m^2 per day. This lower range in scores observed at cohort homes is likely driven largely by the fact that relatively few homes are located by major highways. This observation may be important for exposure estimation; many models are derived from concentration data collected near major roads, which may inaccurately reflect traffic-concentration associations at the lower end.

We can also use our LUR models (Table 9) to gain insight about the approximate amounts of residential outdoor concentrations of $PM_{2.5}$, EC, and NO_2 that are attributable to local traffic. For $PM_{2.5}$, the mean total roadway length within 100 m of the home (1,110 m) accounted for a marginal contribution of 1.2 $\mu g/m^3$, or 9.7% of predicted $PM_{2.5}$. Applying our predictive models with mean values for all terms, the mean predicted concentration is 13.2 $\mu g/m^3$; an increase of one standard deviation in roadway length within 100 m of the home (371 m) increases concentrations to 13.9 $\mu g/m^3$. Population density, which likely captures some traffic-related influence, adds 1.1 $\mu g/m^3$ on average. When we use nontransformed $PM_{2.5}$ data, the predicted local traffic contribution is somewhat larger (2.6 $\mu g/m^3$).

For EC, the mean total roadway length within 200 m of the home (3,560 m) accounted for approximately 0.17 $\mu g/m^3$, or 36%, of predicted EC. Increasing roadway length by one standard deviation (1,156 m), with all other parameters at their mean values, increases predictions from 0.47 to 0.54 $\mu g/m^3$. Using nontransformed EC concentrations, predicted local traffic contributions are somewhat larger (0.39 $\mu g/m^3$). We observed a gradient of almost 1 $\mu g/m^3$ in EC across sampled homes (somewhat greater across individual measurements) before correcting for temporal variability, which is relatively small compared to the variability observed in European studies (approximately 10 $\mu g/m^3$), as expected given the greater use of diesel passenger vehicles in Europe (Hoek et al. 2002).

Modeled local traffic terms accounted for approximately 2.8 ppb, or 21% of modeled NO_2. The mean total roadway length within 50 m of the home (441 m), with mean values for other terms, predicts a concentration of 17.8 ppb. An increase of one standard deviation (179 m) increases concentrations to 18.9 ppb. The range of 50-m roadway lengths observed predicts a NO_2 range of 15.4 to 20.9 ppb. Population density, again, likely captures some local traffic effect and accounts for 4.4 ppb on average.

Interestingly, for all three pollutants modeled, total roadway length within varying buffer radii provided the greatest explanatory power, although sensitivity analyses and subsequent factor analysis LUR models revealed that diesel covariates, where available, also performed well for EC. Because actual traffic counts for smaller residential roads are generally sparse, length measures may provide more stable traffic indicators in residential areas than estimates based on traffic counts. Also, differential bias in traffic-count accuracy by roadway size is found in most available traffic databases; actual traffic counts are generally collected on a regular basis for highways and major roads, with rough estimates created for smaller residential roads.

In addition, we found that accurate modeling of outdoor concentrations near urban residences required some consideration of site characteristics, such as population density and obstructions, which explained significant variability in concentrations and at times altered traffic-concentration associations. Population density, significant for $PM_{2.5}$ and NO_2, may proxy either for other residential sources or for local traffic and may indicate higher per-mile emission rates from stop-and-go traffic in denser residential neighborhoods. We expected obstruction to modify $PM_{2.5}$ and EC, in keeping with recent findings suggesting that roadside barriers reduce $PM_{2.5}$ concentrations (Chen et al. 2007). Though we did not expect to find this effect for a gaseous pollutant such as NO_2, this result is supported by evidence that residential NO_2 concentrations differed significantly depending on whether the home faced onto the courtyard or street, after accounting for distance to road (Reungoat et al. 2005). Our findings may also be related in part to our passive sampling approach, as an obstruction could reduce the face velocity on the sampler. This effect should be limited, however, because we mounted inverted dishes over samplers in the field, to protect samplers and moderate face velocity.

CONSTRAINED FACTOR ANALYSIS AND LUR MODELS FOR OUTDOOR LATENT VARIABLES

Our approach was effective in deriving source-related factors from measures of particle constituents and gaseous pollutants collected across multiple sites, incorporating both spatial and temporal variability. The GIS-based LUR models for outdoor concentrations largely corroborated the hypothesized factor sources, reinforcing the validity of combining constrained factor analysis and LUR models. One benefit of our approach was that it enabled some reasonable distinction between $PM_{2.5}$ factors driven by temporal variability and those driven by spatial variability.

For example, Factor 1, interpreted as long-range transport, was predicted largely by central site monitoring data and season, which vary only over time; as such, this source factor would not vary in a long-term epidemiologic study in this urban area.

In contrast, Factors 2 and 5, both of which were hypothesized to capture aspects of local traffic, had no significant correlation with central site monitoring data in multivariate LUR models. Interestingly, high-density residential land use significantly predicted lower levels of Factor 2 (brake wear and local traffic) but higher levels of Factor 5 (road dust and resuspension). This could potentially be an indication of the differential effects of traffic characteristics, including fleet composition and vehicle speed. For example, prior studies of brake-wear-related exhaust components indicate that brake-pad wear can accumulate inside the brake housing unit and not be released until the vehicle reaches higher speeds, thereby forcing more air through the chamber and releasing encased dusts (Sternbeck et al. 2002; Schauer et al. 2006). If this is the case, then we might expect relatively weak spatial correlations between denser areas where brake dust is created (e.g., by stop-and-go traffic) and areas of more freely flowing traffic, where these constituents may ultimately be released into ambient air. Road dust would similarly have a complex relationship with vehicle speeds and traffic patterns, which are related to deposition and resuspension patterns. The complexities of road dust and brake-wear patterns are further reinforced by some of the differences between the factor loadings in our primary model and those in the factor analysis model derived from CAPs samples (Appendix C), although these may be attributable to a number of other differences in settings and the timing of sampling.

Our remaining two factors had very clear hypothesized sources given the factor loadings, which were well supported by the LUR models. For Factor 3 (diesel exhaust), the covariate for the fraction of diesel vehicles on the nearest major road was significant, but this covariate would not likely be available in many settings or in a large epidemiologic study, as it was based on our traffic-counter data. The next best univariate traffic-related predictor identified in sensitivity analyses, however, was distance to the nearest truck route. This term may therefore serve as a reasonable proxy in settings where more refined traffic data are not available. For Factor 4 (fuel oil combustion), the univariate and multivariate models emphasize strong seasonality, with higher concentrations during the heating season, corresponding to residential fuel oil usage. Although our residential fuel oil covariate did not retain significance in the multivariate model, population density served as a highly correlated proxy in this setting (indeed, fuel oil use was calculated using population density within 200 m of the home). In addition, the map of Factor 4 (Figure 5) clearly shows a high-concentration cluster close to industrial fuel oil sources, including ports and shipping channels along waterways near East Boston. We did not consider covariates such as proximity to the port in our regression models, as, in this case, such covariates could proxy for neighborhood, thus over-modeling our data. Spatial patterns for this factor as well as other terms could clearly be utilized in the future to develop more refined models.

We can gain further insight by comparing our LUR models of outdoor factors (Tables 9 and 11) with those developed for $PM_{2.5}$, EC, and NO_2 (Table 9). Residential outdoor total $PM_{2.5}$ was predicted by central site $PM_{2.5}$ with little contribution from traffic and local sources, similar to the Factor 1 model presented here, reinforcing the interpretation of Factor 1 as long-range transport. In contrast, residential outdoor EC was poorly predicted by central site EC, possibly owing to intraurban variability in EC sources and differences between EC measurement techniques used in study homes (reflectance) and at the central site (aethalometer). Similarly, central site EC did not predict Factor 3 (diesel exhaust). NO_2 had modest loadings across our five factors (Figure 4) and somewhat low correlations with particle constituents, but the factor with the greatest NO_2 loading (Factor 2, brake wear and local traffic) was associated with roadway length (Table 11), which also served as a significant predictor in the NO_2 regression model, albeit with a slightly different buffer size (Table 9). In general, the predictive power of our regression models for factor analysis results is slightly lower than for the individual pollutants; this observation suggests that our factors are not single well-defined sources, points to the complexity of collecting and analyzing measurements over space and time, and relates to a more restrictive regression modeling approach for the factor analysis outputs as well as uncertainties associated with factor analysis itself.

LUR MODELS FOR INDOOR EC, $PM_{2.5}$, AND NO_2

The predictive power of the LUR models for indoor concentrations of EC, $PM_{2.5}$, and NO_2 was somewhat smaller than the predictive power of the LUR models for outdoor concentrations. This was expected given the strong effects of indoor sources and modification by ventilation as well as the limitations of the covariates associated with these factors (e.g., the dependence on binary variables and the nonspecificity of covariates such as occupant density). Table 19 shows that the three pollutants have different patterns and predictors — indoor NO_2 is predicted by local traffic, indoor sources, and ventilation; indoor $PM_{2.5}$ by indoor sources and ventilation but not by local traffic; and indoor EC by local traffic and ventilation but not by indoor sources.

Looking at each pollutant in detail, we see that outdoor NO_2 was predicted by central site data, total roadway length within 50 m of the home, presence of obstructions between the home and roadway, season, and population density (Table 9). Our more restrictive model-building approach for the indoor environment did not allow all of these terms to be introduced, but there was concordance in the spatial coverage of the traffic variables, as traffic density within 50 m of the home was the strongest predictor of indoor NO_2. Given the range of cumulative unweighted traffic density within 50 m of the home (from 4.1 to 198 vehicle-meters/m^2/day), the indoor LUR model indicates traffic contributions from 0.3 to 14 ppb for homes with open windows, with no significant contribution to homes without open windows. As described earlier in this report, the range of roadway length values implies traffic contributions of approximately 5.5 ppb to outdoor concentrations, with population density (a potential proxy for traffic) contributing another 4.4 ppb, indicating general concordance between the indoor and outdoor models. For comparison, gas stove usage contributed 7 ppb, on average, to indoor NO_2 levels, comparable to observations from previous studies (Lee et al. 1998; Levy et al. 1998). Thus, our model would imply that local traffic is a larger contributor to indoor NO_2 where traffic density is high and windows are opened, whereas indoor sources are a larger contributor when traffic density is low or windows are closed.

For EC, roadway length was significant in multivariate outdoor-concentration LUR models, but covariates representing diesel traffic produced larger model R^2 values in sensitivity analyses. This finding is supported by our indoor LUR model for EC, which found that distance from a designated truck route and other diesel-related variables, such as diesel fraction and trucks per day, had the highest posterior probabilities (Table 18). The contribution of local traffic to indoor EC for homes with open windows was approximately 0.2 μg/m^3 (similar to the estimated contribution from the outdoor LUR model) and was insignificant for homes with closed windows. For both the indoor and outdoor LUR models (Tables 17 and 9, respectively), the R^2 values for EC were lower than for $PM_{2.5}$ or NO_2, potentially reflecting the different measurement methods employed at the central site and the residential monitors, greater spatial variability in EC that is difficult to explain with available traffic covariates (which only weakly capture diesel effects), and, for the indoor models, additional indoor sources not captured by questionnaire data.

Finally, for $PM_{2.5}$, both the indoor and outdoor LUR models (Tables 17 and 9, respectively) emphasize that most variability in outdoor concentrations is explained by temporal factors that are captured well by central site monitoring data. Ambient concentrations contributed an average of 15 μg/m^3 to indoor $PM_{2.5}$ for homes with open windows, and 10 μg/m^3 for homes where windows were closed. Additionally, cooking for more than an hour per day contributed 6.2 μg/m^3, and average occupant density contributed 6.5 μg/m^3. The observed effect of cooking is comparable to results from prior studies (Ozkaynak et al. 1994; Brunekreef et al. 2005). Occupant density is likely a proxy for multiple factors, including resuspension activities. Resuspension has not been a substantial contributor in previous studies, although the smaller volumes and greater crowding of our study homes may increase its relative source strength.

CONSTRAINED FACTOR ANALYSIS AND LUR MODELS FOR INDOOR LATENT VARIABLES

As with the outdoor-concentration factor analysis, the indoor-concentration factor analysis was generally effective in deriving factors that were physically interpretable and logical. In the six-factor model derived from measured indoor concentrations, two of the factors (long-range transport, fuel oil/diesel combustion) represented outdoor sources similar to those from the outdoor-concentration factor analysis model, both in terms of their loadings and interpretations. For those factors hypothesized to have indoor sources (e.g., road dust and resuspension, indoor combustion, indoor smoking, and indoor cleaning), these interpretations were corroborated in part by the findings from the factor analyses on residuals from indoor–outdoor regression models. We can therefore conclude that factor analysis methods can be successfully utilized in the residential indoor environment, and insights from indoor–outdoor regression models may corroborate factor interpretations.

The application of LUR-style modeling techniques to the indoor factors, however, was less successful. Given that our LUR- and questionnaire-based regression models for individual indoor pollutants (Table 19) could only explain a small portion of the variance in their concentrations, limited success in this model-building step was not unexpected. The factors driven by outdoor sources were reasonably interpretable, with some univariate associations further supporting our factor interpretations. Indoor factors were poorly predicted, which was attributable in part to limitations with our questionnaire data and in part to inherent complexities with indoor sources. For example, Factor 4 (indoor combustion) was highly skewed, with six observations (clustered among three homes) showing significantly higher factor scores than other observations. These observations all had Ce and La I/O ratios exceeding 10, with K generally in excess of 2, clearly indicative of indoor sources. While our questionnaire data were not successful in explaining the variability in this factor, the factor analyses clearly indicated an indoor source present

at a subset of homes. Similarly, Factor 5 was hypothesized to capture indoor smoking, but few participants reported smoking during the sampling period (four participants reported 1-to-4 cigarettes per day inside the home, with the rest reporting no smoking). Our inability to explain variability in Factor 5 could be related in part to reporting bias for questions about smoking behaviors (given that we were interviewing pregnant women or those with newborns) as well as the highly skewed distribution of the factor scores.

In general, our ability to use LUR and multivariate modeling to explain variability in indoor factors was likely limited by multiple issues, including the small sample size, the skewed distributions anticipated when a significant indoor source is present in a small number of homes, and the complex nature of ventilation-modified and correlated indoor-source activities.

IMPLICATIONS OF EXPOSURE MISCLASSIFICATION

For all pollutants, using LUR models developed from our air pollution substudy (considered here as a validation study) resulted in better estimates of the true health effect coefficients, with relatively small estimated biases and lower RMSEs compared to exposure models derived from individual exposure proxies. Our simulations help to confirm that individual measures of traffic, often used as exposure indicators in epidemiologic studies (Garshick et al. 2003; Heinrich et al. 2005; Ryan et al. 2005), will not perform well if uncorrelated or poorly correlated with the exposure of interest (and may have reduced power, even given positive correlations, if other factors are not considered). While this point is not surprising based on first principles and other simulation results (Freedman et al. 2008; Szpiro et al. 2008; Kim et al. 2009), our results demonstrate the quantitative implications of surrogate selection with varying degrees of measurement error and emphasize the importance of using validation studies to select the most appropriate traffic indicator or exposure model (Gauderman et al. 2005).

Perhaps most importantly in the context of our analysis, these simulations provide reassurance that regression models with fairly modest R^2 values (approximately 0.3 to 0.4) would substantially improve the ability to detect significant associations compared to models using individual exposure surrogates in the absence of validation studies. This is indicated in equation (15) and Figure 13, which show the increase in power with increasing R^2 values that would be expected with improved exposure modeling. Not only is power improved, but because of the nonlinearities articulated in Appendix G (Derivation of the Power Expression for Exposure Misclassification Analysis, available on the HEI Web site), this power is improved substantially, especially for larger studies with higher ORs. However, our simulation findings for EC suggest that a minimal performance threshold for regression models is necessary to yield significant associations consistent with our analytic calculations, presented in Figure 13 and derived in Appendix G.

Even for those exposure models derived from validation data, some bias still occurs. This bias and the bias seen for the individual exposure proxies is not surprising in light of recent analytic and simulation results in the statistical-measurement-error literature on the complexity of the effects of exposure-measurement error. The regression structure induces Berkson-like error in the predicted exposures because the process of estimating exposures using models functions as regression calibration in the measurement error context (Szpiro et al. 2008; Gryparis et al. 2009). This component of the exposure error should increase uncertainty and decrease power but not induce bias when the health model is a linear regression. However, in logistic regression, limited bias does occur with Berkson error (Carroll et al. 2006). More importantly, imprecision in the exposure-model parameter estimates can induce bias, particularly with small exposure data sets (Szpiro et al. 2008), as is the case in our setting. Note also that Freedman and colleagues (2008) found bias away from the null in simulations where the measurement error was large.

Additionally, because of outliers, the mean bias calculation is not very stable for the exposure models derived for individual surrogates. The weak surrogate–true exposure relationships lead to values of γ_2 (equation 13) near zero, producing little variation in the predicted exposures. This in turn caused numerical instability in the health effect estimates, producing some extreme values and explaining the statistical insignificance of some large mean bias estimates.

