OFFICE OF APPLIED STUDIES

Substance Abuse in States and Metropolitan Areas: Model Based Estimates from the 1991-1993 National Household Surveys on Drug Abuse: Methodology Report

by:

Ralph E. Folsom
David R. Judkins
Research Triangle Institute

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Substance Abuse and Mental Health Services Administration

Acknowledgments

This publication was developed by the Research Triangle Institute (RTI) for the Substance Abuse and Mental Health Services Administration, Office of Applied Studies (SAMHSA/OAS) under Contract No. 283-91-5402. Ralph Folsom was responsible for the development of the methodology, initial drafts, and final oversight and review. Mike Witt developed and implemented the computer programs. David Judkins wrote the final draft, explaining the methodology in terms of basic statistical concepts throughout the report, expanding the content of the chapter on evaluation, and providing further derivations in Appendices C, G and H. Joseph Gfroerer and Douglas Wright (OAS) guided the overall development and documentation of the project, reviewed all drafts, and contributed to various chapters. Tom Virag was the RTI project director, and Joseph Gustin was the OAS project officer.

Obtaining Additional Copies of Publication

For additional copies, please write or call:

National Clearinghouse for Alcohol and Drug Information (NCADI)
P.O. Box 2345
Rockville, MD 20847-2345
(301) 468-2600 1-800-729-6686TDD 1-800-487-4889

Electronic Access to Publication

This publication can be accessed electronically through the Internet World Wide Web connections listed below.

World Wide Web Internet Connections: http://www.samhsa.gov
http://www.health.org

Originating Office:
SAMHSA, Office of Applied Studies
5600 Fishers Lane
Rockville, Maryland 20857

June 1997

Table of Contents

Page

Table of Contents
(continued)

Page

List of Exhibits

Page

1. Introduction

In 1996, the Substance Abuse and Mental Health Services Administration (SAMHSA) published a report on substance abuse for 26 States and 25 metropolitan statistical areas (MSAs).[1] That report did not provide much detail on the methodology used to produce those estimates. The report you are now reading has been written to serve two purposes. The first is to give policy analysts a deeper understanding of the properties of the estimates and of the confidence intervals on the estimates. The second is to allow other researchers to replicate the approach and to apply the method to other tasks.

The impetus for developing the local area statistics was based on the increasing demand for such estimates by State and Federal substance abuse planners. While there are existing data systems that provide small area indicators of drug abuse for selected MSAs, like the emergency room drug-related episode reports from the Drug Abuse Warning Network (DAWN) and the voluntary urine test reports among arrestees (DUF), these data systems do not provide estimates of the desired total population statistics on use, dependence, treatment and the need for treatment. Local area surveys designed to provide reliable direct estimates tend to be costly and generally do not provide comparable estimates between areas due to different sampling and data collection methodologies.

The National Household Survey on Drug Abuse (NHSDA) has been the federal government's chief source of information on the magnitude of substance use and abuse in the United States and has been used by policy makers to monitor trends in drug use and to identify problem areas. In addition to assessing prevalence and trends in drug use, SAMHSA uses the NHSDA data to investigate many special topics that are important for understanding the nature of substance abuse, and numerous researchers from around the nation conduct special analyses using the public use files that are produced from the survey. Reports based on the NHSDA are highly valued due to the use in the survey of a rigorous sampling and data collection methodology that has been developed and tested over a number of years.

Despite the value of the NHSDA for national statistics, it is not by itself an adequate source of information about drug abuse at the state and MSA levels. The local area statistics that were published in 1996 rely upon a complex algorithm for combining data from the NHSDA with local data from the Decennial Census and other sources. The algorithm is justified

[1] SAMHSA, 1996.

through the use of models that relate the available local data to drug abuse as measured in the NHSDA. The use of models was required since national surveys like the NHSDA do not provide large enough samples for direct State level estimation, a limitation that persists even after pooling three years of NHSDA data together. The model-based approach to deriving the small area estimates is similar to approaches that have been used by other government agencies as the demand for local area data has increased and the required statistical and computational tools have improved.

The algorithm that was used to combine NHSDA and local data to produce small area estimates is new. It was developed by a contractor, the Research Triangle Institute. The earlier summary report that provided the estimates also had a simplified description of the algorithm and a report on an evaluation of the algorithm, but that discussion was not detailed enough to allow other researchers to replicate the method. This new report provides a detailed explanation of the methods. Although most of this report is quite technical, the introduction was written for a broad audience of social science researchers and policy analysts.

This report has four main chapters and several appendices. The balance of the introduction provides background on the NHSDA, a description of the variables and areas for which estimates were formed, a review of the need for special methodology different from that used to produce national estimates for the NHSDA and most other federal demographic surveys, a review of other methodologies that have been developed in the past for small areas such as states and MSAs and the need for a new methodology, a discussion of evaluation strategies, and guidance on interpretation of the state and MSA estimates.

Chapter 2 explains the methodology in detail. The material in this chapter is quite technical. Readers will need to be comfortable with calculus, linear algebra, and statistics. However, readers do not need to be national experts on the subject of small area estimation to appreciate this chapter. Specifically, it does not assume past experience on small area estimation.

Chapter 3 discusses the evaluation strategy developed for this project and presents the results. This chapter largely repeats the material in the summary report (SAMHSA, 1996), but there is some new material.

Chapter 4 discusses the lessons learned over the course of this project and ideas that the researchers will pursue in future applications of the methodology.

The appendices provide detail on the classification of drug and alcohol dependence, the definitions of the MSAs involved in the study, more technical information on the software used to create the estimates, more information on geographic predictors of drug and alcohol abuse, tables of actual estimates for the states and MSAs, and tables of confidence intervals on the estimates.

1.1 Background on the NHSDA

In this section, the basic design of the NHSDA is reviewed: the sample design, the questionnaire, interviewing procedures, and the data collection schedule.

1.1.1 Sample Design

The respondent universe for the NHSDA is the civilian noninstitutionalized population 12 and older within the United States. Since 1991, the NHSDA has included residents of noninstitutional group quarters (e.g., shelters, rooming houses, dormitories), residents of Alaska and Hawaii, and civilians living on military bases. Persons excluded from the universe include those with no fixed household address, e.g., homeless transients not in shelters, and residents of institutional group quarters such as jails and hospitals.

The sample design is a stratified, multistage probability design. The first stage of sampling involves selection of primary sampling units (PSUs) located across the United States. These PSUs are either very large MSAs or individual counties. Then, within each PSU, smaller land areas, or segments, are selected. Once the segments are selected, professional field staff list all of the dwelling units in each selected segment. From the resultant listings, specific dwelling units are selected for screening and/or interviewing. The NHSDA uses an area sampling frame based on block-level geographical units defined by the Decennial Census. Dwelling unit undercoverage bias is largely eliminated by using the half-open interval technique, which asks questions of screened dwelling units to identify unlisted units.

The NHSDA oversamples blacks and Hispanics to ensure an adequate representation of these minority populations in the sample. Young people are also oversampled. From 1991 to 1993, six major metropolitan areas—New York City, Washington DC, Miami, Chicago, Denver and Los Angeles—were oversampled to provide prevalence estimates for those areas. Due to the association between smoking and illicit drug use, oversampling of current smokers aged 18 to 34 was initiated for the 1993 NHSDA.

Exhibit 1.1 summarizes the sample design/size and response rate data for the NHSDA between 1991 and 1993, the years upon which the small area estimates were based.

Exhibit 1.1 NHSDA Sample Size and Response Rate by Year

	1991	1992	1993
# PSUs	125	118	117
# Segments	3,509	3,218	3,124
# DUs screened	105,311	69,996	100,340
Screener response rate	95.0%	93.6%	94.0%
# Interviews	32,594	28,832	26,489
Interview response rate among screened eligibles	84.3%	82.5%	79.5%

1.1.2 Interviewing Procedures

Prior to the interviewer's arrival at the sample dwelling unit (SDU) a letter was mailed to the resident briefly explaining the survey and requesting their cooperation. Upon arrival at the SDU a few days later, the interviewer referred the respondent to this letter and answered any questions. If the respondent had no knowledge of the lead letter, the interviewer provided another copy, explained that one was previously sent, and then answered any questions. If no one was at home during the initial call at the SDU, the interviewer left a "Sorry I Missed You" card alerting the SDU that the interviewer planned to make another callback at a later date/time. Interviewers made at least four callbacks (in addition to the initial call) to each SDU to complete the screening process and possibly obtain an interview. As necessary and appropriate, the interviewer could make use of the Appointment Card for scheduled return visits with the respondent. When an in-person contact was made with an adult member of the SDU and introductory procedures were completed, the interviewer presented a "Statement of Confidentiality" and answered questions if required. Assuming respondent cooperation, a screening of the SDU was then initiated through administration of the Housing Unit Screening Form for housing units, or the Group Quarters Unit Screening Form for group quarters units.

If a potential respondent refused to be screened, the interviewer was trained to accept the refusal in a positive manner, thereby avoiding the possibility of creating an adversarial relationship and precluding future opportunities for conversion. A conversion was usually attempted by supervisory field staff or specially selected veteran interviewers with established conversion records. If the respondent proceeded with the screening process, the interviewer answered any questions that the screening respondent may have had concerning the study. A number of informational handouts also may have been given to the respondent at this time.

Interviewers screened each sample dwelling unit within a sample segment. The interviewer listed all dwelling unit members on the screening form in order of age, beginning with the oldest. The dwelling unit type was determined based upon the race/ethnicity of the head of household, and the age groups represented in the household. Once the dwelling unit type was determined, the interviewer referred to a selection table that indicates which (if any) age groups were to be interviewed. None, one or two persons may have been selected for interview. The decision was made to limit the number of interviewed persons per household in order to avoid household fatigue, thereby keeping household response rates high.

Assuming the within dwelling unit sampling process selected a member for participation in the study, and if the selected individual was 18 or older and currently available to do the specified study questionnaire, the interviewer moved immediately to begin administering the questionnaire in a private setting within the dwelling unit. If the selected individual was 12 to 17 years of age, parental consent was obtained from the selected individual's parent or legal guardian; then, the minor was asked to participate. Once consent was obtained, the interviewer began the interview process.

The questionnaire and interview methods were designed to retain respondent interest, ensure confidentiality, and maximize the validity of response. The questionnaire was administered in such a way that interviewers would not know respondents' answers to the questions on substance use. Except for tobacco, questions for all substances were self-administered, as respondents recorded their answers on answer sheets. If a respondent was unable to read the questions or preferred not to read them, the interviewer read them out loud. However, only under circumstances of the respondent's physical inability could the interviewer record the answers on the answer sheets. Separate English and Spanish version of the questionnaire were available for each round of the NHSDA.

As the respondent recorded his/her answer choices and completed each answer sheet, they were placed in an envelope. At the end of the interview process, all materials (screening form, study questionnaire, and answer sheets) were enclosed in this envelope, sealed, and mailed to the data processing site.

For verification purposes, respondents were asked to complete a Verification Form that requested their address and telephone number for possible follow-up to ensure that the interviewer did his/her job appropriately. Respondents were apprised that completion of the Verification Form was voluntary, and they were given the opportunity to decline to complete the verification form. These forms were then placed in a separate envelope and mailed to the data processing site. Respondents were invited to travel with the interviewer to the nearest mailbox to verify that the envelopes were immediately mailed.

A random sample of those who completed verification forms received a telephone call expressing appreciation for their participation in the study. Each respondent also was asked to answer a few questions verifying that the interview took place, whether the answer sheets were used properly, and the amount of time required to administer the interviews. Mail-out of verification letters was used when telephone numbers were unavailable.

1.1.3 Data Collection Schedule

Data collection for the 1991 NHSDA took place over a 6 month period from January to June of that year. In 1992, the NHSDA began to operate on a quarterly data collection schedule. That is, data collection extends over the entire calendar year, and is divided into four quarterly components based on the four calendar quarters. This quarterly data collection schedule allows for the realization of several associated benefits and efficiencies, such as increased flexibility in the analysis of trends and in the ability to produce data suitable for public release. Furthermore, such a schedule reduces the burden on the field staff by spreading the data collection over a longer period of time, thus allowing procedures from year to year to be implemented on more of an on-going rather than "start-stop" basis.

1.2 Selection of Variables and Areas for Estimation

It was not feasible to develop estimates of all substance abuse rates. It was necessary to pick a few behaviors. Also, there was a reluctance to prepare estimates for areas with very little NHSDA sample in them. In this section, the selected variables and areas are documented, and the choices are explained.

1.2.1 Selection of Behavior Measures

The NHSDA collects information on many aspects of substance usage, dependence, and treatment along with background demographic variables that can help provide context. Due to the cost of fitting the area estimation models, it was not feasible to produce estimates for all behaviors of interest. Accordingly, SAMHSA chose 11 critical behavior measures for which state and MSA estimates would be produced.

Alcohol and cigarette use were chosen because of their importance for public health and because independent estimates of the use of these substances are available at the state level for comparison. Three measures of usage of illicit substances were chosen, one broad measure that covers all illicit substances, a narrower measure that excludes marijuana, and a measure that focuses exclusively on cocaine.

Two measures of dependence were chosen: a measure of dependence on illicit drugs with or without dependence on alcohol and a measure of dependence only on alcohol without any dependence on illicit drugs. Persons classified as dependent are among the most severely affected substance abusers who are likely to require some kind of treatment or intervention. The dependence measure developed for the NHSDA is based on an algorithm that approximates the DSM-III-R criteria (Appendix A).

Three measures related to treatment were chosen. Two are based on questions in the NHSDA questionnaire that ask about the receipt of treatment. One covers treatment for illicit drugs (with or without treatment for alcohol abuse), and the other covers treatment for alcohol abuse (without treatment for illicit drugs). A measure of need for treatment for illicit drugs was also chosen that is based on an algorithm developed by SAMHSA's Office of Applied Studies. By subtracting the count of those actually reporting receipt of treatment from those classified as needing treatment, it is possible to measure unmet need. Note, however, that unmet treatment need was not explicitly estimated.

Finally, arrest during the past year for any reason was also chosen. This behavior measure was chosen because of its importance and its correlation with substance abuse. Note that data on arrests are also available from the FBI Uniform Crime Reports, but that the UCR data are not directly comparable with NHSDA data since the UCR counts all arrests while the NHSDA counts persons with one or more arrests.

The 11 chosen behaviors are:

Use of legal (licit) substances:

1. **Cigarette Use In Past Month** - Smoked cigarettes at least once within the past month.

2. **Alcohol Use In Past Month** - Had at least one drink of an alcoholic beverage, that is, beer, wine or liquor or a mixed alcohol drink within the past month.

Use of illicit substances:

3. **Any Illicit Drug Use In Past Month** - Use within the past month of hallucinogens, heroin, marijuana, cocaine, inhalants, or the nonmedical use of sedatives, tranquilizers, stimulants or analgesics.

4. **Any Illicit Drug Use Other Than Marijuana In Past Month** - Past month use of any illicit drug excluding those whose only illicit drug use was marijuana.

5. **Cocaine Use In Past Month** - Use within the past month of cocaine in any form, including crack.

Drug or alcohol dependence:[2]

6. **Dependent On Illicit Drugs In Past Year** - Dependent on marijuana, inhalants, cocaine, hallucinogens, heroin, opiates or nonmedical use of sedatives, tranquilizers, analgesics, or stimulants. Those who are dependent on both alcohol and another illicit substance are included, but those who are dependent on alcohol only are not.

7. **Dependent On Alcohol But Not On Illicit Drugs In Past Year** - Dependent on alcohol but not dependent on any illicit drugs.

Treatment for drug and alcohol problems:

8. **Received Treatment For Illicit Drugs In Past Year** - Received treatment in the past 12 months at any location (including hospitals, clinics, self-help groups, doctors) for any illicit drug. These estimates include those who received treatment in the past 12 months for both drinking and illicit drugs.

[2]Detailed definitions of these dependence measures can be found in Appendix A.

9. **Received Treatment For Alcohol Use, But Not Illicit Drugs, In Past Year** - Received treatment in the past 12 months for drinking (including hospitals, clinics, self-help groups, doctors). These estimates exclude those who received treatment in the past 12 months for both drinking and illicit drugs.

10. **Needed Treatment For Illicit Drug Use In Past Year (whether or not treatment received)** - Persons who either: were dependent on illicit drugs in the past year; were a past year heroin user; received treatment in the past 12 months for any illicit drugs; were a needle user of heroin, stimulants or cocaine in the past 12 months; were a daily marijuana user during the past 12 months; or in the past 12 months were weekly users of hallucinogens, cocaine, inhalants or had weekly nonmedical use of stimulants, sedatives, tranquilizers or analgesics.

Past year arrest:[3]

11. **Arrested For Any Crime In Past Year** - Arrested and booked at least once for breaking a law in the past 12 months.

1.2.2 States and MSAs for Which Estimates Were Produced

Estimates were only produced for those States and MSAs which had some minimum NHSDA sample in them. Although estimates could have been produced even for States with no NHSDA data, it was thought that these estimates would not have been reliable enough to meet program needs. The States and MSAs selected for small area estimation are presented in ***Exhibit 1.2***. This exhibit shows the number of people who responded to the combined 1991-1993 NHSDA surveys plus information on the NHSDA sample including numbers of the sample MSA/County units, sample block groups, and the estimated 1992 population.

The States and MSAs presented in ***Exhibit 1.2*** were chosen for small area estimation because:

- The NHSDA sample size was close to 400 persons or more, a number selected so that at least some instances of even the rarest behaviors of interest such as cocaine abuse would be observed at least once in the state sample,

[3] Adjustments of NHSDA substance abuse statistics based on ratios of administrative record arrest counts divided by NHSDA survey estimated arrest reports have been proposed by D. Wright and J. Gfroerer (1994). "The use of external data sources and ratio estimation to improve estimates to hard-core drug use from the NHSDA," in Harrison, L. and Hughes, A., eds. *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates.* NIDA Research Monograph 167, NIH Pub. No. 96-4147, Washington, DC, Supt. Of Docs., U.S. Government Printing Office. However, these adjustments were not done for this report.

- The number of distinct sample PSUs was greater than or equal to 4 units, and
- The number of distinct sample segments was greater than 40 segments.

Exhibit 1.2 State and MSA Small Areas Selected for Inclusion in the Study Population Size and NHSDA Sample Characteristics

STATE	SAMPLE MSA/ COUNTIES[1]	SAMPLE BLOCK GROUPS[2]	SAMPLE RESPONDING PERSONS[3]	1992 POPULATION PROJECTION[4]
TOTAL UNITED STATES	213	8,942	84,974	205,945
NORTH EAST REGION	34	1,489	13,681	42,236
New Jersey	6	167	1,523	6,443
New York	7	843	8,505	14,892
Pennsylvania	11	269	2,133	9,945
SOUTH REGION	85	3,271	32,346	71,396
Florida	13	964	10,066	11,265
Georgia	5	118	1,061	5,442
Kentucky	6	112	1,081	3,053
Louisiana	6	133	1,099	3,387
North Carolina	12	222	1,863	5,641
Oklahoma	4	63	494	2,563
South Carolina	4	48	330	2,939
Tennessee	4	92	780	4,117
Texas	13	503	5,082	13,751
Virginia	9	362	3,538	5,227
West Virginia	3	44	394	1,497
NORTH CENTRAL REGION	49	1723	15,456	48,968
Illinois	6	799	8,088	9,378
Indiana	6	114	978	4,581
Kansas	4	60	515	2,010
Michigan	5	175	1,187	7,615
Minnesota	3	76	684	3,576
Missouri	6	120	1,059	4,223
Ohio	12	264	2,021	8,946
Wisconsin	3	55	475	4,021
WEST REGION	45	2,459	23,491	43,346
California	22	1,320	12,364	24,342
New Mexico	5	74	676	1,199
Oregon	4	59	412	2,397
Washington	3	73	690	4,094

[1]MSA/Counties refers to geographic entities formed to estimate random effect terms in the logistic model and which are generally analogous to NHSDA primary sampling units (PSUs). The exceptions are the distinct MSA constituents of PSUs that crossed State boundaries or combined more than one MSA.

[2]Block groups refers to the sample segments that were selected at the second stage of selection in the 1991-1993 NHSDA.

[3]87,915 people responded to the 1991-1993 NHSDA, however 2,941 people were omitted from the small area estimation research because of missing local area indicator variables that were used as potential predictors in the models.

[4]Population projections presented in 1000s.

Exhibit 1.2 State and MSA Small Areas Selected for Inclusion in the Study Population Size and NHSDA Sample Characteristics (continued)

MSA	SAMPLE BLOCK GROUPS[2]	SAMPLE RESPONDING PERSONS[3]	1992 POPULATION PROJECTION[4]
Anaheim-Santa Ana, CA	*	*	1,996

Atlanta, GA	*	*	2,425
Baltimore, MD	*	*	1,996
Boston, MA	*	*	3,145
Chicago, IL	735	7,537	4,981
Dallas, TX	*	*	2,120
Denver, CO	719	7,585	1,346
Detroit, MI	*	*	3,593
El Paso, TX	*	*	456
Houston, TX	*	*	2,661
Los Angeles, CA	768	7,533	7,127
Miami-Hialeah, FL	725	8,142	1,600
Minneapolis-St. Paul, MN	*	*	2,035
Nassau-Suffolk, NY	*	*	2,178
New York, NY	730	7,676	7,086
Newark, NJ	*	*	1,500
Oakland, CA	*	*	1,727
Philadelphia, PA-NJ	*	*	4,037
Phoenix, AZ	*	*	1,770
San Antonio, TX	*	*	1,040
San Bernardino, CA	*	*	2,122
San Diego, CA	*	*	2,089
St. Louis, MO-IL	*	*	2,006
Tampa-St. Petersburg, FL	*	*	1,822
Washington, DC	725	7,795	3,345

*Number of sample block groups ranged from 40 to 110; number of respondents ranged from 400 to 1200 in these MSAs.

[2]Block groups refers to the sample segments that were selected at the second stage of selection in the 1991-1993 NHSDA.

[3]87,915 people responded to the 1991-1993 NHSDA, however 2,941 people were omitted from the small area estimation research because of missing local area indicator variables which were used as potential predictors in the models.

[4]Population projections presented in 1000s.

See Appendix B for a list of counties included in each MSA.

Four exceptions to the four or more PSU unit rule were allowed for States that met the sample person and area segment minimums.

1.3 Limits of Design-Based Estimates

Design-based estimators are the standard approach used in most national surveys to describe the population. Under the designed-based approach, one treats the substance use status of each person in the population at a particular time as an intrinsic characteristic of the person. A fixed measurement protocol, consisting of a sampling design, a sample size, a questionnaire, a set of procedures for interviewers, a weighting procedure, and an estimator is selected to describe the fixed national population. The goal is to find estimators for various parameters of the national population such as the total number of people using a particular substance, the prevalence of substance abuse in the population, and so on. Estimates of these parameters are calculated based on the sampling design used for the survey, and their expected values and variances are derived by averaging across all possible samples that could be generated using the particular sampling design. (Särndal, Swensson, and Wretman, 1992). The variance across all possible samples can be estimated based upon the data and used to form "randomization-based" or "design-consistent" confidence intervals. With these confidence intervals, statements are possible along the lines of, "If the survey were independently repeated under the same general conditions and a large number of times and a point estimate and a confidence interval were constructed for each replicate, then approximately 95 percent of the confidence intervals formed would include the average of the point estimates." This approach is commonly known as design-based estimation and statisticians who use it can be called randomization-based frequentists or design-based statisticians. This is the procedure used in almost all demographic surveys in the U. S. to produce national estimates.

A key feature of the sampling designs that are used in national surveys is that each member of the population has a nonzero probability of being in the survey. Thus, a national probability sample is also a probability sample of any subgroup within the nation, and one can, in principle, make estimates for any small area within the nation. However, because of the small sample sizes that are obtained for these areas (including some zero sample sizes), the design-consistent confidence intervals can be extremely wide and the normal approximation required for the confidence intervals to be valid may not hold.

The NHSDA sample design was developed with the aim of producing highly precise estimates for the nation, for the four Census regions, and for various demographic subgroups defined in terms of age, race/ethnicity, and gender. It was not designed to produce state estimates. Thus, although the NHSDA is a valid probability sample of the population of each state, design-consistent confidence intervals for state estimates are likely to be very wide because the portion of the national sample that will happen to fall in a state is small or even zero, particularly for small states. In addition, there are other features of the design that make it less suitable for state estimates. A multi-stage clustered design was used. There was no attempt to construct state level strata to ensure a minimum representation within each state. Rather, the stratification focused on controlling the representation of the important demographic subgroups and employed extensive oversampling of young people, blacks, and Hispanics. Clusters at the first stage (the primary sampling units, PSUs) were large MSAs, individual counties in medium and small MSAs, individual nonmetropolitan counties, and small groups of nonmetropolitan counties. For the 1991-1993 samples, a total of 2,939 PSUs were formed. The largest 45 were selected with certainty. The remainder were then sorted by several variables such as percent minority. A sampling algorithm was used to select an additional 72-80 sample PSUs in a manner that reduced the variance for variables correlated with the sort variables. The set of sort variables did not include a marker for state.