LIMITATIONS

There were multiple limitations within our individual analyses as well as in the study as a whole. Our relatively small sample size was an obvious limitation throughout the model-building processes. While measuring and analyzing numerous pollutants both indoors and outdoors at residences provided unique and important data, it also implied that a larger sample size would have been cost-prohibitive and that sampling at numerous sites simultaneously would have been logistically challenging. Our samples incorporated both spatial and temporal variability, and many of our analytic efforts were devoted to attempts to control for time-varying factors to better understand spatial and home-specific contributors to variability. Most prior LUR studies have taken measurements — generally unobstructed roadside measurements — at numerous sites simultaneously, obviating the need to control for temporal factors, but such

an approach is generally infeasible for studies with equipment-intensive, multipollutant-sampling designs, especially given the need to interact with residents throughout the sampling process. That said, our study did include more sites than have been incorporated in $PM_{2.5}$ factor analyses and source-apportionment studies to date, and both our individual-pollutant and our outdoor factor analysis regression models were able to separate to some extent pollutants or factors driven by long-range transport and temporal terms from pollutants or factors driven by spatial terms. Moreover, our effective sample size was somewhat greater than the number of samples collected, given the number of pollutants characterized in our analysis.

There were also limitations in many of the covariates used in the LUR models. The GIS-based traffic covariates can be highly correlated with one another and proxy for a variety of sources, and it is unclear a priori which covariates would be most appropriate within a given geographic area. In our study, length of roadway within various buffers was often a more robust predictor of pollutant concentrations than covariates accounting for traffic counts, but this was in part due to the degree of missing data for traffic counts as well as our focus on dense urban areas. The length of road segments within a defined buffer is a weak proxy of emissions from those road segments, and it would have a different interpretation in a dense urban neighborhood with diesel bus and truck traffic than in a suburban area with mostly smaller roads but possibly near a major highway. Other geographic areas may also have more robust local traffic counts, which could increase the predictive power of covariates calculated using traffic counts. More generally, the available indicators do not necessarily capture the characteristics of traffic that are most relevant to concentrations of different pollutants. For example, dense stop-and-go traffic may create more emissions per vehicle mile, and total traffic counts fail to capture such realities. In principle, GIS-based traffic covariates that include prevailing winds are more physically interpretable and should theoretically better capture the influence of local traffic. However, these terms were not always more significant in our analyses, although this may have been due to our consideration of only two candidate traffic variables with meteorological weighting, and we encourage the further development and exploration of GIS covariates that combine source strength and meteorology.

We used data from a central site monitor in Roxbury to capture temporal trends and help isolate spatial contributions to variability. However, this monitor may have been influenced by local traffic as well as by the same meteorological factors that can influence our measured concentrations at individual homes, complicating the interpretation of this term. The central site monitor also used different measurement methods for EC than were utilized in our sampling, potentially explaining in part the weaker performance for EC in both indoor and outdoor models. However, the monitor used in our primary analyses is at the center of our relatively small monitoring region and is well correlated with other ambient monitors in and around Boston, and our findings were not improved by using other available monitors.

The indoor-source terms in our models were developed from questionnaires that are surrogates for the source emissions rate, have limited resolution, and may be subject to recall bias or other errors. Some of the indoor-source covariates that significantly predicted indoor concentrations, such as number of occupants or number of occupants per room, are difficult to interpret and may represent a variety of occupant activities. However, most of these limitations are inherent in developing exposure estimates based on publicly available or questionnaire data and emphasize the value of more detailed time-activity and source-utilization data. Similarly, the open-windows variable was the best available predictor of the I/O sulfur ratio, but it is a crude proxy for that ratio, which itself is a proxy of the infiltration factor. Measuring AERs in each home clearly would have reduced the uncertainty in this term, but that was logistically impractical; it would have required being able to collect such measurements for all participants in a cohort study over multiple seasons, which would have been unlikely. Given these requirements, a simpler proxy based on questionnaire or home-characteristics data would be necessary even if measured AERs were available within a validation study. Our results suggest that window-opening behavior coupled with season is physically interpretable and reasonably predictive.

Some additional covariates in our models, such as obstruction between the home and nearest major road, would not ordinarily be available or interpretable for all members of a cohort, complicating interpretation of our models. These terms can be considered as correction factors for the restrictions associated with residential monitoring — samplers often need to be set up behind the buildings, wherever power sources are available, on a porch where smoking or grilling has also occurred, or at some other nonideal or nonrepresentative location. These parameters may not be appropriate for extrapolation, as they may not reflect mean concentrations near the home, but are important for correctly interpreting residential data. This complication reinforces the limitations discussed previously about residential sampling in relation to near-roadway sampling. Capturing variability in residential exposures requires accounting for numerous complexities

not present in other study designs and limits the predictive power of our models. On the other hand, our models arguably better reflect some of the real-world complexities associated with personal exposures.

There were some limitations in our exposure misclassification analysis as well. Our analytical framework presumed that individual pollutants are the causative agents in question, but most surrogates do not uniquely identify a causal pollutant. For example, a traffic indicator represents a complex mixture of pollution elevated near roadways (Jerrett et al. 2005) and may also be a marker for other factors such as socioeconomic status or noise. It may be a more appropriate exposure surrogate for simultaneous exposure to multiple factors associated with traffic. Similarly, some indoor-source terms such as occupant density may not only represent resuspension activities, but may also be associated with socioeconomic status, housing type, or related occupant-activity patterns. Our analysis also does not include a true validation study, in the sense that the gold standard does not reflect measured exposures across the entire time window that may be relevant to the development of health effects. However, rarely in epidemiology is the gold standard a perfect measurement of true exposure (Wacholder et al. 1993; Brenner 1996; Spiegelman et al. 1997). Direct measurements are logistically infeasible for a cohort study considering long-term exposures, so we view our validation study (our short-term monitoring across multiple seasons with models that explicitly address seasonally varying factors) as a reasonable attempt to represent measured exposure. In addition, carrying out the estimation step for all 4,500 simulated data sets addresses the effect of uncertainty in estimating model coefficients, but does not reflect uncertainty deriving from the original model-selection process. More generally, the aim of this simulation study was not to determine the merits of our LUR models per se, but to provide a quantitative framework within which various exposure models could be formally evaluated.

Finally, while many aspects of our methods and analysis can be generalized to other settings, we do not recommend that our specific LUR models be directly applied elsewhere without additional data collection and analysis. While some studies have demonstrated transferability of LUR models across cities (Briggs et al. 2000; Poplawski et al. 2009), this was only for NO_2, given additional local sampling and with demonstrated between-city concordance in input data. Moreover, other studies (Jerrett et al. 2005) have shown poor transferability of LUR models. At a minimum, use of our LUR models would be recommended only with local sampling at a subset of sites to validate or modify our models. In terms of the exposure misclassification analysis, our results should not be used to directly reinterpret previously published epidemiologic findings (e.g., to draw inferences about studies using distance to the nearest major road as an exposure surrogate). Multiple traffic covariates did not predict measured concentrations in our study, but terms such as distance to the nearest major road may be more robust predictors of exposures in settings with many large highways and limited surface traffic, as well as for pollutants with limited indoor sources or high effective penetration efficiencies or with both. However, our study offers a methodology that is generalizable and can be applied to other studies and data sets.

GENERALIZABLE INSIGHTS AND LESSONS LEARNED

Our study focused on exposure characterization for an urban cohort, a topic of numerous previous investigations, but it adopted somewhat different strategies than some studies, with respect to sampling design (i.e., measurements in indoor and outdoor residential settings), scope (i.e., consideration of a large number of pollutants at a smaller number of sites), and analytical approaches (i.e., use of both factor analyses and LUR). It is therefore valuable to consider which of these strategies would be recommended for future investigations, and which would not be recommended in light of the resources deployed relative to the information gained.

First, there are significant tradeoffs associated with characterizing the indoor environment. Monitoring indoors at all sites involved substantial resources and reduced the number of sites at which we could take measurements, reducing the statistical power of outdoor LUR models. An alternative strategy would involve developing spatiotemporal models for outdoor concentrations and using a smaller number of indoor–outdoor measurements to develop models of effective penetration efficiency, as is being done in the MESA Air study (Cohen et al. 2009). This approach has some appeal, and as our study has shown, would likely yield more-predictive models. However, capturing only indoor air pollution of ambient origin will add exposure misclassification to resulting epidemiologic analyses. Whether it makes sense to pursue characterization of indoor sources depends on the particle constituents of interest given the health outcome of concern as well as the degree of exposure misclassification induced by a weaker model that includes a measure more closely related to personal exposure versus a stronger model that relies on a measure more distantly related to personal exposure. Our models would suggest that characterizing indoor air pollution of ambient origin is a more viable approach in the near term, but we encourage researchers to continue to refine questionnaires, explore tax assessor databases, and to use other strategies to better capture indoor sources.

Our study reflects a multipollutant exposure-assessment strategy that is potentially responsive to the substantial interest in understanding the relative toxicity of various particle constituents, as well as in developing a hierarchy of sources for air pollution regulations. In general, studies that are able to measure or model pollutants beyond $PM_{2.5}$ and studies that use factor analysis and related techniques to determine source contributions are clearly informative. In particular, it is valuable to understand the combination of pollutants that may be correlated with either a single pollutant or a GIS-based traffic variable, to inform the interpretation of epidemiologic investigations. We recommend that any study collecting particle filters in the context of LUR modeling conduct some form of elemental or chemical analysis followed by factor analysis, to help advance the understanding of pollutants and sources contributing to observed health outcomes. While the cost of such analysis may not be feasible in some settings, the value of the information gained is likely to be substantial relative to the marginal costs incurred. Finally, we encourage investigators pursuing this direction to use techniques for joint factor analysis–LUR modeling, which can help in source identification and can help avoid statistical issues associated with modeling numerous individual pollutants.

CONCLUSIONS

Our study offers some insight into intraurban variability in residential exposures to multiple air pollutants and demonstrates the utility of various methods to evaluate these concentrations. We demonstrated, within an urban area, significant outdoor spatial variability in NO_2 and in multiple particle constituents. LUR techniques combined with constrained factor analyses for outdoor concentrations helped to disentangle the contributions of local sources from long-range transport and other sources of temporal variability. These two techniques have separately been successful in characterizing different temporal and spatial aspects of $PM_{2.5}$, but our analysis leveraged insights from both approaches, using the factor analysis to provide interpretable source factors and GIS-based LUR modeling to corroborate source interpretations.

Our analytic methods would allow exposures to a defined source category to be investigated in epidemiologic studies, potentially moving beyond roadway proximity measures to examine the differential effects of diesel exhaust, brake wear and exhaust, and resuspended road dust, thus reducing exposure misclassification and ultimately facilitating the identification of constituents that are causally related to health effects and enabling more effective interventions. More generally, our analytic methods can be used in settings in which multipollutant sampling is desired but only a small number of homes can be sampled simultaneously, resulting in a need to distinguish temporal from spatial effects for exposure modeling. Our findings suggest that the development of sampling designs with more temporally structured sampling intervals, especially designs that maximize temporal overlap (e.g., systematically staggered sampling with a predictable amount of temporal overlap at multiple homes), might allow for improved distinction of spatial and temporal variability and improved characterization of intraseason variability. Such theoretically optimal sampling designs should be considered, recognizing that residential sampling includes many complications, including visit rescheduling. Similarly, site selection can be optimized given initial characterization of site and traffic characteristics, especially of local sources, site characteristics, and obstructions that may influence concentrations.

Our outdoor findings also show that multiple categories of traffic contributions can be isolated and predicted by different GIS covariates, emphasizing the value of a multipollutant approach and the need to interpret individual traffic proxies with caution. Within our data set, measures of roadway length or traffic density were generally more predictive of exposures than land-use type, which may be more useful for studies of large geographic domains for which land-use and other GIS variables display greater variation. Our findings for outdoor concentrations, with LUR models that include terms for site and meteorological characteristics, also emphasize the importance of incorporating into study models small-scale spatial and temporal predictors to accurately capture exposure variability in urban residential settings.

In our analysis of the indoor residential environment, we demonstrated that indoor–outdoor relationships varied substantially among particle constituents and that information from public databases (e.g., central site monitoring data, GIS-based traffic data) combined with focused questionnaire data could predict indoor concentrations for a subset of key pollutants. Factor analysis methods applied to the indoor environment, coupled with indoor–outdoor modeling, helped to separate indoor sources from outdoor sources to some extent. Although explaining variability in indoor-source contributions was challenging, our analytical framework provides direction for future studies characterizing indoor exposure sources and source-activity patterns. Because of the differential indoor–outdoor relationships across constituents and the likelihood that indoor concentrations are a more meaningful proxy of personal exposures than are outdoor concentrations, formal consideration of the implications of measurement error on epidemiologic study results is necessary. Our analysis provided a technique by

which these implications could be quantified. This analysis also underlined the fact that, because both indoor and outdoor sources importantly influence concentrations of multiple air pollutants, individual exposure proxies are likely to be less effective predictors of true individual exposures than are multiple regression models developed through formal validation studies.

ACKNOWLEDGMENTS

This study was funded by the Health Effects Institute (4727-RFA04-5/05-1), with additional support provided by the National Institutes of Health (R03 ES013988). The overall ACCESS project, which supported participant recruitment and sample collection, was supported by the National Institutes of Health (R01 ES10932, U01 HL072494, and R01 HL080674). Dr. Baxter received additional support from PHS 5-T42-CCT1229661-02 and PHS 1-T42-OH008416-01, and Dr. Clougherty received additional support from an American Lung Association Lung Health dissertation grant. Funding for the collection of CAPs was provided by the Harvard NIEHS Center for Environmental Health (P30ES00002) 2007 pilot grants program and the Harvard/EPA PM Center. The authors wish to thank Prashant Dilwali, Robin Dodson, Shakira Franco, Lu-wei Lee, Rebecca Schildkret, and Leonard Zwack for their sampling assistance; Monique Perron for both her sampling and laboratory assistance; Jonathan Natkin for his assistance in constructing sampling boxes; and Steven Melly for his GIS support. We thank Dr. Rosalind Wright of the Channing Laboratory, the principal investigator of the ACCESS project, as well as all of the technicians associated with the ACCESS project. For their assistance with collecting temporal CAPs data, we wish to thank Joy Lawrence, Christina Rossi, Mark Long, Edgar Diaz, John Godleski, and Petros Koutrakis. Helpful feedback and contributions to manuscripts were provided by Drs. Brent Coull, Francine Laden, John Spengler, and Helen Suh. We greatly appreciate the hospitality and participation of all study participants.

REFERENCES

Allen G. 2006. The aethalometer™ for BC: What it is, what it does, where it's going. In: 2006 National Air Monitoring Conference, November 6, 2006, Las Vegas, Nevada. Available from *http://epa.gov/ttn/amtic/files/ambient/2006conference/allenaethalometer.pdf*.

Arain MA, Blair R, Finkelstein N, Brook JR, Sahsuvaroglu T, Beckerman B, Zhang L, Jerrett M. 2007. The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. Atmos Environ 41:3453–3464.

Beelen R, Hoek G, Fischer P, van den Brandt PA, Brunekreef B. 2007. Estimated long-term outdoor air pollution concentrations in a cohort study. Atmos Environ 41:1343–1358.

Belanger K, Beckett WS, Triche E, Bracken MB, Holford TR, Ren P, McSharry JE, Gold DR, Platts-Mills TAE, Leaderer BP. 2003. Symptoms of wheeze and persistent cough in the first year of life: Associations with indoor allergens, air contaminants, and maternal history of asthma. Am J Epidemiol 158:195–202.

Belisle CJP. 1992. Convergence theorems for a class of simulated annealing algorithms on Rd. J Appl Prob 29:885–895.

Brauer M, Hoek G, van Vliet P, Meliefste K, Fischer P, Wijga A, Koopman L, Neijiens H, Gerritsen J, Kerkhof M, Heinrich J, Bellander T, Brunekreef B. 2002. Air pollution from traffic and the development of respiratory infections and asthmatic and allergic symptoms in children. Am J Respir Crit Care Med 166:1092–1098.

Brauer M, Hoek G, van Vliet P, Meliefste K, Fischer P, Gehring U, Heinrich J, Cyrys J, Bellander T, Lewne M, Brunekreef B. 2003. Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. Epidemiology 14:228–239.

Breiman L. 2001. Random forests. Machine Learning 45:5–32.

Brenner H. 1996. Correcting for exposure misclassification using an alloyed gold standard. Epidemiology 7:406–410.

Briggs DJ, de Hoogh C, Gulliver J, Wills J, Elliott P, Kingham S, Smallbone K. 2000. A regression-based method for mapping traffic-related air pollution: Application and testing in four contrasting urban environments. Sci Total Environ 253:151–167.

Brown KW. 2006. Characterization of particulate and gaseous exposure of sensitive populations living in Baltimore and Boston [dissertation]. Department of Environmental Health, Harvard School of Public Health, Boston, MA.

Bruinen de Bruin YB, Koistinen K, Yli-Tuomi T, Kephalopoulos S, Jantunen M. 2006. Source apportionment techniques and marker substances available for identification of personal exposure, indoor and outdoor sources of chemicals.

EUR 22349 EN. European Commission - Joint Research Centre Institute for Health and Consumer Protection, Physical and Chemical Exposure Unit. European Communities, Luxembourg.

Brunekreef B, Janssen NA, Hartog J, Harssema H, Knape M, van Vliet P. 1997. Air pollution from truck traffic and lung function in children living near motorways. Epidemiology 8:298–303.

Brunekreef B, Janssen N, de Hartog J, Oldenwening M, Meliefste K, Hoek G, Lanki T, Timonen K, Vallius M, Pekkanen J, Van Grieken R. 2005. Personal, Indoor, and Outdoor Exposures to $PM_{2.5}$ and Its Components for Groups of Cardiovascular Patients in Amsterdam and Helsinki. Research Report 127. Health Effects Institute, Boston, MA.