Because there was no attempt to control the distribution within state boundaries, state estimates are extremely unstable. Furthermore, there was 1) extensive oversampling of important demographic domains, 2) clustering at the block level during the second stage of selection to reduce the cost of data collection, and 3) limits placed on the number of persons per sample household to increase the privacy of the interview resulting in people in large households having smaller probabilities of selection. The second-stage clustering and the differential sampling tend to make the estimates even more unstable. Thus, while these procedures are near optimal for national and regional estimates, they are not at all conducive to state estimation.

For example, there were just 330 people in 4 PSUs in the NHSDA in South Carolina from 1991 through 1993. Suppose that the design-based estimated rate of cocaine abuse during this period was 0.66 percent (the same as the estimate provided by the small area estimation project). Using a logit transformation and assuming that the logit of the estimated rate is normally distributed and that the 330 people had come from a *simple random sample*, the standard design-based 95 percent confidence interval would range from 0.26 percent to 1.65 percent. (Using the logit transform avoids the embarrassment of a negative lower bound on the confidence interval.) The design-based variance based on the NHSDA was not directly estimated for this project but is probably on the order of 2 to 6 times larger than the design-based variance from a simple random sample. Assuming that the factor is 4, the 95 percent confidence interval would run from 0.10 percent to 4.05 percent. Such an estimate and confidence interval would be of little use to South Carolina officials.

In addition to the problems that the instability of the design-consistent estimators pose for state planning purposes, this instability also leads to problems for the comparisons and ranking of states. Let π_i^D be the design-consistent estimate of the drug abuse rate for the i-th state and π_i be the limit of that estimate as the state sample size approaches the state population. The “D” stands for design consistent. The parameter π_i is defined as the limit of the design-consistent estimator rather than the “true” rate for the state due to the limitations of the measurement protocol. It can be fairly easily demonstrated that the small state sample sizes and large design effects cause

$$E\left[\underset{i}{\text{Range}}\{\pi_i^D\}\right] \gg \underset{i}{\text{Range}}\{\pi_i\}$$

and that

$$E\left[\sum_i \left(\pi_i^D - \pi\right)^2\right] \gg \sum_i \left(\pi_i - \pi\right)^2,$$

meaning that the range among the states and the variance among the states is exaggerated by the design-consistent estimates. This exaggeration of differences could have important consequences on public perceptions and official policy unless corrected. Thus, one of the properties that a sound alternative approach should possess is a capability to shrink the state estimates together toward the national average.

1.4 Review of Alternative Approaches to Small Area Estimation

The best current review of alternative approaches to small area estimation can be found in Ghosh and Rao (1994). In order to place the SAMSHA small area estimation project in context with other work on small area estimation, it is recommended that the reader consult this article. A book of case studies edited by Schaible (1996) is also interesting and useful for providing a background on approaches to small area estimation. The article by Ghosh and Rao is terse, requiring a fairly high level of statistical expertise on the part of the reader. For this report, the goal is to extract the kernels of the ideas presented by Ghosh and Rao and to present these ideas at a level accessible to the majority of social policy researchers. It is hoped that this material will leave such readers with an intuitive impression of the shortcomings of the other methods and the benefits of the new method presented here. The discussion in this chapter is at a fundamental level. Chapter 2 and Appendices C and G contain heavier mathematics.

The alternative approaches do not treat the substance abuse status of a person at a particular point in time as an intrinsic characteristic of the person. Instead, these approaches assume that it is a matter of chance for each person but that there are underlying factors that make this chance larger for some people than for others.[4] Before discussing the details of these alternative approaches, it will be useful to review their theoretical underpinnings. After reviewing these underpinnings, each of the most popular alternative approaches is reviewed including fixed effect models, composite estimators, and mixed effect models. The discussion in Section 1.4 closes with reasons for believing that the new approach is superior to those other alternatives.

1.4.1 Theoretical Underpinnings: Conceptions of Reality and Probability

In order to use national data to estimate state and local substance abuse rates, it is necessary to develop a theoretical framework for applying the observed patterns of substance use to groups of persons who were not interviewed. This can be done by considering each person's substance abuse status to be a random variable rather than a fixed characteristic of the person. The actual process leading to an individual's use of a particular substance may, in fact, not have a random component; it could be a complex deterministic function of genetics, family environment, local environment, government activities, and so on. However, since this process is too complex for us to understand, it is very useful to think of the process as being random within groups of people.

Under this approach, the statistician's goal is to identify groups of people within each of which there are no discernible differences in their patterns of substance abuse. It is then assumed that the probability of substance abuse is homogenous within each group. This probability can be thought of in two different ways.

One approach is to assume that there is a fixed but unknown probability for a group. With this approach, one thinks of the collection of persons in a homogenous group as an outcome

[4] Whether or not a person is using a certain drug at a certain time T can be thought of as a random variable until time T, at which point it becomes a fixed characteristic of the person. The choice of a distribution for that random variable depends upon prior information. Examining one extreme, if the person's drug using habits are known for all $t<T$ as well as the habits of all close friends, associates and family, and if it is known whether the person was involved in some sort of treatment plan in the time leading up to time T, then the propensity is close to either one or zero. Examining the other extreme, if nothing is known about the person other than that they live in the U. S. and are at least 12 years old, then the propensity will be some sort of national average. The more we know about a person's prior history and present situation, the more we tend to view current drug use as a fixed state of existence rather than as a random event.

of an underlying random process--as just one possible manifestation of an underlying superpopulation. The statistician defines the fixed but unknown probability of drug abuse for the group as the long-run average that would be obtained if this process were repeated independently a large number of times. The collection of possible manifestations of the underlying process is called the superpopulation. This approach to inference is termed the *model-based frequentist* approach. Applying this approach, the statistician uses the data to estimate the fixed but unknown probability and also to provide a confidence region about his estimate. The interpretation of such a confidence interval (e.g., a 95 percent interval) is that over all possible manifestations of the superpopulation, 95 percent of all confidence intervals constructed in this manner from samples of the given size from the group will include the fixed but unknown true probability for members of this subgroup. A statement can thus be along the lines of, "Given a sufficiently large number of observed manifestations, 95 percent of confidence intervals constructed in this manner will include the true probability of substance abuse for people in this group."

A second way to think of this probability is to consider it as a random variable as well. This assumes that there is a random process behind the random process theorized by model-based frequentists. In other words, the probability of substance abuse in a certain group at a certain time isn't simply some fixed p; instead, there is a random process that determines whether the probability of substance abuse for the group at a certain time is above or below p. This is called the *Bayesian* approach to inference after an 18th century theologian/mathematician who discovered a theorem that demonstrates how to use observed data to update subjective estimates of probabilities of various events. The Bayesian starts out the inference task with a prior distribution on p. This means that he has a function F (the prior distribution) such that $\Pr\{p \le \xi\} = F(\xi)$ for $0 \le \xi \le 1$. This prior is based upon past research, intuition, and personal judgment. He then uses the sample data and Bayes theorem to update his belief about the distribution of p and to derive a posterior distribution on p. This posterior distribution will depend to some extent on the prior distribution, but, if the sample size is very large, then this dependence will be slight. The Bayesian can use the posterior distribution to make point estimates and intervals (called prediction intervals) similar to the confidence intervals of the parametric frequentists. The point estimate is taken to be the mean or modal value of the posterior distribution (the expected and the most likely value of p, respectively) and the prediction interval is defined by the percentiles of the posterior distribution. For example, if a 95 percent prediction interval is desired, the lower point will be the 2.5th percentile of the posterior distribution and the upper point will be the 97.5th percentile of the same. Prediction intervals are interpreted along the lines of, "After considering the observed data, I now believe that there is a 95 percent probability that the probability of people in this group abusing the substance in question is in the stated range."

1.4.2 Using the Sampling Weights

Having now briefly presented the model-based frequentist and Bayesian perspectives, it is useful to contrast the treatment of sampling weights in these approaches with the treatment of sampling weights in the randomization-based approach.[5] This is critically important for an understanding of the small area estimation approach adopted for this study since both sampling weights and elements of the model-based frequentist's and Bayesian approaches have been used in developing the method described herein.

Neither the model-based frequentist nor the Bayesian uses sampling weights. They discard information on probabilities of selection and on joint probabilities of selection. They are, in fact, often just as happy to work with a purposely drawn sample as with a random sample. The design-based statisticians, of course, insist on a random sample with known marginal probabilities of selection and prefer having information on the joint probabilities of selection as well.

These differences in the use of sampling weights arise from the different conceptions discussed above. The design-based statistician is interested in estimating the drug abuse rate amongst a certain real population at a particular point in time. The model-based frequentist and the Bayesian on the other hand are not as interested in the real population at a point in time. They are more interested in the process that gave rise to that population. The latter two therefore construct models of the process and are interested in obtaining the most efficient estimates of the parameters of those models. Research has shown that using the sampling weights when estimating the parameters of a *correct* model only increases the variance. However, research has also shown that if the model is wrong, then using the weights could have reduced or eliminated the bias in the estimates due to model error (Holt, Smith and Winter, 1980). Since the design-based statistician believes that every model is wrong, he would rather have the less efficient but more robust estimates that are obtained by using the sampling weights. The other two groups reply that if the model is wrong, then there is no point in estimating the parameters at all; that all one can do is to use the best models available and then estimate the parameters of those models with as little variance as possible.

---

[5]For more information on the contrasts between the different schools of statisticians on this issue, see Smith (1994).

The practical importance of these differences in inferential approaches can be illustrated with a simple example. Suppose that there is a group that the modeler (whether frequentist or Bayesian) believes is homogenous, but that there is, in fact, a small difference between two subgroups that is undetected due to small sample sizes and that the sampling weights are different for the two subgroups. Let subgroup *A* have a process mean rate of substance abuse of 0.3, a sample size of 5 from a population of 10,000 (resulting in a sampling weight for all 5 sample persons of 2000). Let subgroup *B* have a process mean rate of substance abuse of 0.35, a sample size of 25 from a population of 12,500 (resulting in a sampling weight for all 25 persons of 500). The model-based frequentist will tend to estimate a substance abuse rate for the entire group of 0.3417 with an expected standard error of 0.0867 while the designed-based statistician will tend to estimate a substance abuse rate of 0.3278 with an expected standard error of 0.1054. Note that the expected estimate of 0.3278 is correct for the combined group but that it does not accurately describe the process for either subgroup and that it has substantially higher variance than that obtained by the model-based frequentist. With this small sample size, the estimate and standard error obtained by the Bayesian would depend fairly strongly on the prior distribution assumed.

For national estimates, it is more important to estimate the substance abuse rate for the population unbiasedly rather than to get more efficient estimates by assuming models either with or without prior distributions on the model parameters. The NHSDA sample was designed to do just that. However, as discussed above in Section 1.3, this approach is not feasible at the state or MSA level using current NHSDA data; so an alternative was needed.

In developing small area estimates of substance abuse rates, the various approaches were blended. Models have to be assumed to construct estimators for substance abuse at the state and local level, but it is convenient that the state estimates sum to national estimates that are close to the national estimates obtained using design-based techniques without the models. Furthermore, it was desired that state and MSA estimators, although model-based, would converge to design-consistent estimators where state/MSA sample sizes were sufficiently large.

1.4.3 Fixed Effect Models for Small Area Estimation

Fixed effect models include synthetic and regression estimators (domain-indirect estimators in the language of Schaible, 1996). Versions exist for design-based inference, model-based frequentist inference and Bayesian inference, but the development is perhaps simplest within the model-based frequentist approach.

The model-based frequentist basically assumes that one can model the substance abuse rate in a small area as a function of other information that exists for the small area. For example, suppose that a number of domains have been identified such as demographic subsets of the population for which current estimates are maintained for small geographic areas. The Census Bureau has divided the U.S. land area into a set of nonoverlapping very small areas called block groups and has published considerable information about every block group. Furthermore, information is available at the county level on fairly current arrest rates for illegal drug possession and for illegal drug sales or manufacture, deaths related to alcohol abuse, and alcohol and substance treatment rates. All of this information can be useful in developing models that predict local substance abuse rates.

Let π_{ib}^{d} be the true substance abuse rate for persons in domain *d* who live in block group *b* of state *i*.[6] Let N_{ib}^{d} be the number of people in domain *d* who live in block group *b* of state *i*. Let X_{ib}^{d} be a row vector of characteristics for people in domain *d* who live in block group *b* of state *i*. Examples of entries could be binary indicator flags for membership in domain *d* and lack of membership in other domains and median income for the domain within the block group, the percent below poverty, the percent with a high school degree, and so on. A common model would be something like

$$\operatorname{logit}\left(\pi_{ib}^{d}\right) = \ln\left(\frac{\pi_{ib}^{d}}{1-\pi_{ib}^{d}}\right) = X_{ib}^{d}\beta \qquad (1)$$

where β is a vector of unknown parameters. Maximum likelihood methods would be used to obtain an estimate β for β without paying any attention to the sample design. The substance abuse rate for the state would then be estimated as

[6] Recall that this is interpreted not as the actual rate at any particular point in time but as the mean of the underlying process. Also note that the word "true" must be interpreted in light of the inherent limitations of the NHSDA measurement protocol, meaning that only substance abuse that is voluntarily reported to interviewers is of interest.

$$\pi_i^F = \frac{\sum_d \sum_b N_{ib}^d \pi_{ib}^d}{\sum_d \sum_b N_{ib}^d} = \frac{\sum_d \sum_b N_{ib}^d \frac{1}{1+e^{-X_{ib}^d \beta}}}{\sum_d \sum_b N_{ib}^d} \qquad (2)$$

where the “F” stands for fixed-effect model.

As an example of how to set up such a model, suppose that there were three domains, two states, two block groups in each state, and no prediction variables other than domain flags. Then the *X* matrix would be as shown on the left below.

			Applies to People in		
			Domain	State	Block Group
X =	1	0	1	1	1
	1	0	1	1	2
	1	0	1	2	1
	1	0	1	2	2
	0	1	2	1	1
	0	l	2	1	2
	0	1	2	2	1
	0	1	2	2	2
	0	0	3	1	1
	0	0	3	1	2
	0	0	3	2	1
	0	0	3	2	2

Also, β would consist of just two parameters, β_1 and β_2. Comparing a person in domain 1 with a person in domain 3, the difference between the two domains in the logit of the propensity to engage in the behavior of interest is β_1. This translates into odds ratio of e^{β_1}. For example, if β_1=1, then the odds ratio is 2.7, meaning that someone in the first domain has odds of engaging in the behavior 2.7 time higher than for someone in the third domain. For more background on logistic regression, see Chapter 6 of Breslow and Day (1980), Chapter 4 of McCullagh and Nelder (1989) or all of Cox and Snell (1989).

In presenting such an estimate to an analyst, one could say that the national rate of substance abuse was estimated for combinations of a number of demographic domains in certain types of block groups (“neighborhoods”). One could also say that data from the Decennial Census were used to estimate how many inhabitants the state has in each combination of demographic group and block group. The national rates were applied to these state-specific population counts to estimate expected substance abuse rates for the state. Since there are many variables that influence the likelihood that a person will abuse substances other than those measured in the decennial census, the estimates would generally be different from those that would be obtained if a large state-specific sample were interviewed using the national procedures. The analyst could then use this estimate for planning purposes, along the lines of, “If there are this many people in need of assistance, then we need to do ... to meet those needs.” However, one should be cautious about using the estimates to correlate state action plans with estimated differences across the states. The estimates will not reflect differences across the states due to differences in laws, enforcement activities, advertising campaigns, outreach activities and so on, and will tend, rather, to be quite close to each other since the mix of demographic groups and block groups does not vary dramatically across the states. In essence, the fixed modeling approach shrinks the state estimates too close together. Also, if the estimates were repeated every few years, it would be entirely incorrect to try to correlate changes in the state estimates with any changes in policies or other state activities. The only changes would be due to changes at the national level.

An additional difficulty of using fixed effect models is that the estimated variances on these estimates would be highly misleading. The estimated variances would reflect sampling variability on the national rates only. Since the national rates are estimated with small variance, these estimates would appear to be quite accurate, even though they are not -- at least not in the broader sense of being close to what a large sample for the state would have found. Approaches have been suggested for replacing the variance with a mean squared error across all the states, but such approaches are not very satisfactory, as is discussed further in Section 1.5, Chapter 3, and Appendix H.

Thus, it is apparent that there are several important shortcomings with this approach. Most seriously, analysts who used them to assess the impact of state interventions would end up with seriously flawed analyses. Nonetheless, if there were no state specific data in the NHSDA, this would be the best that could be done. For example, there are no NHSDA sample data for Hawaii for the period of 1991 through 1993. This approach is therefore the best that could be done for Hawaii. However, there are state-specific data for all the states for which substance abuse rates were estimated in this study.

The fact that state-specific data do exist for each of the states and MSAs that were used in this analysis allows consideration of a model along the lines of

$$\text{logit}\left(\pi_{ib}^{d}\right)=\ln\left(\frac{\pi_{ib}^{d}}{1-\pi_{ib}^{d}}\right)=X_{ib}^{d}\beta+\eta_{i}, \tag{3}$$

where η_i is a fixed effect for the *i*-th state. However, whereas model (1) will tend to make the states look too much like each other, model (3) will tend to overstate the differences among the states, much as the design-based estimates do. Such a model would produce estimates of π_i that would be only marginally more precise than the design-consistent estimates. Like the design-consistent estimates, the range and variance of substance abuse rates would be exaggerated. What is needed is a method that shrinks the estimates closer together than the design-consistent estimators but not as close together as the fixed-effect model given in equation (1).

1.4.4 Composite Estimators

When at least some survey information is available for a small area, it is possible to use composite estimators that combine information from the survey and from models. The randomization-based approach can be used to create a design-consistent estimator, π_i^D, for the substance abuse rate in each state. As discussed previously, this estimator will be subject to intolerably high variance for most states due to small state sample sizes and the lack of state stratification in the first stage of sample selection. However, this estimator is unbiased (or at least design consistent) for actual substance abuse rates in the state at the point in time of interest. If, as is probable, the model alone does not adequately describe the propensity to abuse substances, then π_i^D contains information about whether the state was higher or lower than projected from a model based only on demographic distributions across block groups. A common approach is to average π_i^D with π_i^F. Let Γ_i be the averaging factor for the i-th state. Then the composite estimator of π_i is

$$\pi_i^C=\Gamma_i\pi_i^F+(1-\Gamma_i)\pi_i^D \tag{4}$$

The variance of π_i^C is larger than the variance of π_i^F, but π_i^C is closer to what would have been obtained with a large sample in the state.

This approach has considerable merit, but the problem is how to decide on values for the Γ_i. Additionally, it becomes more complicated to define what is meant by variance since the randomization-based approach has been mixed with the model-based frequentist approach. Various approaches have been suggested for choosing the Γ_i, but none are very satisfying. It is desired to pick Γ_i close to 1 when the model applies well for the state and/or the sample size is small for a state and to pick Γ_i close to 0 when the model doesn't fit well for the state and/or the sample size is large for the state. It is easy enough to decide when the state has a large or small sample size, but deciding when the model fits well in a particular state is impossible unless the state sample size is very large, in which case, it is known that Γ_i should be selected close to zero without even examining its fit for the state. Ad hoc rules can be developed, but better approaches to the entire problem have recently been developed.

1.4.5 Mixed Effect Models for Small Area Estimation

With a mixed effect model, one endeavors to produce estimates for each small area that reflect not only the level of substance abuse expected for an area given the national process means for different domains and types of block groups but to also reflect the impact of unmeasured variables, such as family environments, on the local substance abuse rates. The goal is thus quite similar to the goal for composite estimation -- smaller variance than the design-based estimates but expected values closer to what a large sample in each state would yield. The difference between using a mixed effect model and composite estimation is that fitting a mixed effect model is a somewhat more objective process than picking the Γ_i in composite estimation.

A one-stage mixed effect model looks like:

$$\text{logit}\left(\pi_{ib}^{d}\right)=\ln\left(\frac{\pi_{ib}^{d}}{1-\pi_{ib}^{d}}\right)=X_{ib}^{d}\beta+U_i, \qquad (5)$$

where U_i is a random variable with zero expected value and positive variance that is the same for all states and it is assumed that the variables are independent across states; i.e., $E\left(U_i\right)=0$, $\text{Var}\left(U_i\right)=\sigma_u^2$, and U_i $U_{i'}$ are independent for every $i\neq i'$..

Note the difference between models (3) and (5). In model (3), each state has a fixed deviation on the logit scale from the baseline propensity expected based just on demographic group and prediction variables.   In model (5), each state has normal random variation on the logit scale from the baseline propensity.   Over all possible manifestations of the state process, it is expected that each U_i will have an average value of zero.   Nonetheless, after the period in question, it is possible to talk about the realized value of the random variation.   In a particular past year, the deviation for each state must have been positive or negative.   Model (5) can be used to get estimates, U_i, of the value of U_i for a particular manifestation of the state process. For example, if a state had a social climate that effectively suppressed substance abuse, one would expect to see a value of $U_i<0$.   Conversely, a climate that was conducive to substance abuse would lead to a value of $U_i>0$.

Unless state sample sizes are large, the U_i from model (5) will be closer to zero than the η_i from model (3).   If all the state sample sizes are large, then the models yield similar estimates. This regression to the mean with small state sample sizes is desirable since each of the state processes is being measured with error, thereby exaggerating the differences among the states, and so there is a need to “shrink” the state estimates somewhat toward the national average. This is demonstrated in Section 3.2.2.   For a more thorough review of the benefits of making the state effect a random effect rather than a fixed effect, see Robinson (1991).

With this approach, the overall estimated rate of substance abuse for the state is

$$\pi_i^M = \frac{\sum_d \sum_b N_{ib}^d \pi_{ib}^d}{\sum_d \sum_b N_{ib}^d} = \frac{\sum_d \sum_b N_{ib}^d \frac{1}{1+e^{-\left(X_{ib}^d \beta \; 55\wedge + U_i\right)}}}{\sum_d \sum_b N_{ib}^d} \qquad (6)$$

where the “M” stands for mixed effect.   From the standpoint of being able to compare the states to determine what sorts of social climates are more resistant to substance abuse than others, it is desired that π_i^M is close to $\lim_{n_i \to N_i} \left(\pi_i^D \right)$, the limit of the randomization-based estimate for the state as the state sample size, n_i, approaches the total state population, N_i.   Note that this is not the same as desiring that π_i^M is close to π_i^D given the actual state sample size from a national sample.   Since π_i^D is very unstable with the small state sample sizes that are available from a national sample, the π_i^M should vary substantially from the π_i^D.

Prior efforts to fit this sort of model did not have the same goal of attempting to approximate the asymptotic values of the design-based estimators. Model-based frequentists have developed techniques to estimate $\boldsymbol{\beta}$, U_i, and σ_u^2 using maximum likelihood techniques. These techniques require that the U_i be normally distributed and run into both memory space and time problems on even the largest computers available. Bayesians place prior distributions on $\boldsymbol{\beta}$ and σ_u^2 but are free to use distributions other than the normal distribution for U_i. Bayesians use new techniques called Monte Carlo Markov Chain (MCMC) methods (of which the most well known example is Gibbs sampling) to obtain posterior distributions for $\boldsymbol{\beta}, U_i$, and σ_u^2. These techniques use huge amounts of computer time but are conceptually easy to program. A third approach is called empirical Bayesian, in which approximations are used to obtain posterior distributions for $\boldsymbol{\beta}, U_i$, and σ_u^2 with less computer time than a full Bayesian approach and without assuming prior distributions. For a review of these methods, see Chapter 9 of Diggle, Liang and Zeger (1994). The Bayesian approach is also discussed in Chapter 8 of Schaible (1996).

All of these techniques assume that the basic structure of the model given in (5) is "true." This means that the propensity to abuse substances can be totally explained by knowing the demographic group that a person belongs to, the type of block group that they live in (as characterized by a few variables from the prior Decennial Census), characteristics of the county, and the state that they live in. Even more strongly, all these techniques assume that the logit transform of this propensity is a linear function of fixed effects for the demographic group and type of block group and of a random effect for the state. The actual propensity to abuse substances may, of course, be considerably more complex reflecting things such as the quality of the local transportation network, the quality of the local police department, the proximity to major drug transportation routes, and a score of other fixed and random characteristics of the milieu in which each person lives. Since these techniques assume that the model is true, they make no use of the sampling weights. If the model is not true, they thus cannot be expected to converge to the substance abuse rates that would be estimated by very large state substance abuse surveys. To deal with this lack of convergence, a survey weighted approach was developed for this study.