Burton RM, Suh HH, Koutrakis P. 1996. Spatial variation in particulate concentrations within metropolitan Philadelphia. Environ Sci Technol 30:400–407.

Carroll R, Ruppert D, Stefanski L. 1995. Measurement Error in Nonlinear Models. 1st edition. Chapman & Hall, London, UK.

Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. 2006. Measurement Error in Nonlinear Models: A Modern Perspective. 2nd edition. Chapman & Hall, London, UK.

Chan Y-C, Cohen DD, Hawas O, Stelcer E, Simpson R, Denison L, Wong N, Hodge M, Comino E, Carswell S. 2008. Apportionment of sources of fine and coarse particles in four major Australian cities by positive matrix factorization. Atmos Environ 42:374–389.

Chen F, Dimmick F, Williams R, Baldauf R. 2007. Particulate matter concentration measurements from near road studies (nrs) in Raleigh, North Carolina. In: 17th Annual Conference of the International Society for Exposure Analysis, October 14–18, 2007, Durham, NC.

Chipman H, George EI, McCulloch RE, Lahiri P (eds). 2001. Model Selection. Institute of Mathematical Statistics Lecture Notes, pp. 65–116. Institute of Mathematical Statistics, Beachwood, OH.

Chow J, Watson J. 1998. Guideline on speciated particulate monitoring. Third draft report. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC. Available from *http://epa.gov/ttnamti1/files/ambient/pm25/spec/drispec.pdf*.

Clayton CA, Perritt RL, Pellizzari ED, Thomas KW, Whitmore RW, Wallace LA, Ozkaynak H, Spengler JD. 1993. Particle total exposure-assessment methodology (PTEAM) study: Distributions of aerosol and elemental concentrations in personal, indoor, and outdoor air samples in a southern California community. J Expo Anal Environ Epidemiol 3:227–250.

Cohen MA, Adar SD, Allen RW, Avol E, Curl CL, Gould T, Hardie D, Ho A, Kinney P, Larson TV, Sampson P, Sheppard L, Stukovsky KD, Swan SS, Liu L-JS, Kaufman JD. 2009. Approach to estimating participant pollutant exposures in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Environ Sci Technol 43: 4687–4693.

Cyrys J, Heinrich J, Hoek G, Meliefste K, Lewne M, Gehring U, Bellander T, Fischer P, van Vliet P, Brauer M, Wichmann H, Brunekreef B. 2003. Comparison between different traffic-related particle indicators: Elemental carbon (EC), $PM_{2.5}$ mass, and absorbance. J Expo Anal Environ Epidemiol 13:134–143.

Duhme H, Weiland SK, Keil U, Kraemer B, Schmid M, Stender M, Chambless L. 1996. The association between self-reported symptoms of asthma and allergic rhinitis and self-reported traffic density on street of residence in adolescents. Epidemiology 7:578–582.

Dzubay T, Stevens R, Gordon G, Olmez I, Sheffield A, Courtney W. 1988. A composite receptor method applied to Philadelphia aerosol. Environ Sci Technol 22:46–52.

Edwards J, Walters S, Griffiths RK. 1994. Hospital admissions for asthma in preschool children: Relationship to major roads in Birmingham, United Kingdom. Arch Environ Health 49:223–227.

Finkelstein MM, Jerrett M, Sears MR. 2004. Traffic air pollution and mortality rate advancement periods. Am J Epidemiol 160:173–177.

Freedman LS, Midthune D, Carroll RJ, Kipnis V. 2008. A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. Stat Med 27:5195–5216.

Garshick E, Laden F, Hart JE, Caron A. 2003. Residence near a major road and respiratory symptoms in U.S. veterans. Epidemiology 14:728–736.

Gauderman WJ, Avol E, Lurmann F, Kuenzli N, Gilliland F, Peters J, McConnell R. 2005. Childhood asthma and exposure to traffic and nitrogen dioxide. Epidemiology 16:737–743.

George EI, McCulloch RE. 1997. Approaches for Bayesian variable selection. Statistica Sinica 7:339–374.

Gilbert NL, Woodhouse S, Stieb DM, Brook JR. 2003. Ambient nitrogen dioxide and distance from a major highway. Sci Total Environ 312:43–46.

Gilbert NL, Goldberg MS, Beckerman B, Brook JR, Jerrett M. 2005. Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model. J Air Waste Manag Assoc 55:1059–1063.

Gilbert NL, Goldberg MS, Brook JR, Jerrett M. 2007. The influence of highway traffic on ambient nitrogen dioxide concentrations beyond the immediate vicinity of highways. Atmos Environ 41:2670–2673.

Gordian ME, Haneuse S, Wakefield J. 2006. An investigation of the association between traffic exposure and the diagnosis of asthma in children. J Expo Sci Environ Epidemiol 16:49–55.

Gryparis A, Paciorek CJ, Zeka A, Schwartz J, Coull BA. 2009. Measurement error caused by spatial misalignment in environmental epidemiology. Biostatistics 10:258–274.

Hammond DM, Dvonch JT, Keeler GJ, Parker EA, Kamal AS, Barres JA, Yip FY, Brakefield-Caldwell W. 2008. Sources of ambient fine particulate matter at two community sites in Detroit, Michigan. Atmos Environ 42:720–731.

Hastie R, Tibshirani R, Friedman J. 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York, NY.

Heinrich J, Topp R, Gehring U, Thefeld W. 2005. Traffic at residential address, respiratory health, and atopy in adults: The national German health survey. Environ Res 98:240–249.

Henderson SB, Beckerman B, Jerrett M, Brauer M. 2007. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. Environ Sci Technol 41:2422–2428.

Highsmith VR, Hardy RJ, Costa DL, Germani MS. 1992. Physical and chemical characterization of indoor aerosols resulting from the use of tap water in portable home humidifiers. Environ Sci Technol 26:673–680.

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

Hopke PK, Ramadan Z, Paatero P, Norris GA, Landis MS, Williams RW, Lewis CW. 2003. Receptor modeling of ambient and personal exposure samples: 1998 Baltimore Particulate Matter Epidemiology-Exposure Study. Atmos Environ 37:3289–3302.

Houseman EA, Coull BA, Ryan LM. 2006. A functional-based distribution diagnostic for a linear model with correlated outcomes. Biometrika 93:911–926.

Hoyer P. 2004. Non-negative matrix factorization with sparseness constraints. J Machine Learning Res 5:1457–1469.

Iijima A, Sato K, Yano K, Kato M, Kozawa K, Furuta N. 2008. Emission factor for antimony in brake abrasion dusts as one of the major atmospheric antimony sources. Environ Sci Technol 42:2937–2942.

Ito K, Xue N, Thurston G. 2004. Spatial variation of $PM_{2.5}$ chemical species and source-apportioned mass concentrations in New York City. Atmos Environ 38:5269–5282.

Jerrett M, Arain MA, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, Morrison J, Giovis C. 2005. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 15:185–204.

Jerrett M, Arain MA, Kanaroglou P, Beckerman B, Crouse D, Gilbert NL, Brook JR, Finkelstein N, Finkelstein MM. 2007. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. J Toxicol Environ Health-Part A-Current Issues 70:200–212.

Johannesson S, Gustafson P, Molnar P, Barregard L, Sallsten G. 2007. Exposure to fine particles ($PM_{2.5}$ and PM_1) and black smoke in the general population: Personal, indoor, and outdoor levels. J Expo Sci Environ Epidemiol 17:613–624.

Kim SY, Sheppard L, Kim H. 2009. Health effects of long-term air pollution: influence of exposure prediction methods. Epidemiology 20:442–450.

Kinney PL, Aggarwal M, Northridge ME, Janssen NAH, Shepard P. 2000. Airborne concentrations of $PM_{2.5}$ and diesel exhaust particles on Harlem sidewalks: A community-based pilot study. Environ Health Perspect 108:213–218.

Koistinen KJ, Edwards RD, Mathys P, Ruuskanen J, Künzli N, Jantunen MJ. 2004. Sources of fine particulate matter in personal exposures and residential indoor, residential outdoor and workplace microenvironments in the Helsinki phase of the EXPOLIS study. Scand J Work Environ Health 30 Suppl 2:36–46.

Laden F, Neas LM, Dockery DW, Schwartz J. 2000. Association of fine particulate matter from different sources with daily mortality in six U.S. cities. Environ Health Perspect 108:941–947.

Lagakos SW. 1988. Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable. Stat Med 7:257–294.

Lai HK, Bayer-Oglesby L, Colvile R, Gotschi T, Jantunen MJ, Künzli N, Kulinskaya E, Schweizer C, Nieuwenhuijsen

MJ. 2006. Determinants of indoor air concentrations of $PM_{2.5}$, black smoke, and NO_2 in six European cities (EXPOLIS study). Atmos Environ 40:1299–1313.

Lall R, Thurston GD. 2006. Identifying and quantifying transported vs. local sources of New York City $PM_{2.5}$ fine particulate matter air pollution. Atmos Environ 40:S333–S346.

Larson T, Gould T, Simpson C, Liu L-JS, Claiborn C, Lewtas J. 2004. Source apportionment of indoor, outdoor, and personal $PM_{2.5}$ in Seattle, Washington, using positive matrix factorization. J Air Waste Manag Assoc 54:1175–1187.

Lee K, Levy JI, Yanagisawa Y, Spengler JD, Billick IH. 1998. The Boston residential nitrogen dioxide characterization study: Classification and prediction of indoor NO_2 exposure. J Air Waste Manag Assoc 48:736–742.

Lee S, Liu W, Wang YH, Russell AG, Edgerton ES. 2008. Source apportionment of $PM_{2.5}$: Comparing PMF and CMB results for four ambient monitoring sites in the southeastern United States. Atmos Environ 42:4126–4137.

Lee SC, Wang B. 2006. Characteristics of emissions of air pollutants from mosquito coils and candles burning in a large environmental chamber. Atmos Environ 40:2128–2138.

Levy JI, Lee K, Spengler JD, Yanagisawa Y. 1998. Impact of residential nitrogen dioxide exposure on personal exposure: An international study. J Air Waste Manag Assoc 48:553–560.

Li Z, Hopke PK, Husain L, Qureshi S, Dutkiewicz VA, Schwab JJ, Drewnick F, Demerjian KL. 2004. Sources of fine particle composition in New York City. Atmos Environ 38:6521–6529.

Lin S, Munsie P, Hwang S-A, Fitzgerald E, Cayo MR. 2002. Childhood asthma hospitalization and residential exposure to state route traffic. Environ Res 88(A):73–81.

Little RJ, Rubin DB. 2002. Statistical Analysis with Missing Data, 2nd edition. New York: John Wiley.

Long CM, Sarnat JA. 2004. Indoor-outdoor relationships and infiltration behavior of elemental components of outdoor $PM_{2.5}$ for Boston-area homes. Aerosol Sci Technol 38:91–104.

Madsen C, Carlsen KCL, Hoek G, Oftedal B, Nafstad P, Meliefste K, Jacobsen R, Nystad W, Carlsen KH, Brunekreef B. 2007. Modeling the intra-urban variability of outdoor traffic pollution in Oslo, Norway: A GA(2)LEN project. Atmos Environ 41:7500–7511.

Maheswaran R, Elliott P. 2003. Stroke mortality associated with living near main roads in England and Wales: A geographical study. Stroke 34:2776–2780.

Marbury MC, Maldonado G, Waller L. 1996. The indoor air and children's health study: Methods and incidence rates. Epidemiology 7:166–174.

Marple VA, Rubow KL, Turner W, Spengler JD. 1987. Low flow rate sharp cut impactors for indoor air sampling: Design and calibration. J Air Pollution Control Assoc 37:1303–1307.

Martuzevicius D, Grinshpun SA, Reponen T, Gorny RL, Shukla R, Lockey J, Hu S, McDonald R, Biswas P, Kliucininkas L, LeMasters G. 2004. Spatial and temporal variations of $PM_{2.5}$ concentration and composition throughout an urban area with high freeway density: The greater Cincinnati study. Atmos Environ 38:1091–1105.

Meng QY, Turpin BJ, Lee JH, Polidori A, Weisel CP, Morandi M, Colome S, Zhang J, Stock T, Winer A. 2007. How does infiltration behavior modify the composition of ambient $PM_{2.5}$ in indoor spaces? An analysis of RIOPA data. Environ Sci Technol 41:7315–7321.

Moore DK, Jerrett M, Mack WJ, Künzli N. 2007. A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. J Environ Monit 9:246–252.

Morgenstern V, Zutavern A, Cyrys J, Brockow I, Gehring U, Koletzko S, Bauer CP, Reinhardt D, Wichmann HE, Heinrich J. 2007. Respiratory health and individual estimated exposure to traffic-related air pollutants in a cohort of young children. Occup Environ Med 64:8–16.

Nethery E, Teschke K, Brauer M. 2008. Predicting personal exposure of pregnant women to traffic-related air pollutants. Sci Total Environ 395:11–22.

Ogulei D, Hopke PK, Zhou L, Pancras JP, Nair N, Ondov JM. 2006. Source apportionment of Baltimore aerosol from combined size distribution and chemical composition data. Atmos Environ 40:S396–S410.

Ozkaynak H, Xue J, Weker R, Butler D, Koutrakis P, Spengler J. 1994. The particle team (PTEAM study): Analysis of the data. Final report, vol. III. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Ozkaynak H, Xue J, Spengler J, Wallace L, Pellizzari E, Jenkins P. 1996. Personal exposure to airborne particles and metals: Results from the particle TEAM study in Riverside, California. J Expo Anal Environ Epidemiol 6:57–78.

Pekney NJ, Davidson CI, Robinson A, Zhou LM, Hopke P, Eatough D, Rogge WF. 2006. Major source categories for

$PM_{2.5}$ in Pittsburgh using PMF and UNMIX. Aerosol Sci Technol 40:910–924.

Poplawski K, Gould T, Setton E, Allen R, Su J, Larson T, Henderson S, Brauer M, Hystad P, Lightowlers C, Keller P, Cohen M, Silva C, Buzzelli M. 2009. Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. J Exp Sci Environ Epidemiol 19:107–117.

Qin YJ, Kim E, Hopke PK. 2006. The concentrations and sources of $PM_{2.5}$ in metropolitan New York City. Atmos Environ 40:S312–S332.

Reungoat P, Chiron M, Gauvin S, Zmirou-Navier D, Momas I. 2005. Retrospective assessment of exposure to traffic air pollution using the ExTra index in the VESTA French epidemiological study. J Expo Anal Environ Epidemiol 15:524–533.

Rizzo MJ, Scheff PA. 2007. Fine particulate source apportionment using data from the US EPA speciation trends network in Chicago, Illinois: Comparison of two source apportionment models. Atmos Environ 41:6276–6288.

Roorda-Knape MC, Janssen NAH, de Hartog J, Van Vliet PHN, Harssema H, Brunekreef B. 1999. Traffic related air pollution in city districts near motorways. Sci Tot Environ 235:339–341.

Rosenlund M, Forastiere F, Stafoggia M, Porta D, Perucci M, Ranzi A, Nussio F, Perucci CA. 2008. Comparison of regression models with land-use and emissions data to predict the spatial distribution of traffic-related air pollution in Rome. J Expo Sci Environ Epidemiol 18:192–199.

Ross Z, English PB, Scalf R, Gunier R, Smorodinsky S, Wall S, Jerrett M. 2006. Nitrogen dioxide prediction in Southern California using land use regression modeling: Potential for environmental health analyses. J Expo Sci Environ Epidemiol 16:106–114.

Ross Z, Jerrett M, Ito K, Tempalski B, Thurston GD. 2007. A land use regression for predicting fine particulate matter concentrations in the New York City region. Atmos Environ 41:2255–2269.

Ryan P, LeMasters G, Biagini J, Bernstein D, Grinshpun SA, Shukla R, Wilson K, Villareal M, Burkle J, Lockey J. 2005. Is it traffic type, volume, or distance? Wheezing in infants living near truck and bus traffic. J Allergy Clin Immunol 116:279–284.

Ryan PH, LeMasters GK, Biswas P, Levin L, Hu S, Lindsey M, Bernstein DL, Lockey J, Villareal M, Khurana Hershey GK, Grinshpun SA. 2007. A comparison of proximity and land use regression traffic exposure models and wheezing in infants. Environ Health Perspect 115:278–284.

Sahsuvaroglu T, Arain A, Kanaroglou P, Finkelstein N, Newbold B, Jerrett M, Beckerman B, Brook J, Finkelstein M, Gilbert NL. 2006. A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario, Canada. J Air Waste Manag Assoc 56:1059–1069.

Sanchez BN, Houseman EA, Ryan LM. 2009. Residual-based diagnostics for structural equation models. Biometrics 65:104–115.

Sarnat JA, Long CM, Koutrakis P, Coull BA, Schwartz J, Suh HH. 2002. Using sulfur as a tracer of outdoor fine particulate matter. Environ Sci Technol 36:5305–5314.

Schauer JJ, Lough GC, Schafer MM, Christensen WF, Arndt MF, DeMinter JT, Park JS. 2006. Characterization of Metals Emitted from Motor Vehicles. Research Report 133. Health Effects Institute, Boston, MA.

Sioutas C, Koutrakis P, Godleski JJ, Ferguson ST, Kim CS, Burton RM. 1997. Fine particle concentrators for inhalation exposures: Effect of particle size and composition. J Aerosol Sci 28:1057–1071.