1.4.6 The Survey-Weighted Empirical Bayesian Approach

This is the name given to the approach developed for this study although it could also be called a "survey weighted Penalized Quasi-likelihood" approach since the estimators can be motivated from either an empirical Bayesian viewpoint or from a frequentist view point.[7] The goal of this study was to be able to fit a model like (5) in such a manner that estimates would be close to design-based estimates for states with large NHSDA samples and for national domains with large sample sizes such as men and women. In terms of equations, it was desired that

$$\lim_{n_i \to N_i} \frac{\pi_i^W}{\pi_i^D} = 1, \text{ and} \tag{7}$$

$$\pi_d^W = \pi_d^D, \tag{8}$$

where π_i^W is the survey-weighted empirical Bayesian estimate of the substance abuse rate for the **state** and π_d^W is the survey-weighted empirical Bayesian estimate of the substance abuse rate for the **demographic group** across all states.[8]

The method was based upon the empirical Bayesian approach because it was thought to be the least expensive in terms of computer time. Even with this approach, large quantities of computer time were consumed in fitting all the models. Equation (8) is useful in terms of the face validity of the estimates. It is certainly awkward to explain why small area estimates do not aggregate up to national estimates. Equation (7) is somewhat useful in states with large sample sizes such as California, but questions do remain about how close π_i^W and π_i^D are to each other when the state sample size is not large. If the model is not true (as it almost certainly is not), the authors believe that their estimates will be better than those from model-based frequentist, Bayesian or Empirical Bayesian approaches, but it is not claimed that π_i^W will be extremely close to $\lim_{n_i \to N_i} \left(\pi_i^D \right)$ when the state sample size is small. It is not realistic to think that modeling exercises of this type can provide as much information about complex social phenomena as large local surveys. This question of evaluation is discussed further in Section 1.5. Guidance on interpreting the estimates is given in Section 1.6. It is not entirely straightforward to say what

---

[7]The approach follows the approach of Breslow and Clayton (1993) quite closely. The phrase penalized quasi-likelihood appears due to Green (1987).

[8]Within each of the four census regions, those states for which small area estimates were not produced were grouped together in a regional residual. These four regional residuals were treated the same as states in all the procedures.

is meant by $\text{Var}\left(\pi_i^W\right)$. More information on the mechanics of the survey-weighted empirical Bayesian approach is given in Chapter 2.

1.5 Evaluation Strategies

In any small area estimation project, questions arise as to the quality of the estimates. The purpose of using a model-based approach that combined NHSDA data with local data was to produce estimates of substance usage behaviors for States and MSAs that would have smaller variance than the direct estimates based solely upon the NHSDA. As discussed in Section 1.3, the direct design-consistent estimates provided confidence intervals for most States and MSAs that were too wide to be useful to researchers, policy analysts, and public health officials. It was recognized that in order to achieve the desired variance reduction, it would be necessary to let the validity of the resulting estimates depend on the models that were created to relate local data to NHSDA data. To the extent that the local data have good predictive power for the substance abuse behaviors of interest and that the models accurately reflect those relationships, the resulting model-based estimates for States and MSAs should be better than either the design-consistent estimates or simply assuming that the national rates apply to every State and MSA. Several meanings can be given to the word “better” in this context, but, regardless of the definition, actually demonstrating that the model-based estimates are better is a challenging task.

This issue is also being confronted by other government agencies. It is, in fact, at the heart of the debate about whether the Decennial Census should be adjusted for undercount. One of the chief methods for estimating the undercount is a national survey taken shortly after the census. Although there is some controversy about the adequacy of the undercount estimates at the national level, the most vigorous debate is whether adequate models can be developed to apply the national information about undercount to small areas in a way that makes the small area estimates better than the raw counts. This remains an open question despite years of major efforts by some of the best statisticians in the country.

This study had a fairly intensive evaluation component in it, but the results were more suggestive than conclusive. In Chapter 3, there is a thorough discussion of the various approaches that were used and the limitations of each. Results of the various approaches are also presented and discussed. More conclusive results might be achieved with a large scale simulation or perhaps the application of the methodology to variables that are measured well on the Decennial Population Census.

1.6 Guidance on Interpretation of Results

As mentioned earlier, the interpretation of confidence intervals provided for the state and MSA estimates of substance abuse is not entirely straightforward. First, it may be better to call these intervals prediction intervals since they are based more on Bayesian concepts than on frequentist concepts. The intervals are based on a model characterized by a national fixed pattern and a series of three random events that jointly determine whether a person abuses substances or engages in other behavior of interest. The national fixed pattern says that a person of a particular age, race and sex living in particular type of block group is going to have a certain baseline propensity to engage in the behavior of interest. The first random event is at the state level. It is assumed to affect everyone in the state in exactly the same manner by adding a small offset to the logit of the propensity to engage in the behavior of interest. The second random event is at the PSU level. It is assumed to affect everyone in the PSU in exactly the same manner by adding a second small offset to the logit of the propensity to engage in the behavior of interest. The third random event is at the person level. Given that the person has a propensity determined by his or her race, sex, age, type of block group of residence, specific PSU and specific state, it is assumed that his or her actual behavior is random with the probability of engaging in the behavior equal to the person's propensity. There was no attempt to explain a mechanism by which the random decision is reached. Statistically, it was treated as if a random number between 0 and 1 was drawn and if the random number was less than the propensity, then the person engaged in the behavior.

One can imagine averaging up the true propensities for some State, MSA or other group of interest and calling the average π. That hypothetical average π is a random variable subject to variance because of the random events at the state and PSU levels. The model states that π is normally distributed.

If all the estimated person-level propensities are averaged up for the same group, then another random variable results that is approximately normally distributed. This average is π, the estimated propensity for the group. The random variable π is subject to variance not just because of the random events at the state and PSU levels but also because of uncertainty in the estimated parameters of the fixed national model. The variable π is used as an estimate of π, but it is known that given π, the conditional expectation of π is not π. Since the conditional variance of π is a measure of how close it is to its own conditional expected value, interest does not focus on that conditional variance. There is even less interest in the unconditional variance of π since that measures the deviance of the π from its expected value over all possible state outcomes, PSU outcomes, and person outcomes -- rather interest focuses on measuring the deviance of the π from π given the state and PSU random events that actually occurred.

According with this interest, the “variance” is defined as the expected value of $(\pi-\pi)^2$ with respect to the distribution of the state and PSU level random effects and the distribution of the estimated parameters of the fixed national model. This might be called a mean square error, but that would lead to confusion with the design-based concept of mean square error. It would probably be most accurate to call it a mean squared prediction error, but that becomes cumbersome, and so it is referred to in this report as simply the variance.

The validity of the prediction intervals depends largely but not entirely on the correctness of the model. The model can be wrong in various ways. For example, the variance of the state level random effects could be larger in one region than in another; the variance of the PSU level random effects could be larger for larger metropolitan areas than for small rural counties or vice versa; the random effects might follow a different distribution from the normal; there could be correlations between state and PSU-level random effects; there might be additional random events at the county, tract, block group, block or household levels; there might be additional random effects associated with schools and places of work; it might be possible that the behavior of interest has a strong genetic component; everyone may have their own unique propensity that they share with no one else; or everyone’s behavior may be predestined with no chance variability at all.

Under some of the milder alternative models, the computed prediction intervals still have meaning, but their coverage properties may be lower than claimed. The use of weights partially protects against model failure since it is known that for large enough samples of large enough populations, the inference procedures come close to design-based inference. However, for predictions for small groups where only small sample sizes are available, the validity of the prediction intervals depends fairly strongly on the validity of the model.

For more information about the procedures that were used to estimate the variances, see Section 2.6 and Appendix G. These sections have some additional information on some acknowledged problems in the variance estimation. Those problems have probably led the prediction intervals to be somewhat too narrow. The degree of the inappropriate shortening is unknown but thought to be small. The estimated prediction intervals are shown in Appendix I.

2. The Survey-Weighted Empirical Bayesian Method

This method requires that a model be developed for predicting the propensity to engage in the behavior of interest. As a result of the modeling, it is possible to estimate the propensity to engage in the behavior of interest for every person in a state, whether or not they are in the sample. In Section 2.1, the structure of the models is described. In Section 2.2, the procedures for selecting predictor variables for the models are described. In Section 2.3, there is a sketch of the procedures used to fit the models once the variables had been selected — more details are given in Appendix C. In Section 2.4, there is a description of how separate models were fit for different segments of the population due to problems that prevented the fitting of a single model for the entire U. S. population. In Section 2.5, there is a description of how the final model was used to develop an estimate of the prevalence of the behavior of interest in targeted states and metropolitan areas. In Section 2.6, there is a discussion of how the variance of the estimator was approximated -- more details are given in Appendix G.

2.1 Structure of the Model

The full model used for most of the nation is specified here. In part of the nation, it was necessary to use a simpler model. The simpler model is described in Section 2.4.

Let the states be indexed by the subscript *i*, the sample PSUs within the states by the subscript *j*, and the sample persons within the sample PSUs by the subscript *k*. Then

$$y_{ijk} = \begin{cases} 1 & \text{if the } k\text{th person in the jth PSU of the } i\text{th state reported engaging in the behavior of interest and} \\ 0 & \text{otherwise .} \end{cases}$$

Let $\pi_{ijk} = (y_{ijk}) = y_{ijk}=1$. This π_{ijk} is the propensity for the indicated person to engage in the behavior of interest. A value of π_{ijk} close to 0 indicates that the person is very unlikely to engage in the behavior of interest, while a value of π_{ijk} close to 1 indicates that the person is nearly certain to engage in that behavior.

Let U_i be a normal random variable associated with the ith state and U_{ij} be another normal random variable associated with the *j*-th PSU of the ith state. These variables cannot be directly observed. They represent the impact of unmeasured conditions in the state and PSU on the propensity to engage in the behavior of interest, as will be explained more fully below. It is assumed that $E(U_i)=0$, $Var(U_i)=\sigma_1^2$, $E(U_{ij})=0$, $Var(U_{ij})=\sigma_2^2$, and that all the U_i and U_{ij} are mutually independent. The expectations and variances here are with respect to a superpopulation as was discussed in Section 1.4.

Let X_{ijkt} be the value of the t-th background variable for the indicated sample person. This can be the average income for the block group that the person lives in, the arrest rate for the county, an indicator variable for a particular type of block group or county, or an indicator variable for a particular type of person (i.e., an indicator for a particular demographic domain). By indicator variable, it is meant that

$$X_{ijkt} = \begin{cases} 1 & \text{if the person lives in the type of block group or county of interest or belongs to the demographic group or interest and} \\ 0 & \text{otherwise .} \end{cases}$$

Let $\lambda_{ijk} = \text{logit } \pi_{ijk} = \ln \left(\frac{\pi_{ijk}}{1-\pi_{ijk}} \right)$. Note that λ_{ijk} is referred to as the logit propensity for the individual to engage in the behavior of interest. Also note that the propensity, π_{ijk}, must be strictly between 0 and 1 for λ_{ijk} to be defined.

With this notation, the full model can be written as

$$\lambda_{ijk} = \sum_{t=1}^{p} X_{ijkt}\ \beta_t + U_i + U_{ij} \quad ,$$

where β_t is an unknown parameter that specifies the impact of the t-th background variable on the logit propensity. Also note that p is simply the number of background variables in the model.

The model assumes that over all possible manifestations of states and PSUs, for people with characteristics $X_{ijkl}, \ldots, X_{ijkp}$, there is a central tendency for the logit propensity to engage in the behavior of interest. That central tendency is $\sum_{t=1}^{p} X_{ijkt}\ \beta_t$.

However, the model assumes that this logit propensity varies across the possible manifestations of states and PSUs. This variation is due to unmeasured variables such as family background and peer group influences, efficacy of state and local programs to influence behavior, cultural differences and random chance. It is assumed that the variation follows a normal law, that the process causing the state variation is independent across the states, and that the process causing the PSU variation is independent across the PSUs. The use of the normal law seems reasonable since so many natural phenomena appear to obey it such as human height, weight, and IQ. However, it is not possible to directly test the reasonableness of this normal assumption. It is possible that these state and PSU random disturbances follow nonnormal distributions.

Some additional aspects of the model are important to note. The model assumes that everyone with the same values for the *p* background variables has the same central tendency in the logit of the propensity to engage in the behavior of interest. This is somewhat difficult to believe. Variables about the county and about the block group may have a very strong effect on this central tendency, but it seems likely that variables on family background would exert even stronger effects. This disbelief in the model led to the decision to include the sampling weights in the model-fitting procedure as was mentioned in Section 1.4.

This project was probably the first small area estimation project outside of the Census Bureau to use block group level data. Using county level data is a more common approach that is far less expensive to implement, but then the assumption of homogeneity in the central tendency of the logit propensity would have been even harder to believe. Block group-level variables were included in the modeling process in order to make the homogeneity assumption more credible.

The objective of the modeling process was to obtain posterior Bayes predictions of β_1, ..., β_p; all the U_i; and all the U_{ij}. (Note that the β_t are called fixed effects and that the U_i and U_{ij} are called random effects.) Variances and covariances on these predictions were also required in order to be able to form prediction intervals around the small area estimates. The parameters σ_1^2 and σ_2^2 were not directly of interest, but had to be estimated in order to obtain the desired predictions of the fixed and random effects.

Section 2.3 discusses how the long list of potential background variables was narrowed down to the final list of predictors. A sketch of the model fitting procedures is given in Section 2.4, and a detailed description of the model-fitting procedure is given in Appendix C.

This section concludes with an interpretation of the final model for cigarette smoking during the previous month among 18-25 year olds outside of the big cities. There were 117 fixed effects in this model. Their predicted values are given in Appendix E. The first effect is an intercept term entitled “DUMMY”. This means that $X_{ijkl} = 1$ for everyone in the sample. The predicted value of β_1 is = -0.560. This means that for a person for whom all the other background variables are equal to zero, the predicted central tendency of the logit of the propensity to have used cigarettes in the last month is -0.560. Taking the inverse of the logit transform, this means that the propensity is

$\pi = \frac{e^{-.560}}{1+e^{-.560}} = \frac{1}{1+e^{.560}} = 0.364$. On average then, 36.4% of persons with zero values of the

other background variables are expected to report smoking in the past month.

The second effect is for being female. This means that

$$X_{ijk2} = \{ 1$$

0 .

The predicted value of β_2 is β 55^$_2$=0.083. This means that females with zero values on all the remaining background variables have a predicted logit propensity of β 55^$_1$ + β 55^$_2$ =-0.477. Again taking the inverse of the logit transforms, this means that the propensity is

$\pi = \frac{1}{1+e^{.377}}$=0.383. On average then, 38.3 percent of females with zero values on the other

background variables are expected to report smoking in the past month. A property of the logit transform is that the ratio of the odds for males and females is constant under this model for males and females with identical remaining background variables. The ratio is

$\left(\frac{\pi_F}{1-\pi_F}\right) / \left(\pi \frac{subM}{1-\pi_M}\right) = \frac{e^{-.477}}{e^{-.560}} = 1.087$, meaning that females are estimated to have an

odds

for smoking in the past month that is 8.7% higher than the odds for similar males. Note that this result does not seem intuitive. It was expected that males would have a higher propensity to smoke at this age. When the significance probabilities in Appendix F are examined one sees that the *p*-value for the female term is a rather large 0.1047, meaning that standard procedures of hypothesis testing with a maximum tolerance for false positive rates of 0.05 or even 0.10 would accept the null hypothesis of no effect of sex on the propensity to smoke in this age range.

Continuing with the third through sixth effects, the modelling yielded a predicted value of β 55^$_{3}$= -.273 for black females and β 55^$_{4}$= -1.145 for other females, β 55^$_{5}$= -.903 for blacks, and -.008 for others. The central tendencies of the logit propensities for all combinations of race and sex are:

Logit (π)

	Male	Female
White	-.560	-.477
Black	-1.463	-1.653
Hispanic	-.560	-.477
Other	-.568	-1.630

and the corresponding propensities are:

	Male	Female
White	.364	.383
Black	.188	.161
Hispanic	.364	.383
Other	.362	.164

It is important to bear in mind that these are not population estimates. These propensities apply only to persons with zero values for all the other background variables. Since those remaining variables include block-group-level variables for which no one in the sample has a value of zero, these absolute propensities are actually not very interesting. The odds ratios are more interesting. These are shown below with white males as the reference group.

Odds Ratios for Smoking in Last Month

	Male	Female
White	1.000	1.087
Black	0.405	0.335
Hispanic	1.000	1.087
Other	0.992	0.343

The next three effects are for regions. The omitted Midwest region has a zero value for all three of these indicator variables. The next four effects are for type of setting the person lives in: large MSA, medium MSA, small MSA, or urban nonmetropolitan (small town distant from any major city). The omitted category here is rural nonmetropolitan (on a farm or out in the wilderness).

The next three effects are for the prevalence of children in the tract that the person lives in. Tracts are roughly the size of ZIP codes areas. There does not appear to be large variation in the percent of the population that is 0 to 18 years of age at such a high geographic level, but some tracts do have unusually high or low proportions of children. In this case, $X_{ijk,\,14}$ = for a linear trend effect across five categorized levels of percent of population in tract of person ijk that was 0-18 years of age at the time of the last decennial census (April 1, 1990). Also, $X_{ijk,15}$ = a quadratic trend effect $X^2_{ijk,14}$ and $X_{ijk,16} = X^3_{ijk,14}$. So β 55^$_{14}$= 0.174, β 55^$_{15}$= -0.020 , and β 55^$_{16}$= .094. Clearly, these parameters are of almost no interest to policy analysts. The most important point of this discussion is to illustrate the mechanisms of using the parameter estimates. The reason for estimating the parameters is only to be able to estimate the average propensity in every block group in the states for which small area estimates were produced.

The random effects are perhaps of somewhat greater interest. They were intended to measure the cumulative impact at the state and PSU levels of all the variables that could not be included in the model, such as official policies and intangible variables such as cultural patterns. In essence, the fixed effects tell us about relationships between known environmental factors and the behavior of interest and the random effects tell us about what is still unexplained. However, this interpretation is muddied somewhat by having county-level arrest and substance treatment data included amongst the background variables, since these variables are more naturally viewed as the results of substance abuse rather than as causes. Instead of representing shifts in the logit propensity to engage in the behavior of interest after controlling for age, race, sex, and local environmental variables, the random effects come to represent shifts after controlling for all those variables and county-level arrest and treatment data on substance abuse. Shifts of this type appear to be difficult to interpret. Although the random effects are not being published, they are available from SAMHSA.

2.2 Procedures for Selecting Fixed Effects

A wide range of possible predictor variables were considered. This section presents a description of how the variables were created and a discussion of how the list of actual fixed effects in the model was determined.

NHSDA Data: Although there are many useful background variables collected through the NHSDA survey instrument, most of these cannot be used in the modeling because it is impossible to get reliable estimates at the block group level of how many people possess these characteristics. Accordingly, the only predictor variables that could be used from the NHSDA interview were age, sex, race, and Hispanic origin. From geographic information from the sampling system, additional available variables included: region, size of metropolitan area (if metropolitan), and whether the block group was urban or rural. The four age groups were 12-17, 18-25, 26-34, and over 34. As explained in Section 2.4, separate models were fit for each age group, so age and all interactions of age with other predictor variables were forced into all the models. Sex, race, Hispanic origin, and geography were recoded into the 11 main-effect variables listed in ***Exhibit 2.1***. Additionally, the interactions of sex with race/ethnicity, sex with region, and race/ethnicity with region were tested.

Exhibit 2.1 NHSDA Predictor Variables (fixed effects)

Female
Hispanic (of any race)
Black but not Hispanic
Neither white, black nor Hispanic
Northeast
South
West
Large MSA (metropolitan statistical area)
Medium MSA
Small MSA
Urban block group outside of any metropolitan area

Note that one category has been excluded from each of the classifications that the variables induce (i.e., male, white but not Hispanic, Midwest, and rural nonmetropolitan). This was done to get unique parameter estimates. It was possible to have set extra constraints to make the parameters unique, but it was easier to exclude categories.

Census data: ***Exhibit 2.2*** lists Census data elements that were considered for use in the modeling. There were eighteen groups of these variables. All of the attributes come from the 1990 U.S. Census long form sample.[9] About 8 percent of the population was asked to complete a long form in 1990. For most of the variable groups, the possibilities were considered of using average responses both at the block group level and at the tract level.

Note that all of the census variables are approximately continuous at the tract level and block group level. There was concern that the relationship of the logit propensity to the predictor variables might be nonlinear. Frequently, for example, the upper and lower ends of the income distribution have more in common with respect to certain behaviors than the middle of the income distribution.[10] Consideration was thus given to using powers of the continuous variables as predictor variables, such as the square of the poverty rate, the cube of the poverty rate, and so on. However, using these powers of continuous variables as predictor variables probably would have led to outliers when the model was used to predict block-group level propensities. The reason for this is that when a continuous variable is used as a predictor, any outlying values of the predictor variable can lead to outliers in the predicted propensities.

In order to avoid the problems with outliers while simultaneously being able to estimate linear trends, quadratic relationships, and higher order relationships, some special recodes were created. These recodes were done in the same manner for each of the variables. The procedure was as follows:

Sort the units (counties, tracts or block groups, depending on the variable being recoded) by the variable. Using the poverty rate for tracts as example, that means that all the tracts in the nation were sorted by the tract-level poverty rate. The sorted list was then divided into 5

[9] Further information on these long form Census data can be found in the technical documentation for Summary Tape File #3 of the 1990 Census of Population and Housing as well as in the booklet "1990 Census of Population and Housing Tabulation and Publication Program" prepared by the U.S. Bureau of Census.

[10] A well known example of this concerns annual doctor visits. Since the poor have access to Medicaid and the upper end of the income distribution has excellent access to health care, the graph of doctor visits against income is U-shaped.

Exhibit 2.2 1990 Census Variables Used to Model Prevalence of Substance Abuse[11]

1. Race x Hispanic

Percent:
White nonhispanic
Black nonhispanic
Hispanic
Other

2. Education for persons 18 or older

Percent with:
0-8 years
9-12 years and no H.S. diploma
H.S. graduate
some college and no degree
associate degree
bachelors, graduate, or professional degree

3. Age

Percent aged:
0-18 years
19-24 years
25-34 years
35-44 years
45-54 years
55-64 years
65 and over

4. Poverty

Percent of:
families below poverty level

5. Public Assistance

Percent of:
households with public assistance income

6. Disability

Percent of:
persons 16-64 with a work disability

7. Household composition

Percent:
one-person households
of households with female heads (no spouse present) with children under 18

8. Employment

Percent:
of men 16 years and older in the labor force
of women 16 years and older in the labor force

9. Housing value - owner occupied units

Median value of owner occupied housing units

10. Housing rent - rental units

Median rents for rental units

11. Sex by marital status (persons 16 years and older)

Percent of:
Females currently married and not separated
Females separated, divorced, or widowed
Females never married
Males currently married and not separated
Males separated, divorced, or widowed
Males never married

12. Income

Median Household Income

13. Urbanicity

Percent:
of persons residing in an urban place

14. Urbanized Area

Percent:
of persons in an MSA urbanized area

15. Age of Housing

Percent:
of HUs built before 1939
of HUs built from 1940 to 1949

16. High School Dropout Rate (Tract level only)

Percent:
of high school age children who have dropped out

17. Hispanic Subpopulations

Percent:
of Hispanics that are Cuban
of Hispanics that are Puerto Rican

18. Other Race Subpopulations

Percent of:
Population that is Asian and Pacific Islander
Population that is Native American, Alaskan, or Aleut

---

[11] From Summary Tape File 3, 1990 Census of Population and Housing. Summaries were used at the tract level and at the block group level. Where a segment consisted of blocks from multiple block groups or even tracts, the summary variables were a weighted average of the statistics from the block groups or tracts that intersected the segment.

categories, each containing 20 percent of the tracts. Such groups are referred to as quintiles. The 20 percent of tracts with the lowest poverty rates are referred to as the first quintile of the tract-level poverty rate distribution. The fifth quintile consists then of the 20 percent of tracts with the highest poverty rates. The trend variables were then defined as shown in ***Exhibit 2.3***.