Sioutas C, Koutrakis P, Burton RM. 1995. A technique to expose animals to concentrated fine ambient aerosols. Environ Health Perspect 103:172–177.

Spencer MT, Shields LG, Sodeman DA, Toner SM, Prather KA. 2006. Comparison of oil and fuel particle chemical signatures with particle emissions from heavy and light duty vehicles. Atmos Environ 40:5224–5235.

Spiegelman D, Schneeweiss S, McDermott A. 1997. Measurement error correction for logistic regression models with an "Alloyed gold standard." Am J Epidemiol 145:184–196.

Sternbeck J, Sjodin A, Anderasson K. 2002. Metal emissions from road traffic and the influence of resuspension: Results from two tunnel studies. Atmos Environ 36:4735–4744.

Su JG, Brauer M, Ainslie B, Steyn D, Larson T, Buzzelli M. 2008. An innovative land use regression model incorporating meteorology for exposure analysis. Sci Total Environ 390:520–529.

Suh HH, Nishioka Y, Allen GA, Koutrakis P, Burton RM. 1997. The metropolitan acid aerosol characterization study: Results from the summer 1994 Washington, D.C. field study. Environ Health Perspect 105:826–834.

Sutton K, Caruso J. 1999. Liquid chromatography-inductively coupled plasma mass spectrometry. J Chromatogra A 856:243–258.

Szpiro AA, Sheppard L, Lumley T. Accounting for errors from predicting exposures in environmental epidemiology and environmental statistics. 2008. University of Washington Biostatistics Working Paper Series, working paper 330.

Thurston G, Spengler J. 1985. A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. Atmos Environ 19:9–25.

U.S. Environmental Protection Agency. 2001. Candles and incense as potential sources of indoor air pollution: Market analysis and literature review. EPA-600/R-01-001. Office of Research and Development, Research Triangle Park, NC.

Vaeth M, Skovlund E. 2004. A simple approach to power and sample size calculations in logistic regression and Cox regression models. Stat Med 23:1781–1792.

Van Roosbroeck S, Wichmann J, Janssen NA, Hoek G, van Wijnen JH, Lebret E, Brunekreef B. 2006. Long-term personal exposure to traffic-related air pollution among school children, a validation study. Sci Total Environ 368:565–573.

Van Roosbroeck S, Hoek G, Meliefste K, Janssen NA, Brunekreef B. 2008. Validity of residential traffic intensity as an estimate of long-term personal exposure to traffic-related air pollution among adults. Environ Sci Technol 42:1337–1344.

Wacholder S, Armstrong B, Hartge P. 1993. Validation studies using an alloyed gold standard. Am J Epidemiol 137:1251–1258.

Wallace L. 2005. Real-time measurements of black carbon indoors and outdoors: A comparison of the photoelectric aerosol sensor and the aethalometer. Aerosol Sci Technol 39:1015–1025.

Wallace LA. 1996. Indoor particles: A review. J Air Waste Manag Assoc 46:98–126.

Wallace LA, Mitchell H, O'Connor GT, Neas L, Lippmann M, Kattan M, Koenig J, Stout JW, Vaughn BJ, Wallace D, Walter M, Adams K, Liu LS. 2003. Particle concentrations in inner-city homes of children with asthma: The effect of smoking, cooking, and outdoor pollution. Environ Health Perspect 111:1265–1272.

Watson J, Chow J, Frazier C. 1999. X-ray fluorescence analysis of ambient air samples. In: Elemental Analysis of Airborne Particles (Landberger S, Creatchman M, eds). Gordon and Breach Publishers, NJ.

Williams R, Suggs J, Rea A, Leovic K, Vette A, Croghan C, Sheldon L, Rodes C, Thornburg J, Ejire A, Herbst M, Sanders Jr. W. 2003. The Research Triangle Park particulate matter panel study: PM mass concentration relationships. Atmos Environ 37:5349–5363.

Wu J, Lurmann F, Winer A, Lu R, Turco R, Funk T. 2005. Development of an individual exposure model for application to the Southern California Children's Health Study. Atmos Environ 39:259–273.

Yakoleva E, Hopke PK, Wallace L. 1999. Receptor modeling assessment of particle total exposure assessment methodology data. Environ Sci Technol 33:3645–3652.

Yanagisawa Y, Nishimura H. 1982. A badge-type personal sampler for measurement of personal exposure to NO_2 and NO in ambient air. Environ Int 8:235–242.

Yeomans KA, Golder PA. 1982. The Guttman-Kaiser criterion as a predictor of the number of common factors. The Statistician 31:221–229.

Yli-Tuomi T, Lanki T, Hoek G, Brunekreef B, Pekkanen J. 2008. Determination of the sources of indoor $PM_{2.5}$ in Amsterdam and Helsinki. Environ Sci Technol 42:4440–4446.

Zhao W, Hopke PK, Gelfand EW, Rabinovitch N. 2007. Use of an expanded receptor model for personal exposure analysis in schoolchildren with asthma. Atmos Environ 41:4084–4096.

Zhao WX, Hopke PK, Norris G, Williams R, Paatero P. 2006. Source apportionment and analysis on ambient and personal exposure samples with a combined receptor model and an adaptive blank estimation strategy. Atmos Environ 40:3788–3801.

Zhou Y, Levy JI. 2007. Factors influencing the spatial extent of mobile source air pollution impacts: A meta-analysis. BMC Public Health 7:89.

Zmirou D, Gauvin S, Pin I, Sahraoui F, Just J, Moullec YL, Bremont F, Cassadou S, Reungoat P, Albertini M, Lauvergne N, Chiron M, Labbe A, Vesta investigators. 2004. Traffic related air pollution and incidence of childhood asthma: Results of the Vesta case-control study. J Environ Community Health 58:18–23.

Zota A, Adamkiewicz G, Levy JI, Spengler JD. 2005. Ventilation in public housing: Implications for indoor nitrogen dioxide concentrations. Indoor Air 15:393–401.

APPENDIX A. Scatterplots of Central Site Concentration Data and Residential Outdoor Measurements

For our LUR analyses, we matched central site monitoring data with our indoor and outdoor observations for the precise hours for which sampling was conducted. The relevant central site monitors were those with hourly concentration data throughout the sampling period, collected at a location relevant to our study domain (which included portions of Suffolk, Norfolk, and Middlesex counties). While we used the monitor in Roxbury (Harrison Avenue, U.S. EPA site ID 250250042) for our primary analyses for all three pollutants, we considered the following monitors for sensitivity analyses (in addition to the average across monitors and a background monitor for NO_2 for summer months only):

- $PM_{2.5}$ and EC: North End (174 North Street, U.S. EPA site ID 250250043)
- NO_2: Kenmore Square (U.S. EPA site ID 250250002), South Boston (531a East First Street, U.S. EPA site ID 250250040)

The Roxbury monitor is located on a residential roadway within a dense urban neighborhood and is considered a "population-exposure" monitor by the U.S. EPA. The North End monitor is located in a dense urban neighborhood, proximate to a major highway and multiple surface roads, and is also considered a population-exposure monitor. The Kenmore Square monitor is located at ground level on a large urban roadway and is considered a "highest-concentration" monitor by the U.S. EPA. The South Boston monitor is located in an industrial zone and is considered a population-exposure monitor but as part of the industrial monitoring network. Thus, some differences across monitors would be anticipated, based on neighborhood and monitoring objectives.

Of note, our residential sampling was not uniform across the year, with sampling sessions that at times overlapped between homes, and with sampling durations that varied slightly across homes. Thus, a time-series analysis of these central site data was not considered informative, and we instead evaluated scatterplots of central site data against each other and in comparison with our measurements.

Figure A.1 shows that $PM_{2.5}$ concentrations as measured at the central site monitors in Roxbury and the North End were highly correlated ($r = 0.87$), indicating that our LUR models would be relatively insensitive to the choice of monitor. For EC (Figure A.2), the correlation was slightly lower but still substantial ($r = 0.81$), with systematically higher concentrations generally observed at the North End monitor. For NO_2 (Figures A.3 to A.5), there is more scatter in the associations, although the Roxbury monitor is more tightly correlated with the other two monitors ($r = 0.75$ with Kenmore Square, $r = 0.87$ with South Boston) than the other two monitors are with one another ($r = 0.59$). This is clearly reflective of the specific sitings of the Kenmore Square and South Boston monitors, which are meant to capture local traffic and industrial sources, respectively. The largest differentials between these monitors tended to occur during the summertime when the Boston Red Sox were playing (given that the Kenmore Square monitor is proximate to Fenway Park). These analyses emphasize that monitors other than the one in Roxbury are unlikely to be informative for our LUR models, either because they provide no additional information value (for $PM_{2.5}$ and EC) or because they do not provide broadly representative concentrations (for NO_2).

For our residential outdoor concentrations relative to central site concentrations at the Roxbury monitoring site, the correlation was strong for $PM_{2.5}$ (Figure A.6). Had we used the North End monitor, the correlation would have been slightly weaker. For EC, the relationship between the central site monitor and our residential outdoor concentrations was weaker, although with a more significant association during the cooling season (Figure A.7). Correlations

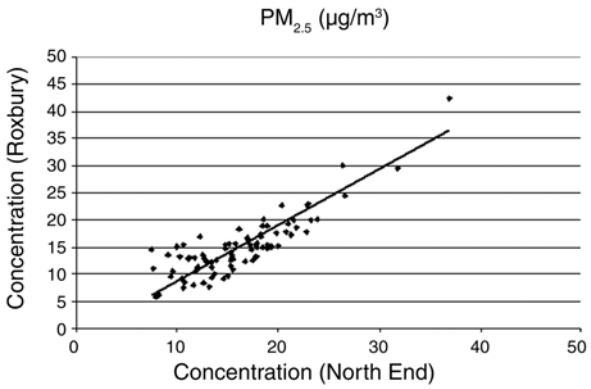

Figure A.1. $PM_{2.5}$ (µg/m³) at Roxbury central site vs. at North End central site.

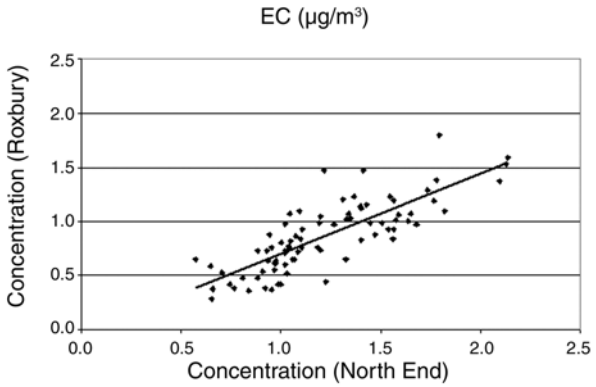

Figure A.2. EC (µg/m³) at Roxbury central site vs. at North End central site.

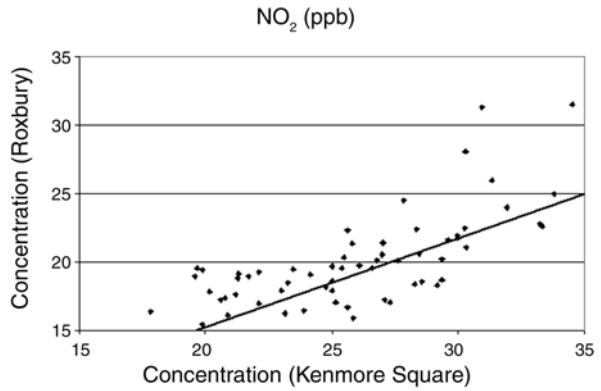

Figure A.3. NO_2 (ppb) at Roxbury central site vs. at Kenmore Square central site.

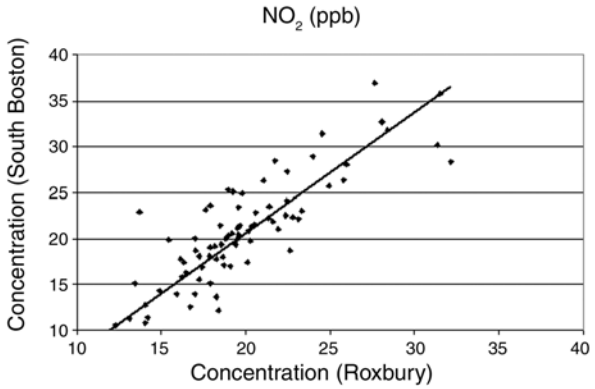

Figure A.4. NO_2 (ppb) at Kenmore Square central site vs. at South Boston central site.

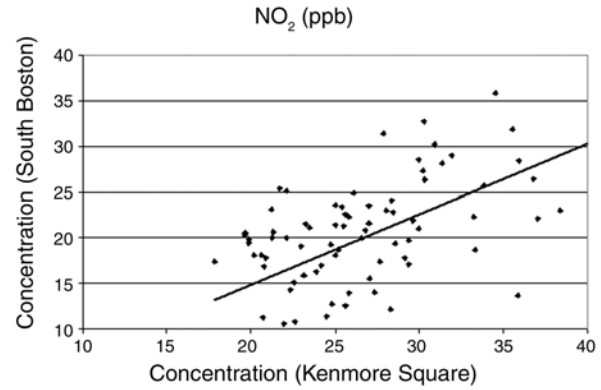

Figure A.5. NO_2 (ppb) at Roxbury central site vs. at South Boston central site.

between the North End monitor and our residential outdoor concentrations were somewhat weaker. Finally, the residential outdoor concentrations of NO_2 were also weakly associated with Roxbury central site concentrations (Figure A.8), and the association was not improved by using the alternative central site monitors.

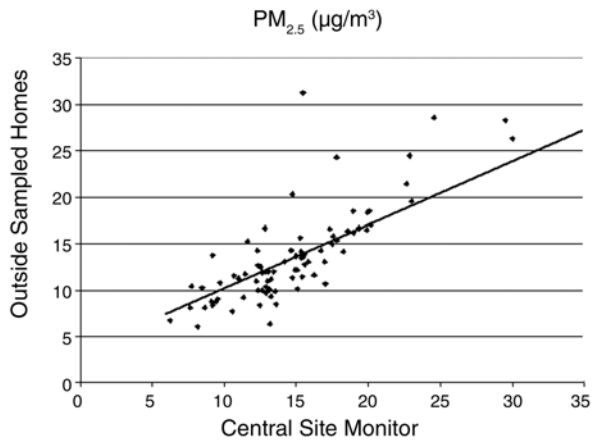

Figure A.6. $PM_{2.5}$ (µg/m³) measured outside of homes vs. at Roxbury central site.

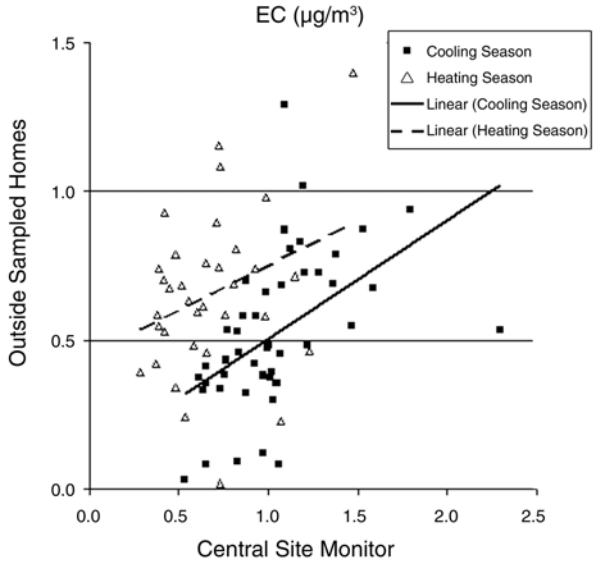

Figure A.7. EC measured outside of homes ($m^{-1} \times 10^{-5}$) vs. at Roxbury central site (µg/m³).

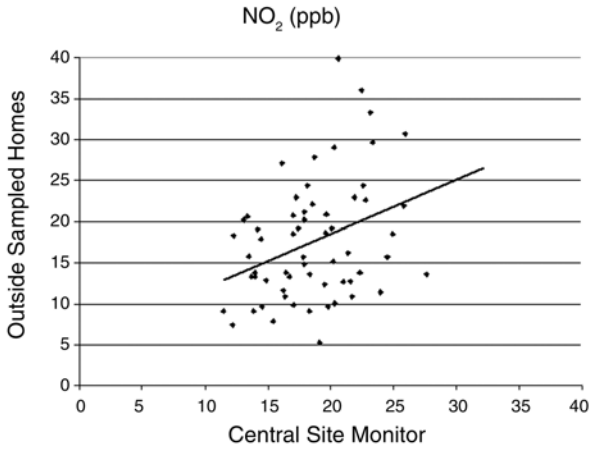

Figure A.8. NO_2 (ppb) measured outside of homes vs. at Roxbury central site.

APPENDIX B. Results of Sensitivity Analyses for Multivariate LUR Model Results for Outdoor $PM_{2.5}$, EC, and NO_2

Tables B.1, B.2, and B.3 consider alternative traffic indicators and show the predictor with the highest significance in each category as well as alternative buffer lengths for the measures of roadway length. For each pollutant, the final model from Table 9 is in bold, and models including diesel-based variables not available for all observations are italicized.

Appendix Table B.1. Sensitivity Analyses for $PM_{2.5}$[a]

Indicator Type	Traffic Indicator	Estimate(s), *P* Value(s) from Multivariate Model	Model R^2
Base model (without traffic)			0.74
Cumulative density scores	Unweighted 500-m traffic density (n = 57)	$\beta1 = 4.74 \times 10^{-4}$ (0.17)	0.75
Summary measures	Total roadway length within:		
	50 m (n = 57)	$\beta1 = 1.31 \times 10^{-4}$ (0.35)	0.74
	100 m (n = 57)	**$\beta1 = 1.48 \times 10^{-4}$ (0.02)**	**0.76**
	200 m (n = 57)	$\beta1 = 1.42 \times 10^{-5}$ (0.56)	0.74
	300 m (n = 57)	$\beta1 = 5.54 \times 10^{-6}$ (0.57)	0.74
Distance-based measures	Distance to nearest designated truck route (n = 57)	$\beta1 = 1.80 \times 10^{-5}$ (0.62)	0.74
Characteristics of nearest major road	ADT (n = 57)	$\beta1 = -3.04 \times 10^{-6}$ (0.25)	0.74

[a] Bold indicates a selected term in the final model from Table 9.

Appendix Table B.2. Sensitivity Analyses for EC[a,b]

Indicator Type	Traffic Indicator	Estimate(s), *P* Value(s) from Multivariate Model	Model R^2
Base model (without traffic)			0.31
Cumulative density scores	Unweighted 500-m traffic density (n = 54)	$\beta1 = 5.39 \times 10^{-4}$ (0.49) $\beta2 = 5.63 \times 10^{-3}$ (0.41)	0.39
Summary measures	Total roadway length within		
	50 m	$\beta1 = 3.94 \times 10^{-4}$ (0.18)	
	× Still winds (n = 54)	$\beta2 = 4.08 \times 10^{-3}$ (0.01)	0.41
	100 m	$\beta1 = 2.15 \times 10^{-4}$ (0.14)	
	× Still winds (n = 54)	$\beta2 = 1.27 \times 10^{-3}$ (0.03)	0.47
	200 m	**$\beta1 = 1.10 \times 10^{-4}$ (0.01)**	
	× Still winds (n = 54)	**$\beta2 = 4.38 \times 10^{-4}$ (0.02)**	**0.52**
	300 m	$\beta1 = 2.99 \times 10^{-5}$ (0.11)	
	× Still winds (n = 54)	$\beta2 = 2.01 \times 10^{-4}$ (0.04)	0.48
Distance-based measures	To nearest highway (> 19,000 cars/day)	$\beta1 = 0.452$ (0.06)	
	× Still winds (n = 54)	$\beta2 = 0.549$ (0.04)	0.45
Characteristics of nearest major road	*Diesel fraction*	*$\beta1 = -1.06$ (0.02)*	
	× Still winds (n = 34)	*$\beta2 = 34.6$ (0.02)*	*0.54*
	Trucks per day	*$\beta1 = -7.41 \times 10^{-5}$ (0.06)*	
	× Still winds (n = 34)	*$\beta2 = 3.51 \times 10^{-3}$ (0.02)*	*0.54*
	Trucks / distance to major road	*$\beta1 = -6.31 \times 10^{-3}$ (0.03)*	
	× Still winds (n = 34)	*$\beta2 = 0.119$ (0.05)*	*0.54*

[a] Bold indicates a selected term in the final model from Table 9.

[b] Italics indicate diesel-based variables not available for all observations.

Appendix Table B.3. Sensitivity Analyses for NO_2[a,b]

Indicator Type	Traffic Indicator	Estimate(s), *P* Value(s) from Multivariate Model	Model R^2
Base model (without traffic)			0.39
Cumulative density scores	Unweighted 100-m traffic density	β1 = 0.055 (0.004)	
	× Obstructed from major road (n = 50)	β2 = −0.051 (0.004)	0.55
	Kernel-weighted 50-m traffic density	β1 = 0.034 (0.02)	
	× Obstructed (n = 50)	β2 = −0.056 (0.002)	0.55
	Density of larger roads (> 8500 cars/ day) within 200 m	β1 = 589.4 (0.049)	
	× Obstructed (n = 50)	β2 = −760.9 (0.0095)	0.52
Summary measures	Total roadway length within		
	50 m	**β1 = 0.0144 (< 0.0001)**	
	× Obstructed (*n* = 50)	**β2 = −0.0094 (0.005)**	**0.56**
	100 m	β1 = 0.0022 (0.34)	
	× Obstructed (n = 50)	β2 = −0.0042 (0.005)	0.54
	200 m	$\beta1 = 9.28 \times 10^{-4}$ (0.22)	
	× Obstructed (n = 50)	$\beta2 = -1.25 \times 10^{-3}$ (0.008)	0.52
	300 m	$\beta1 = 2.31 \times 10^{-4}$ (0.55)	
	× Obstructed (n = 50)	$\beta2 = -6.65 \times 10^{-4}$ (0.0095)	0.50
	ADT × Length within 200 m	$\beta1 = 1 \times 10^{-7}$ (0.21)	
	× Obstructed (n = 50)	$\beta2 = -1 \times 10^{-7}$ (0.10)	0.47
Distance-based measures	To nearest highway (>19,000 cars/day)	β1 = 0.0176 (0.12)	
	× Obstructed (n = 50)	β2 = −0.0195 (0.07)	0.50
Characteristics of nearest major road	*Trucks per day*	*β1 = .0061 (0.01)*	
	× Obstructed (n = 34)	*β2 = .00596 (0.01)*	*0.59*

[a] Bold indicates a selected term in the final model from Table 9.

[b] Italics indicate diesel-based variables not available for all observations.

APPENDIX C. Constrained Factor Analysis for CAPS Samples Collected at One Location

In total, 20 ambient and concentrated CAPs samples were collected from January through February 2008. To provide qualitative comparisons with our primary factor analysis, we focused on the CAPs samples, given fewer LOD issues and the availability of postconcentrator EC concentrations.

For these CAPs samples, many of the highest correlations among constituents were similar to those observed for our ACCESS study samples; S was strongly correlated with Se ($r = 0.73$), Ni with V ($r = 0.90$), and Al with Ba ($r = 0.92$). The correlation structures for the CAPs samples differed for some constituents, however; while EC was positively associated with constituents such as S, P, and Ca, it showed stronger associations with Al, Ba, Ni, and V (correlations not shown). This may have been an indication of the significant contribution of residential fuel oil for these heating-season samples, as well as the fact that these substudy samples were collected at one location with significant diesel traffic. More broadly, there were strong positive correlations among many CAPs constituents, especially those analyzed using ICP–MS, emphasizing that meteorological variability influenced many constituents similarly.

Applying identical factor analytic methods (as described in Constrained Factor Analysis, under Statistical Methods and Data Analysis) to our CAPs samples also resulted in a five-factor model (Figure C.1). In this case, many of the constituents measured (especially those measured by ICP–MS), had relatively high loadings on the first factor, which was likely indicative of the generally high correlation among constituents and was interpreted as meteorological variability influencing all constituents similarly. This would correspond most closely to Factor 1 in our primary model (Figure 4); note, however, that the use of samples from only the heating season in this analysis may have reduced the linkage between meteorological characteristics and long-range transport of emissions from electricity generation. This point is reinforced by the second factor in Figure C.1, which includes S, Se, and As and therefore appears indicative of coal combustion, which was found in Factor 1 in our primary model as well. Loading most heavily on the third factor in Figure C.1 are Ca, Fe (by XRF analysis), K, and Si, crustal elements indicative of resuspended road dust (Table 3). This is a somewhat different patterning than seen in Factor 5 in our primary model, which was interpreted as road dust and resuspension, and is also more closely connected to constituents originally hypothesized to indicate road dust (Table 3). The CAPs samples, collected from a single major roadway, may provide a clearer signal of road dust than samples from our relatively less trafficked residential settings. The fourth factor in Figure C.1 includes some loading of constituents previously associated with fuel oil (Ni, V) and diesel (EC, S, P, Zn, Fe). The fact that these sources could not be separated within the factor analysis may be partly a function of a small sample size, but may also be attributable to the use of measurements at a single location in a single season. Finally, the fifth factor in Figure C.1 includes a number of constituents associated with brake wear or other forms of traffic-related road dust (e.g., Tl, Al, Ba). Thus, while the models in Figure 4 and Figure C.1 differ somewhat (and would not be anticipated to be identical, given differences in study design and issues related to sample size), the CAPs factor analysis generally corroborates the structure and interpretation of our primary factor analysis, while pointing out some uncertainties associated with the interpretation of our brake wear and road dust terms. More generally, both models indicate the viability of disentangling multiple traffic sources and residential fuel oil from long-range transport and meteorological influences.

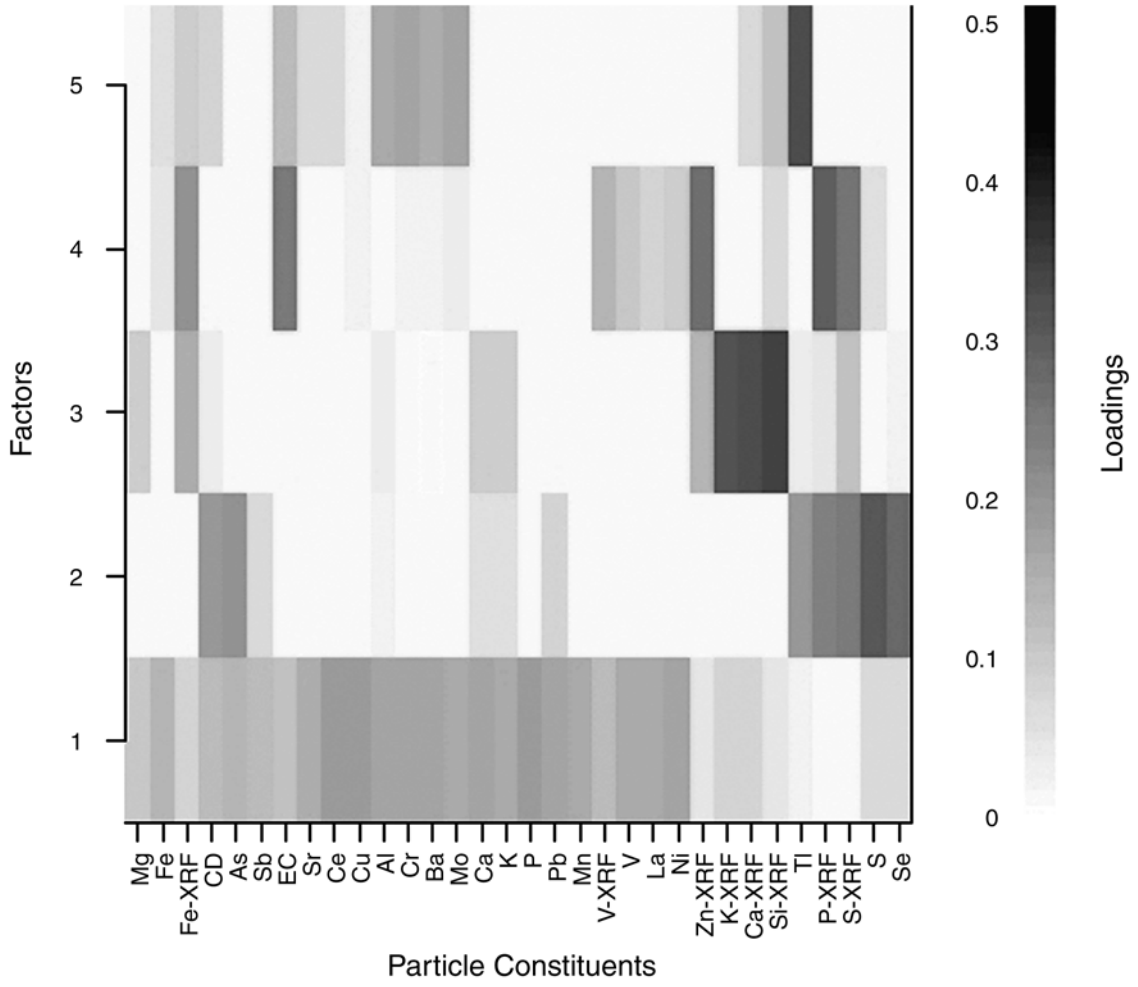

Figure C.1. Factor loading matrix for CAPs samples using a five-factor model. Particle constituents are sorted by their position in a hierarchical clustering dendrogram to optimize visual interpretation by placing correlated constituents together.

APPENDICES AVAILABLE ON THE WEB

Appendices D, E, F, and G contain supplemental material not included in the printed report. They are available on the HEI Web site *http://pubs.healtheffects.org*.

Appendix D. Questionnaire Administered to Air Pollution Study Participants to Gather Housing Characteristics and Occupant Behavior Data

Appendix E. Supplemental Questionnaire Derived from ACCESS Prenatal Questionnaire and Administered Directly to Noncohort Air Pollution Study Participants to Gather Housing Characteristics and Occupant Behavior Data

Appendix F. Results of Cholesky Residual Analysis of Normality for Constrained Factor Analyses

Appendix G. Derivation of the Power Expression for Exposure Misclassification Analysis

ABOUT THE AUTHORS

Dr. Jonathan I. Levy is currently a professor of environmental health in the Department of Environmental Health at the Boston University School of Public Health. He holds an Sc.D. in environmental science and risk management from the Harvard School of Public Health and an A.B. in applied mathematics from Harvard College. Dr. Levy's research interests include development of air pollution exposure and health risk models for urban environments, with an emphasis on spatial and demographic variability in exposure and susceptibility.

Dr. Jane E. Clougherty is currently a Senior Research Scientist at the New York City Department of Health and Mental Hygiene, where she is primarily responsible for the New York City Community Air Survey (NYCCAS), an ongoing assessment of fine-scale variation in multiple air pollutants throughout the five boroughs of New York. She holds an A.B. in economics and environmental studies from the University of Chicago, an M.Sc. from the McMaster University School of Geography and Geology, and an Sc.D. in environmental health from the Harvard School of Public Health. Dr. Clougherty's research interests focus on the combined health effects of social and physical environmental exposures in urban communities, and in the development of methods for fine-scale geographic exposure assessment, to elucidate distinctions among spatially clustered social (e.g., poverty, noise) and physical (e.g., traffic-related air pollution) exposures.

Dr. Lisa K. Baxter is currently a postdoctoral fellow at the U.S. EPA's National Exposure Research Laboratory. She received her M.S. and Sc.D. in environmental health from the Harvard School of Public Health. Dr. Baxter's research focuses on improving air pollution exposure estimates for epidemiologic studies.

Dr. E. Andres Houseman is an assistant professor of community health in the Center for Environmental Health and Technology at The Warren Alpert Medical School of Brown University. He has an Sc.D. in biostatistics from the Harvard School of Public Health and an A.B. in applied mathematics from the University of California at Berkeley. Dr. Houseman's research interests include statistical models (especially latent-variable models) and their application to environmental epidemiology, to molecular epidemiology, and to the combination of data from both fields.

Dr. Christopher Paciorek is a visiting assistant professor in the Department of Statistics at the University of California at Berkeley and an adjunct assistant professor in the Department of Biostatistics at the Harvard School of Public Health. He has a master's degree and a Ph.D. in statistics from Carnegie Mellon University, and a master's degree in botany (ecology) from Duke University. Dr. Paciorek's research interests are in the areas of Bayesian statistics and statistics for spatial and spatiotemporal data. The applied focus of his work is in environmental applications, including environmental health, ecology, and climate. Much of his recent work has been applied to improving exposure assessment and understanding statistical issues in environmental epidemiology.

OTHER PUBLICATIONS RESULTING FROM THIS RESEARCH

Baxter LK, Clougherty JE, Laden F, Levy JI. 2007. Predictors of concentrations of nitrogen dioxide, fine particulate matter, and particle constituents inside of lower socioeconomic status urban homes. J Exp Sci Environ Epidemiol 17:433–444.

Baxter LK, Clougherty JE, Paciorek CJ, Wright R, Levy JI. 2007. Predicting residential indoor concentrations of nitrogen dioxide, fine particulate matter, and elemental carbon using questionnaire and geographic information system based data. Atmos Environ 41:6561–6571.

Baxter LK, Suh HH, Paciorek CJ, Clougherty JE, Levy JI. 2006. Predicting infiltration factors in urban residences for a cohort study. Healthy Buildings 2006, Lisboa, Portugal, International Society of Indoor Air Quality and Climate.

Baxter LK, Wright RJ, Paciorek CJ, Laden F, Suh HH, Levy JI. 2010. Effects of exposure measurement error in the analysis of health effects from traffic-related air pollution. J Exp Sci Environ Epidemiol 20(1):101–111.

Clougherty JE, Houseman EA, Levy JI. 2009. Examining intra-urban variation in fine particle mass constituents using GIS and constrained factor analysis. Atmos Environ 43:5545–5555.

Clougherty JE, Levy JI, Kubzansky LD, Ryan PB, Suglia SF, Canner MJ, Wright RJ. 2007. Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology. Environ Health Perspect 115:1140–1146.

Clougherty JE, Wright RJ, Baxter LK, Levy JI. 2008. Land use regression modeling of intra-urban residential variability in multiple traffic-related air pollutants. Environ Health 7:17.