Exhibit 2.3 Definitions of Trend Variables

Linear	Quadratic	Cubic	Quartic	
-2	2	-1	8	if the tract was in the first quintile
-1	-1	2	-5	if the tract was in the second quintile
0	-2	0	-6	if the tract was in the third quintile
1	-1	-2	-5	if the tract was in the fourth quintile
2	2	1	8	if the tract was in the fifth quintile

Creating four variables for every one of the variables in the list of 18 variable groups would produce more variables than could possibly be simultaneously fit. In addition, there were problems with collinearity between variables,[12] particularly if both the tract-level variables and the block-group level variables corresponding to the same census question (such as average poverty rate for the tract and for the block group) were used. To reduce the collinearity problem, all variables that could be logically determined given other variables were dropped. As an example of this, consider the first group of variables concerning race and ethnicity. If the percent of the tract that is white but not Hispanic, the percent that is black but not Hispanic and the percent that is Hispanic are all known, then one can easily deduce the percent that is neither white nor black nor Hispanic.

To further reduce collinearity problems, block-group level variables were discarded if the correlation between the tract-level version and the block-group level version (across all the block groups in the nation) was above 0.80. Retention of the tract-level variables was favored over the block-group level variables because of the smaller sampling error on the tract level variables. However, this was a difficult decision. If all work had been done only at the tract level, the tasks of assembling the predictor variables and of obtaining population projections would have been much less expensive. There was substantial initial interest in the block group level data

[12]Two variables are collinear if the correlation between them is equal to one. The model fitting procedures would fail if any pair of variables was collinear or even approximately collinear.

because it was thought that they would have greater predictor power than tract level data. Of course, for those variables where the tract-level and block-group level variables are nearly collinear, it didn't matter much which were used since both versions contained roughly equal information in them.

There was also a desire to include some interactions of the Decennial Census variables with the NHSDA variables. For all Census variables that were kept after the preceding step, two-way interactions of the variables with sex and separately with race/ethnicity were generated. No other interactions were tested.

County level (social indicator) correlates of substance abuse: In addition to the Census variables, recoded county level 'social indicators' of substance abuse were also considered. These county level variables were obtained from three sources: The first of these sources was the FBI's Uniform Crime Reports database for 1991. This source yielded data on arrest rates per 10,000 persons for illegal drug possession, drug sales/manufacture by several reported drug categories, and total violent crime arrest rates. The second source combined data from the 1991 and 1992 National Drug and Alcoholism Treatment Unit Survey (NDATUS) conducted by the Substance Abuse and Mental Health Services Administration. From this source, data were obtained on the 1991 and 1992 average treatment rates per 1,000 county residents for (1) alcohol treatment alone and (2) for illicit drug treatment (including treatment for both drug and alcohol use). Thirdly, 1990 alcohol-related death rates per 10,000 county residents were obtained from the National Center for Health Statistics national death certificate registry. Two such rates were considered in this research: (a) the "any related rate" which includes all ICD-8 cause-of-death codes that are deemed to have a significant link to alcohol abuse, and (b) a more restrictive rate which requires explicit mention of alcohol on the death certificate. Appendix D presents the 1990 census block group, census tract, and County level variables that were considered as possible fixed effects. In appendix D, the defining variable labels of block group variables are preceded by the letter B. The tract and County variable labels are preceded by letters T and C respectively.

Underclass Indicator: Finally, there was a variable at the tract level from the Urban Institute that flagged tracts in distressed inner city areas. This was a binary recode of several tract-level census summary variables.

Further reducing list of possible predictor variables

Despite the work to eliminate collinear census variables, there were still too many variables and some of the variables had structural problems. A variable has such a problem if there exists a level of the variable such that no one in the entire NHSDA sample had the characteristic and also engaged in the behavior of interest. A logistic model always predicts a propensity strictly between 0 and 1 -- never exactly 0 or 1. If a predictor variable has a value that is associated with an observed propensity of 0 or 1, then the model fitting will fail. Accordingly, all such variables were dropped. Note that this was done separately for each of the 88 models, as were all the remaining steps in this section.

The next step was to fit a simple **linear** regression model using all the remaining possible fixed effects. The form of this simple linear regression model was

$$Y_{ijk} = \sum_{t=1}^{P} X_{ijkt}\beta_t + e_{ijk},$$

where it was assumed that all the e_{ijk} are independently and identically distributed as $N(0, \sigma^2)$. Since Y_{ijk} can be only 0 or 1, while the right hand side can predict any real number, this model is not at all satisfactory for most inferences. Nonetheless, models of this type do have several important advantages over logistic models when the analysis is in the exploratory stage. First, the linear models do still produce parameter estimates that are approximately correct. Second, they do not fail when some interaction term is associated with an observed propensity of 0 or 1. Third, linear models are far faster to run than logistic models since there is no iteration required. Fourth, the software for implementing variable elimination procedures is better written for linear models than for logistic models[13]. The idea was to quickly cull out poor predictors and predictors that were approximately collinear with other predictors.[14] To be on the safe side, the selection criteria were set loosely for these runs at α=0.1. If any interaction terms associated with a main effect were significant, then the interaction and the associated main effects were kept even if the main effects were not significant themselves. (This makes the model somewhat easier to interpret.) Similarly, if a higher order term was significant for one of the recodes for polynomial trends, then all the recodes for lower order trends were kept even if those lower order trends were not significant themselves.

[13]Elimination procedures are procedures that involve fitting a series of models, where the (n+1)-th model is obtained either adding the most important previously omitted variable from the n-th model or deleting the least important currently included variable in the n-th model.

[14]Collinear predictor variables also cause linear models to fail because the matrix X’X becomes ill-conditioned, making it difficult or impossible to invert it.

To improve the accuracy of the p-values, a modified version of the REGRESS procedure from SUDAAN was used to carry out the linear regression runs. SUDAAN increases the variance estimates to reflect the effects of clustering and unequal weighting in the sample design. The stripped-down version was a custom version made to run faster than the more flexible commercial version. SUDAAN requires the input of cluster and stratification structure for the sample design. The same structure information was used as would be used for national models from the NHSDA. This structure is different from the structure implicit in the mixed model, but this was not very important since the program was only being used to further cull out weak predictor variables.

The next step was to run fixed-effect logistic models with the retained variables. This was done using a stripped-down version of the procedure LOGISTIC from SUDAAN. The algorithm was the same as would have been used to fit a fixed-effect logistic model to any dataset from a complex sample survey, but some of the optional features were removed to make the program very fast. The same sample structure information that would have been used for national models from the NHSDA was again used. In these logistic fits, any variables that failed to pass the significance tests outlined above with α reduced to 0.05 were deleted. The fixed effect parameters obtained from these survey weighted LOGISTIC procedure runs were used as the starting βvalues in the mixed logistic model runs.

2.3 Fitting the Models

The full model described in Section 2.2 can be written in matrix form as

$$\lambda = X\beta + ZU ,$$

where λ, X, β, Z, and U are all matrixes. The structure of these matrices is a bit tedious to describe and understand, but the effort is well worthwhile in terms of the simplicity that they allow in describing the extremely complex model fitting procedure.

Let *m* be the number of states involved in the model fitting. The term “state” is used loosely here. A state-like entity in the model could be either a state, the balance of a state after taking out an MSA, or the balance of a region after taking out all targeted states. Let r_i be the number of sample PSUs in the *i*-th state. Let n_{ij} be the number of sample persons in the *j*-th PSU

of the *i*-th state. Let $n = \sum_{i=1}^{m} \sum_{j=1}^{r_i} n_{ij}$ be the total number of sample persons used in the model fitting. The matrix λ is an n x 1 matrix (i.e., n rows and 1 column). It is formed by listing all the λ_{ijk} with the first subscript varying the slowest, the middle subscript the next slowest, and the last subscript the fastest. (Recall that λ_{ijk} is the logit propensity to engage in the behavior of interest for the *k*-th person in the *j*-th PSU of the *i*-th state.) This is the same as listing all the logit propensities for the first sample PSU of the first state, then those of the second sample PSU and so on until all have been listed for the first state, and the proceeding to the second state and so on. The last entry is for the last person in the last sample PSU of the last state.

The matrix X is an n x p matrix (i.e., n rows and p columns), where p is the number of fixed effects in the model. The element in the j-th column of the i-th row is denoted X_{ij} and signifies the value of the *j*-th background variables for *i*-th sample person where the sample persons are ordered by state, PSU, and person, as in the description of λ above. The matrix β is a p x 1 matrix with elements β_1 through β_p.

Let

$$r = \sum_{i=1}^{m} r_i$$

be the number of PSUs in sample. Let q = m+r. This is the number of random effects in the model. Then U is a qx1 matrix where U_1 through U_m are listed first and then U_{11} through U_{mr_m} are listed next with all the PSU effects for the same state listed next to each other. The matrix Z is an nxq matrix consisting of 0s and 1s. The *j*th element of the *i*th row of Z is denoted Z_{ij} and signifies whether the *i*-th sample person is in the state or PSU associated with the *j*-th random effect. Exactly two elements of each row of Z are 1's since a particular sample person can only belong to one state and one PSU. The other elements of each row are all 0s. Each column of Z corresponds to either a state random effect or a PSU random effect. If a column corresponds to a state random effect then the number of 1s in the column is equal to the state sample size. Similarly, if a column corresponds to a PSU effect, then the number of 1s in the column is equal to the PSU sample size. The distribution assumptions about the random effects can be written compactly as

$$U \sim N_q(0\ 55\sim, G)\ ,$$

where N_q denotes a q-variate normal distribution, 0 60~? is a qx1 matrix of zeroes and G is the variance-covariance matrix for U.

Given the assumptions about independence between the random effects, a uniform variance for the state-level effects, and another uniform variance for the PSU-level effects, the matrix *G* has a particularly simple form. Let I_n be an *nxn* identity matrix. (A matrix with 1s on the main diagonal and 0s off the main diagonal. Then

$$G = \begin{bmatrix} \sigma_1^2 I_m & 0 \\ 0 & \sigma_2^2 I_r \end{bmatrix},$$

where, as before, *m* is the number of states and *r* is the number of sample PSUs. Note that *G* is *q*x*q*.

The objective of the model fitting was to find the best possible estimates of β and *U*. Once these are obtained, they are used to predict a propensity for every person in the population to engage in the behavior of interest. In other words, obtaining best estimates of β and *U* would lead to best estimates of λ, which would in turn lead to best estimates of π for the entire population. If π is known for the entire population, then it is easy to estimate the average propensities for small areas. Complicating this simple objective is the fact that statisticians do not agree on the way to obtain best estimates of β and *U*. This is an area of very active current research. There is a maximum likelihood approach favored by model-based frequentists, another approach favored by Bayesians, and an approach favored by empirical Bayesians. These various approaches are discussed below. That discussion is prefaced by a quick review of probability distributions and a general discussion of the different approaches to estimation.

By way of quick review, the distribution function for a random variable is usually written *F*(y) and is defined as *F*(y)=Pr{Y≤y}. If *F* is differentiable at *y*, then one may define *f*(y) = *F′*(y) and call *f* the probability density function (pdf). If *F* is not differentiable at *y*, then *f* is defined as

$$f(y) = \lim_{\varepsilon \to 0^+} (F(y+\varepsilon)) - \lim_{\varepsilon \to 0^-} (F(y+\varepsilon))$$

and it is then called a probability function (pf) since

$$f(y) = Pr\{Y=y\}.$$

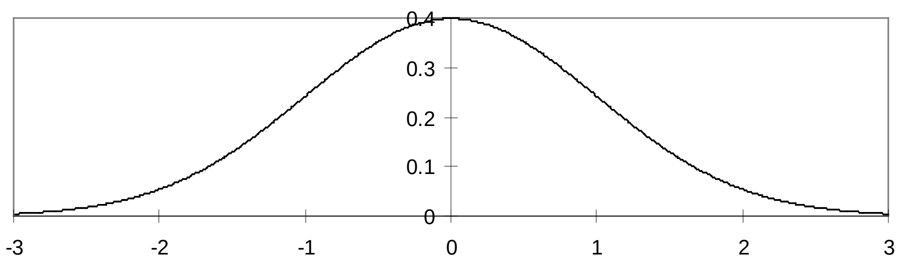
The modeler usually has a family of possible distributions in mind when making inferences based on a sample. That means that he believes the distribution of *Y* is known except for the values of a few parameters, labeled by the vector θ. Writing out f(y) will involve the use of θ in some way. Thus, f(y) is often written as f(y|θ), drawing attention to the fact that the pdf or pf of *y* depends on θ. Instead of viewing it primarily as a function of *y* that depends on θ, the pdf or pf can be viewed as a function of θ that depends on *y*. When it is viewed in this manner, it is called a likelihood function instead of a pdf or a pf and it is usually written as

$$L(\theta|y).$$

Note, however, that L(θ|y)=f(y|θ). Both functions appear identical. The difference is only one of emphasis. When f(y|θ) is used, the writer is, thinking about ways to use knowledge about θ to make predictions about *y*; when L(θ|y) is used, the writer is thinking about ways to use knowledge about *y* to make predictions about θ.