ABBREVIATIONS AND OTHER TERMS

ACCESS	Asthma Coalition for Community, Environment, and Social Stress
ADT	average daily traffic
AER	air exchange rate
BIC	Bayesian information criterion
CAPs	concentrated ambient particles
CART	classification and regression tree
EC	elemental carbon
F_{INF}	infiltration factor
GIS	geographic information system
I/O	indoor–outdoor
ICP–MS	inductively coupled plasma–mass spectrometry
K–G	Kaiser–Guttman criterion
LOD	limit of detection
LUR	land-use regression
MassDEP	Massachusetts Department of Environmental Protection
MHD	Massachusetts Highway Department
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
NNMF	nonnegative matrix factorization
MVN	multivariate normal
NO	nitric oxide
NO_2	nitrogen dioxide
OC	organic carbon
OR	odds ratio
PEM	personal environmental monitor
$PM_{2.5}$	fine particulate matter with aerodynamic diameter $\leq$ 2.5 μm
PMF	Positive Matrix Factorization
Q–Q	quantile–quantile
R^2	coefficient of determination
RMSE	root mean square error
SANN	simulated annealing
SEM	structural equation modeling
U.S. EPA	U.S. Environmental Protection Agency
XRF	x-ray fluorescence spectroscopy
z~	**z** is distributed as

ELEMENTS

Al	aluminum
As	arsenic
Ba	barium
Ca	calcium
Cd	cadmium
Ce	cerium
Cr	chromium
Cu	copper
Fe	iron
K	potassium
La	lanthanum
Mg	magnesium
Mn	manganese
Mo	molybdenum
Ni	nickel
P	phosphorus
Pb	lead
S	sulfur
Sb	antimony
Se	selenium
Si	silicon
Sr	strontium
Tl	thallium
V	vanadium
Zn	zinc

Research Report 152, *Evaluating Heterogeneity in Indoor and Outdoor Air Pollution Using Land-Use Regression and Constrained Factor Analysis*, J.I. Levy et al.

INTRODUCTION

The Health Effects Institute has demonstrated a long-standing interest in research to improve the characterization of exposure to air pollutants for epidemiologic studies. As part of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS*), methodologic work by the investigators advanced our understanding of the effects of error in measuring pollution in time-series studies (Samet et al. 2000). The Relationships of Indoor, Outdoor, and Personal Air (RIOPA) studies examined the relationships between pollutant concentrations in indoor, outdoor, and personal air (Weisel et al. 2005; Turpin et al. 2007) and established the RIOPA database that continues to be a resource for research in this area. Brunekreef and his colleagues (2005) studied correlations between ambient, indoor, and personal air concentrations of fine particulate matter with an aerodynamic diameter ≤ 2.5 μm ($PM_{2.5}$) and its components measured for groups of elderly cardiovascular patients in two European cities. Other HEI-funded studies have focused on improving the characterization of exposures to particulates and gases for sensitive subpopulations in U.S. cities (Koutrakis et al. 2005). The recent HEI-funded extended analysis of the American Cancer Society study included detailed studies of Los Angeles and New York, designed to better understand how refining exposures at the within-city, or intra-urban, level might affect the size and significance of health effects estimates (Krewski et al. 2009). Recognizing growing concern about the role of traffic in urban air pollution, HEI recently published a critical review of the literature on the emissions, exposures, and health effects of traffic-related air pollutants (HEI Panel on the Health Effects of Traffic-Related Air Pollution 2010). The study by Dr. Jonathan I. Levy of the Harvard School of Public Health falls squarely within this body of research.

Levy submitted an application to HEI under Request for Applications 04-5, the Walter A. Rosenblith New Investigator Award. This award was established to provide support for an outstanding new investigator at the assistant professor level to conduct work in the area of air pollution and health and is unrestricted with respect to the specific topic of research. In his application, "Using Geographic Information Systems (GIS) to Evaluate Heterogeneity in Indoor and Outdoor Concentrations of Particle Constituents," Levy proposed an approach to improve existing GIS methods for predicting intra-urban exposures for use in epidemiologic studies. His goals were to (1) develop models for predicting intra-urban variation in outdoor concentrations of $PM_{2.5}$, elemental carbon (EC), nitrogen dioxide (NO_2), and selected particle species related to various sources using monitoring data, GIS data related to traffic and land-use characteristics, and seasonal and meteorological factors and (2) develop models that could explain the influence of housing characteristics and occupant behaviors on indoor levels and sources of these air pollutants using questionnaire data or publicly available information or both. Despite some concerns about whether the models would have adequate predictive power and about their generalizability to other settings, the HEI Research Committee thought the project was an innovative and promising approach to filling critical gaps in the exposure assessment literature.

SCIENTIFIC BACKGROUND

Epidemiologic studies of exposure to air pollution have typically relied on data from centrally located ambient air quality monitors. However, such data are not sufficient for capturing the spatial variability of pollutant concentrations at the local scale, in particular at the intra-urban scale at which traffic-related air pollution is both highest and most variable (HEI Panel on the Health Effects of Traffic-Related Air Pollution 2010). The ideal approach

Dr. Jonathan I. Levy's three-year study, "Using Geographic Information Systems (GIS) to Evaluate Heterogeneity in Indoor and Outdoor Concentrations of Particle Constituents," began in September 2005. Total expenditures were $298,400. The draft Investigators' Report from Levy and colleagues was received for review in February 2009. A revised report, received in September 2009, was accepted for publication in October 2009. During the review process, the HEI Health Review Committee and the investigators had the opportunity to exchange comments and to clarify issues in both the Investigators' Report and the Review Committee's Critique.

* A list of abbreviations and other terms appears at the end of the Investigators' Report.

would be to measure each individual's personal exposure to traffic-specific pollutants over time, but this is difficult, intrusive, expensive, and generally not feasible for very large populations. Investigators have consequently sought ways to predict, or to model, individual-level exposures from more readily available data.

At the time Levy's study was considered for funding, a body of research was emerging that relied on a variety of surrogates (also referred to as "proxies") for such exposures, either alone or in combination (Krämer et al. 2000; Janssen et al. 2001; Hoek et al. 2001, 2002; Gehring et al. 2002). These surrogates included individual pollutants previously associated with traffic emissions (e.g., carbon monoxide, NO_2, $PM_{2.5}$, benzene, and EC) and various measures of traffic levels or density (e.g., vehicle mix and volume, numbers of vehicles over time) or of proximity to traffic (e.g., traffic density within varying distances, distances to roadways, self-reported traffic exposures).

Each type of surrogate has limitations in predicting personal exposures to traffic-related pollutants (Jerrett et al. 2005a). An individual's personal exposure to the few pollutants used as markers for traffic can also be affected by a number of other factors, including time and activity patterns, indoor sources of pollution, meteorological conditions, land-use patterns, and socioeconomic variables. Also, traffic is a source of a complex mixture of pollutants, and the limited set of pollutants studied may not be the causal agents for the health effects observed.

Land-use regression (LUR) models have been developed to take advantage of additional types and sources of data to augment and improve upon these simpler methods (Veen et al. 1997; Lebret et al. 2000; Hoek et al. 2008). LUR uses a variety of nearby land-use patterns, traffic data, physical-site characteristics, and other variables as independent predictors for concentrations of the pollutants of interest. Brauer and colleagues (2007), for example, found that accounting for the compounded influence of multiple roadways and supplementing basic proximity measures with more detailed land-use information, such as locations of heavy traffic and street canyons, substantially improved regression-based estimates of traffic-related $PM_{2.5}$ concentrations in three European cities.

The fundamental problem that these developments in exposure-assessment methods progressively have sought to address was that of exposure-measurement error — the error that can arise when surrogate measures are used in place of "true" individual-level exposures — and the implications this error could have for the findings of observational epidemiologic studies of traffic-related air pollution and health. Exposure-measurement error can take several forms: Berkson type error, which tends to reduce the precision of the estimate; "classical" measurement error, which tends to bias the estimate of the effect toward the null; or, more commonly, some indeterminate combination of the two. Each reduces the overall power of a study to detect an effect of an exposure on health outcomes (Armstrong 1998; Dominici et al. 2000; Zeger et al. 2000).

Levy and colleagues set out to improve upon methods used to characterize individual-level exposure to traffic-related pollutants for epidemiologic studies. They wanted to address the limitations of simple proximity models and other models relying primarily on spatially distributed land-use characteristics and to incorporate other potentially important time-varying predictors of pollutant concentrations. Furthermore, with scientific interest growing in the application of these newer approaches for characterizing spatial variability in air pollutants, they wanted to gain insight into their implications for the findings of epidemiologic studies.

The investigators were also concerned about a differential form of exposure-measurement error — one not addressed in most studies — that could arise from the use of residential outdoor air pollutant levels as a measure of individual-level exposure. People generally spend most of their time indoors (Leech et al. 2002), and studies have shown that taking indoor exposures into account is important in estimating personal exposures, particularly for pollutants with significant indoor sources (for example, NO_2) (Levy et al. 1998; Zipprich et al. 2002). Levy further hypothesized that the relationship between indoor and outdoor concentrations (i.e., indoor/outdoor ratios) of individual pollutants might be correlated with factors such as socioeconomic status, housing quality, or other factors that also independently affect the health outcome of interest. Use of a surrogate for individual-level exposure based on outdoor concentrations at a residence might therefore differ in systematic ways among populations of different socioeconomic status.

Finally, Levy and his colleagues also sought to define more clearly the nature of the pollutants and their concentrations that could be related to traffic and those that could have other sources, both outdoor and indoor. As discussed above, most early studies of traffic-related exposure have relied on individual pollutants believed to be markers for traffic (NO_2, $PM_{2.5}$, EC) but which also have other sources. At the time Levy's study was initiated, few investigators had used source apportionment in epidemiologic studies to attribute exposure to emissions from specific types of sources, particularly indoor sources. The use of source apportionment techniques within epidemiologic studies offers the opportunity to improve the understanding of potential causal relationships between particular exposures and observed health outcomes and thus to better inform policy decisions (Laden et al. 2000; Mar et al. 2000).

SUMMARY OF STUDY

SPECIFIC AIMS

While the overall objectives of the study remained the same as those originally stated in the proposal, the specific objectives were further differentiated and the methods for meeting them were refined over the course of the study. The specific aims stated in the report are as follows:

1. Using monitoring data collected as part of a birth cohort study in Boston, Massachusetts, develop GIS-based LUR models to explain spatial and temporal variability in residential outdoor concentrations of $PM_{2.5}$, EC, and NO_2.
2. Using constrained factor analysis on particle-constituent concentrations and NO_2 measures, determine source types contributing to variability in residential outdoor concentrations.
3. Apply LUR modeling techniques to the factor scores estimated through constrained factor analysis on residential outdoor concentrations, to determine whether GIS covariates and other predictors explain factor variability and thereby support initial factor interpretations.
4. Using indoor and outdoor residential monitoring data, develop physically interpretable regression models exploring indoor–outdoor relationships for EC, $PM_{2.5}$, NO_2, and particle constituents, relying on questionnaire data to characterize indoor sources and ventilation.
5. Develop GIS-based LUR models including terms for indoor sources and ventilation that would be available for all participants in an epidemiologic cohort and that could be used to explain spatial and temporal variability in residential indoor concentrations of $PM_{2.5}$, EC, and NO_2.
6. Using constrained factor analysis on particle-constituent concentrations and NO_2 measures, determine source types contributing to variability in residential indoor concentrations.
7. Apply LUR modeling techniques to factor scores estimated through constrained factor analysis on residential indoor concentrations, to determine whether GIS covariates and other predictors explain factor variability and thereby support initial factor interpretations of indoor and outdoor source types.
8. Use a simulation framework to assess the degree to which the exposure misclassification induced by the use of a proxy variable for exposure would influence epidemiologic study findings.

METHODS

Data Collection

Participant Selection The primary source of study participants was a prospective birth cohort study of etiologic factors in the development of asthma, the Asthma Coalition for Community, Environment, and Social Stress (ACCESS) study in Boston, Massachusetts. At the time the Levy et al. project started, the ACCESS study was in the process of recruiting pregnant women throughout the metropolitan area. Among several etiologic factors under investigation were indoor and outdoor exposures to air pollutants, including those potentially related to traffic.

Air Quality Measurements Levy and colleagues collected detailed measurements at 43 homes, a subset from the ACCESS study plus others selected to increase sample size and capture additional neighborhoods. These homes were selected to reflect a range of potential exposures to traffic and a set of neighborhoods broadly representative of Boston. From 2003 through 2005, consecutive sets of 3- to 4-day samples of NO_2 and $PM_{2.5}$ were collected simultaneously indoors and outdoors at each home in two seasons, a cooling season defined as May through October, and a heating season defined as December through March. Only a limited number of homes were sampled during each sampling period. $PM_{2.5}$ samples were collected on Teflon filters using a Harvard Personal Environmental Monitor, and NO_2 samples were collected using Yanagisawa passive filter badges. Temperature and humidity measurements were also taken at each home.

The particle filters were analyzed for EC using a reflectance method. To analyze individual elements on the filters, the investigators first used x-ray fluorescence spectroscopy (XRF) and then high-resolution inductively coupled plasma–mass spectrometry (ICP–MS). XRF provides a measure of the total mass of an element on each filter, whereas ICP–MS provides a measure of water-soluble metal concentrations and has a lower limit of detection.

In addition to obtaining sampling data from individual homes, the investigators obtained hourly NO_2, $PM_{2.5}$, EC, and meteorological data for the period of the study from a Massachusetts Department of Environmental Protection (MassDEP) monitor centrally located in relation to the individual sampling sites, as well as from other nearby monitors. NO_2 was measured using a chemiluminescence method, $PM_{2.5}$ was measured using a beta-attenuation method, and EC was measured by optical absorbance. These central site monitoring data were obtained to represent temporal variability, in meteorological conditions and in the background concentrations of pollutants in subsequent statistical analyses, given that sampling at individual homes was conducted during short, discrete time periods.

Later in the investigation, Levy and colleagues obtained supplemental $PM_{2.5}$ samples collected as part of a separate study, to assess whether measures of individual particle components at a single location over a longer period of time might yield different results in their factor analyses. Specifically, they obtained ambient $PM_{2.5}$ and concentrated ambient particle samples collected during a 5-hour period each day in January and February of 2008 using a Harvard Ambient Fine Particle Concentrator located at the Harvard School of Public Health on Huntington Avenue in Boston, not far from a MassDEP central site monitor. These samples were analyzed for the same particulate matter components analyzed in the main study.

Traffic, Occupant Behaviors, and Home Characteristics Data The investigators collected several different types of data, to represent each home's proximity to traffic and the potential density of traffic nearby. For each home, they collected continuous traffic counts on the road with the most traffic within 100 m; they also collected data on road networks and traffic counts from the Massachusetts Highway Department. They used these to create several GIS-based "traffic indicators" — for example, traffic density per unit area surrounding each home, distance from the nearest roads of various sizes, and percentage of diesel traffic — that they might use to predict contributions of traffic to observed pollutant concentrations. The investigators also compiled additional GIS data on population density, a potential proxy for traffic and other sources, within an area around each home.

Using a standardized questionnaire administered to participants at each home, investigators also compiled data on occupant behaviors and home characteristics that had been shown previously to influence pollutant concentrations. They obtained from city authorities data on local land use and on the age of each home, its living area, building materials, heating system, and whether or not it had air conditioning. U.S. Census data were used to determine the number of homes heated with oil within 200 m of each sampled home, because home heating oil is a possible source of air pollutant levels measured.

Statistical Approach and Data Analysis

Outdoor LUR Models The investigators first developed separate multivariate LUR models to predict measured outdoor concentrations of EC, NO_2, and $PM_{2.5}$ at participant homes. These pollutants were selected because they have been studied using GIS methods in other epidemiologic studies. The investigators used a systematic approach to decide which of the many possible covariates — representing traffic, other local combustion sources, population density, home characteristics, or possible influences of local meteorological conditions — to include in the individual models. They began with a statistical assessment of each potential covariate's ability to explain variability in the outdoor concentrations of each pollutant at each home. The final multivariate models for each pollutant were then built through a stepwise regression approach, a process in which the strongest individual covariates were progressively included in the model and tested for their ability to explain additional variability in the model. The investigators chose to emphasize indicators that were more physically interpretable and thus more readily generalizable to other locations.

Indoor LUR Models Using a similar multivariate LUR approach, the investigators sought to predict indoor concentrations of the same three pollutants. In building these models, they emphasized the use of covariates that might be available for all participants from existing resources, rather than the development of new covariates specifically for this study. For example, they used central site monitoring data, publicly available traffic data, and other land-use data rather than the outdoor concentrations measured at individual homes. In the absence of home-specific air-exchange rates, they developed indicators for this factor from models associating indoor/outdoor ratios of sulfur (which has limited indoor sources) with data on housing type and age as well as occupant behaviors like window opening. They also incorporated some study-specific questionnaire data on potential indoor sources of the various pollutants (e.g., gas stoves, cleaning activities).

Source Apportionment Using Constrained Factor Analysis To characterize potential contributions of various sources to outdoor and indoor concentrations, Levy and his colleagues evaluated a number of factor analysis approaches. At the outset of the study, they proposed to use structural equation modeling (SEM) as an approach to allow testing of specific hypotheses about the relationship between observed data (in this case, NO_2 and PM speciation data) and factors representing particular sources (latent variables). However, their preliminary analyses revealed several weaknesses in the SEM approach, given the available data, leading them to consider other approaches. Bayesian approaches were considered but offered little advantage over more traditional approaches. They therefore decided to use a constrained factor analysis approach in which the contributions of specific pollutants to specific factors are constrained to be positive, similar to a nonnegative matrix factorization approach or positive

matrix factorization (Hoyer 2004). The investigators reported that this approach both demonstrated desirable statistical properties and led to more easily interpretable factors.

In this analysis, they focused on a set of sources that they hypothesized to be significant contributors to $PM_{2.5}$ concentrations, including traffic, coal combustion, fuel oil combustion, and long-range transport. They focused on NO_2 and those particle constituents that had been associated with one of these sources in other studies and that were above the limit of detection in 60% or more of the samples from the study. The same set of constituents was included in the indoor factor analysis.

The investigators produced models with two to nine source factors (or latent variables). They evaluated several formal statistical criteria for determining the optimal number of source factors for the outdoor and indoor models, but primarily relied on the Kaiser–Guttman criterion and the Bayesian information criterion, with support from a separate cross-validation exercise. They resolved differences in the factors selected by the two criteria based on their judgment about which factors were the most physically interpretable.

LUR Modeling to Predict Factor Scores The last formal modeling step was to develop LUR models that could predict variability in the outdoor and indoor factor scores, which were estimated using the constrained factor analysis described above. Model development followed a systematic, sequential process similar to that for prediction of $PM_{2.5}$, NO_2, and EC concentrations. The investigators used univariate regression to select candidate variables representing the central site monitor, traffic, population, and other source terms. They built final multivariate models using forward stepwise regression, retaining only those covariates that alone, or in interaction terms, met preset significance criteria. They conducted several sensitivity analyses using alternative selection approaches to ascertain the robustness of the covariate selection process for the final models.

The LUR models for predicting indoor factor scores were developed following a similar process but using questionnaire data and proxy measures for ventilation in addition to outdoor source terms. However, the investigators recognized that it would be difficult to separate indoor from outdoor contributions to the indoor factor scores because, among other reasons, individual constituents might have multiple sources. Consequently, they conducted additional exploratory analyses designed to provide insight into the interpretation of indoor factor analysis scores and the possible contributions from indoor sources.

Exposure Misclassification Analysis The final set of analyses Levy and his colleagues conducted was designed to evaluate to what extent the use of various exposure models, representing different degrees of exposure misclassification, affected the power of an epidemiologic study to detect an adverse health effect. They used a simulation approach to explore the association between pollutant concentrations in the home and wheeze in the first year of life in a hypothetical epidemiologic study of 1000 participants, patterned on the ACCESS data set. The "gold standard" for exposure — an ideal not usually attained in epidemiologic studies in which concentrations are measured in each home — was represented in this analysis by predictions of indoor concentrations of EC, NO_2, and $PM_{2.5}$ at every home using the study's indoor LUR models. The investigators developed three alternative sets of exposure estimates to represent decreasing quality of exposure assessment. The first used the same indoor LUR models for $PM_{2.5}$, NO_2, and EC concentrations, but for a sample of homes. The second used models based on a single outdoor source term or a traffic indicator that had performed well in the covariate selection process ("good exposure surrogates"). The third used models based on traffic indicators that had not performed well in the covariate selection process but that had been used by other investigators ("poor exposure surrogates"). Using these different estimates as measures of exposure to indoor pollution, the investigators then explored the estimated bias, and inflation in the predicted standard error, of the coefficient representing the relationship between indoor pollutant concentrations and wheeze in the first year of life. They conducted simulations under three different scenarios for the strength of the "true" association (i.e., odds ratios of 1.05, 1.50, and 2.00).

Overview of Key Results

LUR Models for Outdoor EC, $PM_{2.5}$, and NO_2 The final multivariate LUR models developed were able to explain 52%, 56%, and 76% of the variability in outdoor residential concentrations of EC, NO_2, and $PM_{2.5}$, respectively, with a combination of temporal and spatial terms (see Table 9 in the Investigators' Report). More of the variability in $PM_{2.5}$ was explained by temporal variability (68% for the term for the central site monitor) than it was for EC (30% for the combination of terms for the central site monitor and season) and NO_2 (33% for the combination of terms for the central site monitor and season). EC and NO_2 had stronger relationships with indicators for local traffic than did $PM_{2.5}$. Roadway length within 200 m of the home accounted for about 14% of the variation in EC, with an additional 8% explained by the interaction between

roadway length and wind speed. Roadway length within 50 m, obstruction between the monitor and nearest major road, and population density collectively accounted for about 23% of the variability in NO_2. Roadway length within 100 m of the home explained only about 2% of the variability in $PM_{2.5}$. The investigators noted that their study corroborated the findings of other studies that spatial and temporal variations in $PM_{2.5}$ concentrations are largely dictated by regional patterns of the pollutant in the eastern United States.

Constrained Factor Analysis and LUR Models for Outdoor Source Factors From their constrained factor analyses, Levy and his colleagues selected a five-factor model representing different sources or latent variables. Factor 1 represented long-range transport; Factor 2, brake wear and local traffic; Factor 3, diesel exhaust; Factor 4, fuel oil combustion; and Factor 5, road dust and resuspension.

The investigators reported that the final multivariate LUR models they developed to predict these outdoor source factors had weaker explanatory power than the LUR models for individual pollutants. That is, the LUR models predicted 69% of the variation in Factor 1, 16% in Factor 2, 32% in Factor 3, 41% in Factor 4, and 20% in Factor 5 (see Table 11 in the Investigators' Report). However, they found that the terms that remained significant in the final models were consistent with their interpretation of the different factors. For example, Factor 1 was most strongly predicted by central $PM_{2.5}$ and season, with very little association with local source terms. Factor 4 was explained by both central site NO_2 and by local measures of population density. The authors concluded that the ability of LUR models combined with constrained factor analysis to disentangle the contributions of long-range transport from local sources could allow future epidemiologic studies to evaluate the effect of certain source categories on health outcomes.

LUR Models for Indoor EC, $PM_{2.5}$, and NO_2 Levy and his colleagues reported less success in the ability of their LUR regression models to predict indoor concentrations of the three pollutants relative to outdoor concentrations of these same pollutants. They reported particular difficulty in identifying traffic terms with strong explanatory power to include in their models. In their initial models, which they constructed without accounting for the influence of ventilation, they found higher indoor NO_2 concentrations to be significantly associated with unweighted cumulative traffic density within 50 m of the home and to a lesser degree with gas stove usage and ambient NO_2 levels. Lower EC concentrations were predicted by increasing distance from a designated truck route; indoor EC concentrations were positively associated with ambient EC concentrations but had no identified indoor sources. Indoor $PM_{2.5}$ was not predicted well by any traffic terms but was associated with cooking and occupant density as well as with central site $PM_{2.5}$ concentrations. The indoor LUR models for NO_2, EC, and $PM_{2.5}$ predicted 20%, 21%, and 36%, respectively, of the variation of the three pollutants (see Table 17 in the Investigators' Report). When ventilation terms were introduced into the models, their explanatory powers increased slightly to 25%, 32%, and 40% (see Table 19 in the Investigators' Report). Levy and colleagues reported substantial variability in indoor–outdoor relationships among particle constituents.

Constrained Factor Analysis and LUR Models for Indoor Source Factors The investigators selected a six-factor model from their constrained factor analyses as the best representation of the range of sources explaining the mixture of pollutants measured indoors in their study. Factor 1 represented long-range transport; Factor 2, fuel oil/diesel combustion; Factor 3, road dust and resuspension; Factor 4, indoor combustion; Factor 5, indoor smoking; and Factor 6, indoor cleaning.

As in the LUR model with outdoor source factors, the model predicting Factor 1, long-range transport, had the highest explanatory power (71% of the variance) with an interaction term involving the central site monitor and a variable for open windows playing the strongest role (68%). The models for the remaining indoor factors had very little overall ability to explain the variability in those factors (18% for Factor 2; 8% for Factor 3; 31% for Factor 4; 24% for Factor 5; and 8% for Factor 6 — see Table 21 in the Investigators' Report). The authors note that these results were not unexpected given the poor correlations of several individual variables observed in the model-development process, and given the limited indoor source covariates available. They conducted several sensitivity analyses designed to further understand the basis for the poor performance of the multivariate models and to explore alternative approaches. These analyses confirmed the general interpretation of certain factors as dominated by indoor sources but also the difficulty of predicting these factors (i.e., Factors 4–6) and of separating the influence of indoor and outdoor sources.

Implications of Exposure-Measurement Error From their simulation analysis, Levy and his colleagues reported that their indoor LUR models for NO_2, $PM_{2.5}$, and to a lesser extent EC resulted in better estimates of the "true" health effect coefficients, compared with the simpler "good" or "poor" proxies for exposure represented by individual variables. That is, they resulted in smaller estimated

biases, smaller median standard errors, and lower root mean square errors (see Tables 22–24 in the Investigators' Report). The power of the LUR models to detect a significant association was thus greater than it was for the simpler exposure models. Not surprisingly, the power of these models was greater for the stronger hypothesized associations (i.e., with odds ratios of 1.5 or 2.0), although this relationship varied by pollutant. The NO_2 and $PM_{2.5}$ models performed much better than did the EC model, which even for an odds ratio of 2.0 predicted a significant relationship in only 27% of the simulations (see Table 25 in the Investigators' Report).

From these simulations, the investigators concluded that epidemiologic studies might benefit from the use of exposure models or surrogates that offer even a small reduction in exposure misclassification relative to that associated with the simple use of models of outdoor concentrations. They cite, for example, their indoor LUR models used in this simulation with modest R^2 values of about 0.3 to 0.4.

Levy and his colleagues conclude that the combination of analytical techniques used in their study could eventually be used to refine characterizations of exposure and to evaluate the relative contribution of particular sources to health outcomes in epidemiologic studies. They suggest that the value of reducing measurement error is particularly evident in the study of exposures to traffic-related air pollutants, which may be influenced by both indoor and outdoor sources.

HEI REVIEW COMMITTEE TECHNICAL EVALUATION

Levy and his colleagues took advantage of a small but rich data source focused on understanding the factors that contribute to the incidence of childhood asthma in a major U.S. city to explore important exposure questions that are of broad interest to environmental health science. They undertook a number of challenging methodologic approaches to improving the prediction of personal exposure from indoor and outdoor sources and thus to improving epidemiologic estimates of the effects of traffic-related air pollution on health. Their report marks one of the first efforts to combine LUR models with factor analysis to assess the spatial aspects of source factors and to use these combined models to more fully characterize exposure to both indoor and outdoor sources.

The overall design of the sampling program was well conceived although it faced somewhat inevitable limitations. Repeated measurements of air pollution at each of the homes provided more data, particularly in the temporal dimension. However, the small number of sampling locations (43) limited the spatial extent of sampling and the range of residence types from which data could be collected. The design created challenges given the presence of temporal and spatial variation in pollutant concentrations. Spatial and temporal contributions to variation were not always formally separated in the work (i.e., some terms reflected both spatial and temporal aspects), making it difficult to know which predominated in explaining variation or to what extent each contributed to the overall performance of a prediction model or clustering of pollutants.

The statistical methods used in the study were many, varied, and often sophisticated, creating challenges for both the investigators and for the reader. However, the motivation for and description of each method used were clear and persuasive that each was a defensible approach to meeting the study objectives. In their systematic and careful approach to the development of their models, and particularly in the factor analysis, the investigators struck a good balance between letting the shape of the data and letting prior knowledge dictate the model form. This balance is often a difficult one to achieve.

LUR MODELING

In earlier LUR models, investigators have primarily relied on spatial factors (e.g., geographic measures of traffic density) to predict pollutant levels, dealing with temporal factors in a separate stage of the modeling, if at all (Jerrett et al. 2005a; Ryan and LeMasters 2007). Because the residential multipollutant sampling approach made simultaneous measurements at numerous sites infeasible, Levy and his colleagues explicitly included both temporal and spatial factors in their models by including central site measurements along with geographic predictors.

Their LUR models of outdoor pollutant levels performed reasonably well, explaining most of the variance in $PM_{2.5}$ and to a lesser extent in NO_2 and EC. Their results essentially confirmed previous findings regarding the relative importance of temporal and spatial factors in predicting the concentrations of different pollutants. That is, they found broader-scale temporal variation, represented by measurements at the central site monitor, to be an important determinant of local $PM_{2.5}$ levels. Spatially distributed factors, such as traffic, population density, and other land-use covariates, were more influential in predicting NO_2 and EC variation in the models. The study provided a clear demonstration of these relationships.

Nonetheless, the HEI Review Committee commented that the LUR models used to predict spatial distribution of the outdoor concentrations would have benefited from more direct evaluation and validation. That is, the investigators relied heavily on the assessment of the model fit in their evaluations rather than on a "real-world" check of the resulting spatial patterns of pollutants predicted by the model. The Committee suggested that producing visual spatial surfaces that could be evaluated against the location of busy freeways, point sources of emissions, or influential weather patterns might be informative in modeling exercises of this kind.

The development of LUR models for predicting indoor concentrations was a thoughtful innovation beyond standard LUR models that have focused on outdoor concentrations. However, developing indoor LUR models that would provide better estimates of individual-level exposure proved to be a much greater challenge than expected. As the investigators acknowledge, the predictive value of the indoor LUR models was generally poor; however, their exploration of the possible explanations and implications of this finding is thorough and informative. In particular, their findings that the LUR models' performance was poorer when important indoor sources were present should be highlighted. The related point that — when indoor sources are minor or nonexistent — the predictive value of an indoor LUR model built on outdoor source terms can be improved by the relatively straightforward addition of a proxy term for ventilation (i.e., open windows) is also noteworthy.

APPLICATION OF LUR TO PREDICTION OF OUTDOOR AND INDOOR SOURCE FACTORS

An ultimate, and a most innovative, goal of the study was to see if LUR and factor analysis together would provide new insight about the sources contributing to outdoor and indoor concentrations of pollutants and help explain why their contributions might differ at individual homes. The Committee thought the constrained factor analysis approach decided upon by the investigators was technically well justified and that they had given plausible interpretations of the resulting outdoor and indoor factor scores. The Committee noted, however, that it is important to bear in mind that these interpretations were largely post hoc given the nature of the analysis, that many of the individual constituents were highly skewed, and that consequently other plausible interpretations might exist. As one can see in the dendrograms provided in the Investigators' Report (Figures 4 and 8), individual pollutants and PM constituents identified a priori as markers for particular sources, and thus influential in factor interpretations, also contributed to other factors.

The Committee commented that the application of LUR to predict the outdoor and indoor factor scores was also very carefully conducted, but that the analyses had met with limited success. The analyses did provide some confirmation for the investigators' major source interpretations (long-range transport, traffic, fuel oil combustion). However, they were most successful at explaining variation in sources that are already reasonably well understood. In particular, both the outdoor and indoor LUR models performed best at predicting variability in the source factor identified most closely with $PM_{2.5}$ concentrations at the central site monitor, long-range transport. As for the indoor LUR models developed to predict individual pollutant concentrations, incorporating a proxy for ventilation improved the performance of the indoor LUR models designed to predict variation in the source category of long-range transport. The reasons behind the poorer performance of the LUR models at predicting sources relative to their performance at predicting individual pollutants are difficult to pinpoint. However, the small sample sizes in the study, the complications inherent in predicting indoor sources, and the difficulty in identifying distinct "fingerprints" for different source types noted above all likely contributed to this finding.

ANALYSES OF THE IMPACT OF EXPOSURE MISCLASSIFICATION

The Review Committee thought that the simulation studies of the impact of exposure misclassification, represented by the use of different exposure metrics, on the power of epidemiologic studies were well done and provided a useful perspective. As more sophisticated approaches for estimating spatial variability in concentrations of urban air pollution have been developed, interest in applying them in epidemiologic studies has grown (Jerrett et al. 2005b; Brauer et al. 2007; Slama et al. 2007; Rosenlund et al. 2008). At the time the Levy study was initiated, however, few epidemiologic studies had been conducted in which the impact of different exposure metrics on estimates of human health effects had been directly compared in the same study population (Ryan et al. 2007). Some studies exploring the implications of different measures of traffic-related exposures for health effect estimates have since begun to emerge and have suggested the importance of improving the estimates of individual-level exposures (Jerrett et al. 2005b; Van Roosbroeck et al. 2008).

In the absence of direct comparisons of different exposure metrics in epidemiologic studies, simulation analyses like those conducted by Levy and his colleagues could

provide useful insights that could inform the design of exposure assessments. Another recent simulation-based study has suggested that augmenting estimates of individual-level exposure with information about personal mobility patterns is less likely to underestimate health effects than basing exposures on residential address alone (Setton et al. 2010). Such studies allow exploration of how different exposure-assessment designs might affect the power to detect hypothesized health outcomes. The value of more complex exposure assessments can theoretically then be more directly weighed against the cost of the data necessary to support them.

The general pattern of results from the simulation analyses conducted by Levy and his colleagues was consistent with previous findings on the nature of error in exposure estimates; that is, that the error is a combination of Berkson-type error and classical statistical error. The loss of precision and bias downward in the effect estimates is very evident in the results for the simpler models relative to the "gold standard" exposure measurements. The question of the degree of differential exposure error in the outdoor LUR and indoor LUR models relative to "true" personal exposures remains unanswered by this study.

Despite the better performance of the LUR models relative to that of the simpler exposure proxies in this study, it is important to be aware that these quantitative differences depend on the underlying simulation scenario and therefore may not apply to epidemiologic studies with fundamentally different designs. Where these conditions are not met — that is, where the dominant influences on the variation in a pollutant concentration are not well known — LUR models are likely to perform much worse, and in some cases little better, than a surrogate. Therefore, the results of these simulations should not be assumed to apply to all situations. Nonetheless, the HEI Review Committee concluded that simulation analyses should be encouraged to the extent that they can provide insights beyond what general theory might suggest or the assumptions embedded in a particular study design.

CONCLUSIONS AND IMPLICATIONS

Levy and his colleagues undertook a set of challenging and creative methodologic approaches to answering important questions related to environmental health. They sought to build on existing LUR techniques for estimating individual-level exposures to $PM_{2.5}$, EC, and NO_2 for use in epidemiologic studies, in particular by developing LUR models to predict the concentrations of the pollutants indoors as well as outdoors. They combined LUR techniques with factor analysis of $PM_{2.5}$ components and other pollutants in an effort to improve understanding of the potential contribution of different sources, including traffic, to human exposures.

The HEI Review Committee praised the evident care and competence demonstrated by Levy and his team. The investigators' LUR analyses of outdoor pollutant concentrations were consistent with previously published findings, performing reasonably well at explaining variation in concentrations of $PM_{2.5}$ and to a lesser extent in NO_2 and EC. Their efforts to develop LUR models to predict indoor concentrations, as a closer proxy for personal exposure, proved a greater challenge. The predictive value of these indoor models was generally poor, although the authors did offer some useful insights about the variables that most influenced the performance of the models. The limited predictive value of the LUR models developed to explain variation in the sources identified in the factor analyses, particularly of the indoor models, seemed to be attributable more to limitations of the data than to the methods. However, whether these methods would perform better with a less limited set of data remains to be demonstrated.

The Committee found the results of this study to be relevant to epidemiologic studies of air pollution. However, the Committee thought that the conclusions would have been more useful had the authors translated the findings into more explicit, practical suggestions for the design of questionnaires, exposure monitoring programs, and epidemiologic analyses in future studies.

At the same time, the Committee noted that the generalizability of the findings requires careful caveats. The spatiotemporal sampling design for this study creates particular challenges for the application of the investigators' approach to other epidemiologic studies. Some of the variation in exposures is temporal and some spatial. Most epidemiologic studies are likely to be informed by variation in one or the other dimension only. Even those epidemiologic studies using spatiotemporal designs are unlikely to have quite the same configuration (e.g., number of days in observation periods, extent of repeated sampling). Therefore, the study cannot provide a generalized evaluation of LUR models in other situations. The ultimate challenge for studies of this nature is to provide some demonstration that the increased sophistication of the modeling provides sufficient improvement over simpler approaches to warrant the additional data and computational requirements it imposes. The authors' simulation analyses, in which they explore the implications of different exposure proxies and estimates for exposure misclassification and for the power of epidemiologic studies, are a useful step in that direction. Their conclusion that even the relatively poor estimates of exposure provided by the LUR models might reduce measurement error and thus improve effects estimates in future

studies warrants further scrutiny. Even when a poor surrogate is outperformed by a prediction model, the surrogate may be the epidemiologist's "best buy" if the extent of improved performance is outweighed by the costs of collecting the predictor data necessary for the prediction model.

ACKNOWLEDGMENTS

The Health Review Committee thanks the ad hoc reviewers for their help in evaluating the scientific merit of the Investigators' Report. The Committee is also grateful to Annemoon Van Erp for her oversight of the study, to Katherine Walker for her assistance in preparing its Critique, to Leah Shriro for science editing of this Report and its Critique, and to Fred Howe, Bernard Jacobson, Flannery McDermott, and Ruth Shaw for their roles in preparing this Research Report for publication.

REFERENCES

Armstrong BG. 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup Environ Med 55:651–656.

Brauer M, Hoek G, Smit HA, de Jongste JC, Gerritsen J, Postma DS, Kerkhof M, Brunekreef B. 2007. Air pollution and development of asthma, allergy and infections in a birth cohort. Eur Respir J 29:879–888.

Brunekreef B, Janssen NAH, de Hartog JJ, Oldenwening M, Meliefste K, Hoek G, Lanki T, Tomonen KL, Vallius M, Pekkhanen J, Van Grieken R. 2005. Personal, Indoor, and Outdoor Exposures to $PM_{2.5}$ and Its Components for Groups of Cardiovascular Patients in Amsterdam and Helsinki. Research Report 127. Health Effects Institute, Boston, MA.

Dominici F, Zeger SL, Samet JM. 2000. A measurement error model for time-series studies of air pollution and mortality. Biostatistics 1:157–175.

Gehring U, Cyrys J, Sedlmeir G, Brunekreef B, Bellander T, Fischer P, Bauer CP, Reinhardt D, Wichmann HE, Heinrich J. 2002. Traffic-related air pollution and respiratory health during the first 2 yrs of life. Eur Respir J 19: 690–698.

HEI Panel on the Health Effects of Traffic-Related Air Pollution. 2010. Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. Special Report 17. Health Effects Institute, Boston, MA.

Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, Briggs D. 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos Environ 42: 7561–7578.

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

Hoek G, Fischer P, van den Brandt P, Goldbohm S, Brunekreef B. 2001. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. J Expo Anal Environ Epidemiol 11:459–469.

Hoyer P. 2004. Non-negative matrix factorization with sparseness constraints. J Machine Learning Research 5:1457–1469.

Janssen NAH, van Vliet PHN, Aarts F, Harssema H, Brunekreef B. 2001. Assessment of exposure to traffic related air pollution of children attending schools near motorways. Atmos Environ 35:3875–3884.

Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, Morrison J, Glovis C. 2005a. A review and evaluation of intraurban air pollution exposure models. J Exposure Anal Environ Epidemiol 15:185–204.

Jerrett M, Burnett RT, Ma R, Pope CA III, Krewski D, Newbold KB, Thurston G, Shi Y, Finkelstein N, Calle EE, Thun MJ. 2005b. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16:727–736.

Koutrakis P, Suh HH, Sarnat JA, Brown KW, Coull BA, Schwartz J. 2005. Characterization of particulate and gas exposures of sensitive subpopulations living in Baltimore and Boston. Research Report 131. Health Effects Institute, Boston, MA.

Krämer U, Koch T, Ranft U, Ring J, Behrendt H. 2000. Traffic-related air pollution is associated with atopy in children living in urban areas. Epidemiology 11:64–70.

Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi Y, Turner MC, Pope CA III, Thurston G, Calle EE, Thun MJ. 2009. Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality. HEI Research Report 140. Health Effects Institute, Boston, MA.

Laden F, Neas LM, Dockery DW, Schwartz J. 2000. Association of fine particulate matter from different sources with daily mortality in six U.S. cities. Environ Health Perspect 108:941–947.

Lebret E, Briggs D, van Reeuwijk H, Fischer P, Smallbone K, Harssema H, Kriz B, Gorynski P, Elliot P. 2000. Small area variations in ambient NO_2 concentrations in four European areas. Atmos Environ 34:177–185.

Leech JA, Nelson WC, Burnett RT, Aaron S, Raizenne ME. 2002. It's about time: A comparison of Canadian and American time–activity patterns. J Expo Anal Environ Epidemiol 12:427–432.

Levy JI, Lee K, Spengler JD, Yanagisawa Y. 1998. Impact of residential nitrogen dioxide exposure on personal exposure: An international study. J Air Waste Manag Assoc 48:553–560.

Mar TF, Norris GA, Koenig JQ, Larson TV. 2000. Associations between air pollution and mortality in Phoenix, 1995–1997. Environ Health Perspect 108:347–353.

Rosenlund M, Picciotto S, Forastiere F, Stafoggia M, Perucci CA. 2008. Traffic-related air pollution in relation to incidence and prognosis of coronary heart disease. Epidemiology 19:121–128.

Ryan PH, LeMasters GK. 2007. A review of land-use regression for characterizing intraurban air models pollution exposure. Inhalation Toxicol 19:127–133.

Ryan PH, LeMasters GK, Biswas P, Levin L, Hu S, Lindsey M, Bernstein DI, Lockey J, Villareal M, Khurana Hershey GK, Grinshspun SA. 2007. A comparison of proximity and land use regression traffic exposure models and wheezing in infants. Environ Health Perspect 115:278–284.

Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. 2000. Part I. Methods and methodologic issues. In: The National Morbidity, Mortality, and Air Pollution Study. Research Report 94. Health Effects Institute, Cambridge, MA.

Setton E, Marshall JD, Brauer M, Lundquist KR, Hystad P, Keller P, Cloutier-Fisher D. 2010. The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. J Expo Sci Environ Epidemiol, Epub ahead of print.

Slama R, Morgenstern V, Cyrys J, Zutavern A, Herbarth O, Wichmann H-E, Heinrich J; LISA Study Group. 2007. Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: A study relying on a land-use regression exposure model. Environ Health Perspect 115:1283–1292.

Turpin B, Weisel CP, Morandi M, Colome S, Stock T, Eisenreich S, Buckley B, and Others. 2007. Part II. Analyses of concentrations of particulate matter species. In: Relationships of Indoor, Outdoor, and Personal Air (RIOPA). Research Report 130. Health Effects Institute, Boston, MA.

Van Roosbroeck S, Li R, Hoek G, Lebret E, Brunekreef B, Spiegelman D. 2008. Traffic-related outdoor air pollution and respiratory symptoms in children: The impact of adjustment for exposure measurement error. Epidemiology 19:409–416.

Veen AVD, Briggs DJ, Collins S, Elliott P, Fischer P, Kingham S, Lebret E, Pryl K, Reeuwijk HV, Smallbone K. 1997. Mapping urban air pollution using GIS: A regression-based approach. Int J Geo Info Sci 11: 699–718.

Weisel CP, Zhang J, Turpin BJ, Morandi MT, Colome S, Stock TH, Spektor DM, and Others. 2005. Part I. Collection methods and descriptive analyses. In: Relationships of Indoor, Outdoor, and Personal Air (RIOPA). Research Report 130. Health Effects Institute, Boston, MA.

Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery D, Cohen A. 2000. Exposure measurement error in time-series studies of air pollution: Concepts and consequences. Environ Health Perspect 108:419–426.

Zipprich JL, Harris SA, Fox JC, Borzelleca JF. 2002. An analysis of factors that influence personal exposure to nitrogen oxides in residents of Richmond, Virginia. J Expo Anal Environ Epidemiol 12:273–285.

RELATED HEI PUBLICATIONS: PARTICULATE MATTER AND TRAFFIC

Number	Title	Principal Investigator	Date*
Research Reports			
140	Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality	D. Krewski	2009
139	Effects of Long-Term Exposure to Traffic-Related Air Pollution on Respiratory and Cardiovascular Mortality in the Netherlands: the NLCS-AIR Study	B. Brunekreef	2009
138	Health Effects of Real-World Exposure to Diesel Exhaust in Persons with Asthma	J. Zhang	2009
137	The Influence of Improved Air Quality on Mortality Risks in Erfurt, Germany	A. Peters	2009
133	Characterization of Metals Emitted from Motor Vehicles	J.J. Schauer	2006
131	Characterization of Particulate and Gas Exposures of Sensitive Subpopulations Living in Baltimore and Boston	P. Koutrakis	2005
130	Relationships of Indoor, Outdoor, and Personal Air (RIOPA)		
	Part I. Collection Methods and Descriptive Analyses	C.P. Weisel	2005
	Part II. Analyses of Concentrations of Particulate Matter Species	B.J. Turpin	2007
127	Personal, Indoor, and Outdoor Exposures to $PM_{2.5}$ and Its Components for Groups of Cardiovascular Patients in Amsterdam and Helsinki	B. Brunekreef	2005
124	Particulate Air Pollution and Nonfatal Cardiac Events		2005
	Part I. Air Pollution, Personal Activities, and Onset of Myocardial Infarction in a Case–Crossover Study	A. Peters	
	Part II. Association of Air Pollution with Confirmed Arrhythmias Recorded by Implanted Defibrillators	D. Dockery	
118	Controlled Exposures of Healthy and Asthmatic Volunteers to Concentrated Ambient Particles in Metropolitan Los Angeles	H. Gong Jr.	2003
95	Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations	M. Lippmann	2000
94	National Morbidity, Mortality, and Air Pollution Study		2000
	Part I. Methods and Methodologic Issues	J.M. Samet	

Continued

** Reports published since 1998.*
Copies of these reports can be obtained from the Health Effects Institute and many are available at www.healtheffects.org.

RELATED HEI PUBLICATIONS: PARTICULATE MATTER AND TRAFFIC

Number	Title	Principal Investigator	Date*
Special Reports			
17	Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects		2010
16	Mobile-Source Air Toxics: A Critical Review of the Literature on Exposure and Health Effects		2007
HEI Communications			
10	Improving Estimates of Diesel and Other Emissions for Epidemiologic Studies		2003
HEI Research Program Summaries			
	Research on Diesel Exhaust and Other Particles		2003
HEI Perspectives			
	Understanding the Health Effects of Components of the Particulate Matter Mix: Progress and Next Steps		2002

** Reports published since 1998.*
Copies of these reports can be obtained from the Health Effects Institute and many are available at www.healtheffects.org.

HEI BOARD, COMMITTEES, and STAFF

Board of Directors

Richard F. Celeste, Chair *President, Colorado College*

Sherwood Boehlert *Of Counsel, Accord Group; Former Chair, U.S. House of Representatives Science Committee*

Enriqueta Bond *President Emeritus, Burroughs Wellcome Fund*

Purnell W. Choppin *President Emeritus, Howard Hughes Medical Institute*

Michael T. Clegg *Professor of Biological Sciences, University of California–Irvine*

Jared L. Cohon *President, Carnegie Mellon University*

Stephen Corman *President, Corman Enterprises*

Gowher Rizvi *Vice Provost of International Programs, University of Virginia*

Linda Rosenstock *Dean, School of Public Health, University of California–Los Angeles*

Henry Schacht *Managing Director, Warburg Pincus; Former Chairman and Chief Executive Officer, Lucent Technologies*

Warren M. Washington *Senior Scientist, National Center for Atmospheric Research; Former Chair, National Science Board*

Archibald Cox, Founding Chair *1980–2001*

Donald Kennedy, Vice Chair Emeritus *Editor-in-Chief Emeritus,* Science*; President Emeritus and Bing Professor of Biological Sciences, Stanford University*

Health Research Committee

David L. Eaton, Chair *Associate Vice Provost for Research and Director, Center for Ecogenetics and Environmental Health, School of Public Health, University of Washington–Seattle*

David T. Allen *Gertz Regents Professor in Chemical Engineering; Director, Center for Energy and Environmental Resources, University of Texas–Austin*

Joe G.N. Garcia *Vice Chancellor for Research and Professor of Medicine, University of Illinois–Chicago*

Uwe Heinrich *Professor, Medical School Hannover, Executive Director, Fraunhofer Institute for Toxicology and Experimental Medicine, Hannover, Germany*

Grace LeMasters *Professor of Epidemiology and Environmental Health, University of Cincinnati College of Medicine*

Sylvia Richardson *Professor of Biostatistics, Department of Epidemiology and Public Health, Imperial College School of Medicine, London, United Kingdom*

Richard L. Smith *Director, Statistical and Applied Mathematical Sciences Institute, University of North Carolina–Chapel Hill*

James A. Swenberg *Kenan Distinguished Professor of Environmental Sciences, Department of Environmental Sciences and Engineering, University of North Carolina–Chapel Hill*

Ira B. Tager *Professor of Epidemiology, School of Public Health, University of California–Berkeley*

HEI BOARD, COMMITTEES, and STAFF

Health Review Committee

Homer A. Boushey, Chair *Professor of Medicine, Department of Medicine, University of California–San Francisco*

Ben Armstrong *Reader in Epidemiological Statistics, Public and Environmental Health Research Unit, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, United Kingdom*

Michael Brauer *Professor, School of Environmental Health, University of British Columbia, Canada*

Bert Brunekreef *Professor of Environmental Epidemiology, Institute of Risk Assessment Sciences, University of Utrecht, the Netherlands*

Mark W. Frampton *Professor of Medicine and Environmental Medicine, University of Rochester Medical Center*

Stephanie London *Senior Investigator, Epidemiology Branch, National Institute of Environmental Health Sciences*

Armistead Russell *Georgia Power Distinguished Professor of Environmental Engineering, School of Civil and Environmental Engineering, Georgia Institute of Technology*

Lianne Sheppard *Professor of Biostatistics, School of Public Health, University of Washington–Seattle*

Officers and Staff

Daniel S. Greenbaum *President*
Robert M. O'Keefe *Vice President*
Rashid Shaikh *Director of Science*
Barbara Gale *Director of Publications*
Jacqueline C. Rutledge *Director of Finance and Administration*
Helen I. Dooley *Corporate Secretary*

Kate Adams *Staff Scientist*
Aaron J. Cohen *Principal Scientist*
Maria G. Costantini *Principal Scientist*
Philip J. DeMarco *Compliance Manager*
Suzanne Gabriel *Editorial Assistant*
Hope Green *Editorial Assistant (part time)*
L. Virgi Hepner *Senior Science Editor*
Anny Luu *Administrative Assistant*
Francine Marmenout *Senior Executive Assistant*
Sumi Mehta *Senior Scientist*
Nicholas Moustakas *Policy Associate*
Hilary Selby Polk *Senior Science Editor*
Sarah Rakow *Science Administrative Assistant*
Robert A. Shavers *Operations Manager*
Geoffrey H. Sunshine *Senior Scientist*
Annemoon M.M. van Erp *Senior Scientist*
Katherine Walker *Senior Scientist*
Morgan Younkin *Research Assistant*

HEALTH
EFFECTS
INSTITUTE

101 Federal Street, Suite 500
Boston, MA 02110, USA
+1-617-488-2300
www.healtheffects.org

RESEARCH
REPORT

Number 152
December 2010