Exhibit 2.4 shows the pdf for a standard normal distribution. The pdf reaches its maximum value at the center of the distribution at Y=0. The point that maximizes the pf or pdf is called the mode of the distribution. If one considers trying to balance the graph at a single point along the horizontal axis, the center of gravity is also found at Y=0. The center of gravity is called the mean of the distribution. If one considers trying to find the value of Y such that half the area under the graph is on the left and half the area is on the right, that point is also found at Y=0. The point with half the area on either side is called the median of the distribution. For the normal distribution, all the measures are the same; i.e., the mode, mean and median are all equal to each other. This is not true for all random variables, but since the central limit theorem

Exhibit 2.4 Standard Normal Probability Density Function



states that the mean of a large random sample of values of a random variable will be approximately normally distributed, the mode of the distribution is close to the mean in many applications.

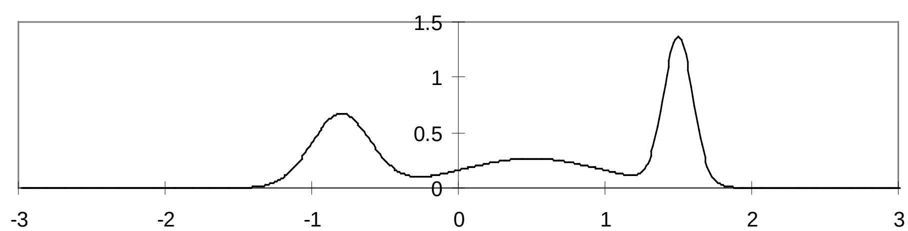
The shape of the curve in ***Exhibit 2.4*** is similar to the shape of the curve of the likelihood function for the unknown mean of a normal distribution given a simple random sample from that distribution, except for the fact that the likelihood function will be centered at the sample mean.

The method of maximum likelihood is to estimate a parameter by finding the mode of its likelihood function given the sample data. Suppose for example, that it is believed that Y is normally distributed with known variance but unknown mean and that a sample of 30 observations of Y yielded an average value of -34. The value that maximizes the likelihood function for the unknown mean given that particular random sample is -34.

Although this method works fine for many applications, there can be problems when the likelihood function is not symmetric or has multiple local maxima as in Exhibit 2.5. The likelihood function can look like this when the variable being measured (the *Y* variable) is not normal and the sample size is small. Note that for this random variable, the mean, median, and mode are all different. The median is 0.5, the mean is 0.4, and the mode is 1.5. For such a random variable, maximum likelihood estimation would not be a good method for finding the mean based on a small sample.

Despite the dangers of using maximum likelihood estimators for small samples, they have excellent properties when based on large samples. It has been demonstrated that for many classes of random variables, the maximum likelihood estimators of their parameters are asymptotically unbiased and exhibit minimum possible variance. In addition, the MLE

Exhibit 2.5. Probability Density Function for Mixture of Normal Variables



estimators are asymptotically normally distributed which is useful for forming confidence intervals.

The Bayesian approach to estimation would be to place a prior distribution on the parameter θ, then derive the posterior distribution of θ given the sample data, then find the mean of the posterior distribution. If p(θ) is the prior distribution on θ, then the posterior distribution of θ given Y is

$$p(\theta|y)=\frac{f(y|\theta)p(\theta)}{\int f(y|\theta)p(\theta)d\theta}$$

Note that if p(θ)=1 for all θ and if $\int f(y|\theta)d\theta$ exists, then p(θ|y)=L(θ|y), so the two methods are very similar for a flat p(θ). Such a prior distribution is called an uninformative prior because it states that all real numbers are equally likely to be the correct value of θ. Such a prior is an improper prior because a proper probability distribution has finite area underneath its graph. For some applications, use of an uninformative prior will result in the denominator of the posterior being infinite so that the posterior is not well defined. For other applications, the use of an uninformative prior is adequate. In general, use of uninformative priors is the biggest difference between the empirical Bayes approach used in this study and a fully Bayesian approach. With a fully Bayesian approach, only proper priors would be used.

If there are two random variables of interest, *Y* and *Z*, then the joint distribution of the two together is frequently of interest, as well as the marginal distributions of each by itself, and the conditional distributions of each given the other. These various distributions are related to each other as follows. Let the marginal probability functions or probability density functions for *Y* and *Z* be f(*y*) and g(*z*), let the joint pf or pdf of *Y* and *Z* be h(*x*,*y*), and let the conditional pf or pdf of *Y* given *Z* be f(*y*|*z*), and the conditional pf or pdf of *Z* given *Y* be g(*z*|*y*). Then

$$f(y|z)=\frac{h(y,z)}{g(z)} \quad ,$$

$$g(z|y)=\frac{h(y,z)}{f(y)} \quad ,$$

$$h(y,z)=f(y|z)g(z)=g(z|y)f(y) \quad ,$$

$$f(y)=\int h(y,z)dz=\int f(y|z)g(z)dz \quad ,$$

$$g(z|y)=\int h(y,z)dz=\int g(z|y)f(y)dy \quad ,$$

where the integrals are Lebesgue integrals that correspond to summation when *h* is a pf instead of a pdf.

If there is a sample of *n* random observations of Y, then the joint pdf or pf for the entire set of observations is

$$f_n(y_1, DOTSAXIS, y_n | \theta) = \prod_{i=1}^{n} f(y | \theta)$$

where the Π notation means to multiply all *n* factors together.

Returning from general review to this specific study, the pf for a single observation of substance abuse or other behavior is

$$f(y_i \mid \beta, U, G) = [\pi_i^{y_i}(1-\pi_i)^{1-y_i}] \ ,$$

for $y_i = 0$ or 1. (This is a standard result for Bernoulli random variables.) Using the general rule given above, the pf for the entire y vector of behavior observations given β, U, and G is

$$f(y \mid \beta, U, G) = \prod_{i=1}^{n} [\pi_i^{y_i}(1-\pi_i)^{1-y_i}].$$

The pdf for the vector of random effects given the assumption of a q-variate normal distribution is

$$f(U|G) = \frac{1}{\left(\sqrt{2\pi}\right)^q \sqrt{|G|}} \exp\left\{-\frac{1}{2}U^t G^{-1} U\right\} \ ,$$

where $|G|$ denotes the determinate of G, U^t denotes the transpose of U, and G^{-1} indicates the inverse of G (a matrix such that $GG^{-1} = G^{-1}G = I$, the identity matrix).

Using the rules given above for conditional and joint distributions, the joint pf/pdf for y and U is

$$f(y, U \mid \beta, G) = f(y|\beta, U, G) f(U|\beta, G) \qquad .$$

Since the distribution of U does not depend on β, this can be simplified to

$$f(y, U \mid \beta, G) = f(y|\beta, U, G) f(U|G)$$

$$= \left\{ \prod_{i=1}^{n} [\pi_i^{y_i}(1-\pi_i)^{1-y_i}] \right\} \frac{\exp\left\{-\frac{1}{2} U^t G^{-1} U\right\}}{(\sqrt{2\pi})^q \sqrt{|G|}}$$

If this function is now viewed as a function of β and U given fixed y and G instead of a function of y and U given fixed β and G, something very similar to a likelihood function is obtained:

$$L(\beta, U \mid y, G) = \left\{ \prod_{i=1}^{n} [\pi_i^{y_i}(1-\pi_i)^{1-y_i}] \right\} \frac{\exp\left\{-\frac{1}{2} U^t G^{-1} U\right\}}{(\sqrt{2\pi})^q \sqrt{|G|}}$$

Finding β and U that maximize this penalized quasi-likelihood has substantial intuitive appeal as a means for estimating β and U. However, there are two problems. The first is that G is unknown. The second is that $L(\beta, U|y, G)$ is not a true likelihood function since U is a latent variable rather than a parameter. A latent random variable is a random variable that, by definition, can not be observed. When the existence of a latent random variable is theorized, it is standard practice to also speculate on the nature of the joint distribution of the observable and the latent random variables. However, maximum likelihood estimation requires that the marginal distribution of the observable random variables given the distribution parameters be found, and that this marginal distribution be maximized to find the maximum likelihood estimates of the parameters. Using a procedure that is similar to maximum likelihood estimation but not exactly the same raises some questions about the properties of the estimators. As was mentioned above, maximum likelihood estimators are known to have good properties when sample sizes are large. Little is known about the properties of estimates obtained by maximizing $L(\beta, U|y, G)$ even when G is known.

This is the situation that was faced in this project. The random Y is observable, but the random vector U is not. So U is a latent random vector. The distributions of both Y and U depend on the parameter G. In addition, the distribution of Y depends of the parameter β. The parameter β is the real parameter of interest. It governs the relationship between the census background variables and the observed behaviors. The marginal conditional distribution of Y given the parameters β and G is, however, quite computer intensive to calculate. Using the relationships on marginal and conditional distributions given above, the marginal distribution of of Y given the parameters β and G is

$$f(y|\beta,G)=\frac{\left\{\prod_{i=1}^{n}[\pi_i^{y_i}(1-\pi_i)^{1-y_i}]\right\}\dfrac{\exp\left\{-\frac{1}{2}U^t G^{-1}U\right\}}{(\sqrt{2\pi})^q\sqrt{|G|}}}{\int\left\{\prod_{i=1}^{n}[\pi_i^{y_i}(1-\pi_i)^{1-y_i}]\right\}\dfrac{\exp\left\{-\frac{1}{2}U^t G^{-1}U\right\}}{(\sqrt{2\pi})^q\sqrt{|G|}}dU}$$

This can be turned around into a likelihood function for β and G given Y:

$$L(\beta,g \mid y)=\frac{\left\{\prod_{i=1}^{n}[\pi_i^{yi}(1-\pi_i)^{1-y_i}]\right\}\frac{\exp\left\{-\frac{1}{2}U^t G^{-1} U\right\}}{(\sqrt{2\pi})^q \sqrt{|G|}}}{\int\left\{\prod_{i=1}^{n}[\pi_i^{yi}(1-\pi_i)^{1-y_i}]\right\}\frac{\exp\left\{-\frac{1}{2}U^t G^{-1} U\right\}}{(\sqrt{2\pi})^q \sqrt{|G|}}dU}$$

The model-based frequentist would evaluate the integral in the denominator and then find the values β 55^ and$G$ that jointly maximize L(β,G|y). He would then use G to make predictions$U$ for U, and then useβ 55^ and$U$ to make predictions for λ and hence for π.

The difficulty with this approach is that the integration to find $L(\beta,G|y)$ is q-dimensional and each of those q dimensions must be integrated with numerical techniques since closed-form integrals of $L(\beta,U|y, G)$ do not exist. Some methods exist to reduce the dimensionality of the integration to the number of levels of random effects. For just a state effect, it is possible to reformulate the task as one involving q=m single-dimensional integrations; for a model with both state and PSU effects, the task can be reformulated into some number of two-dimensional integrations. This approach was, in fact, tried for this project. However, the computer time required for the integration was found to be unacceptable even with the 2-level nested random effect model. Since at the time, there were plans to develop 3-level and even 4-level nested random effect models (with random effects for separate segments and even households), this approach was abandoned.

The estimation approach of maximizing$L(\beta,U|y,G)$ can be easily justified from an empirical Bayes point of view. (That is the reason why the method is called an empirical Bayes method.) To see this, note that if an uninformative joint prior is placed on β and U, then the joint posterior of β and U given the data y and the variance matrix G is

$$p(\beta,U \mid y,G)=\frac{\left\{\prod_{i=1}^{n}[\pi_i^{y_i}(1-\pi_i)^{1-y_i}]\right\}\frac{\exp\left\{-\frac{1}{2}U^t G^{-1} U\right\}}{(\sqrt{2\pi})^q \sqrt{|G|}}}{\int\left\{\prod_{i=1}^{n}[\pi_i^{y_i}(1-\pi_i)^{1-y_i}]\right\}\frac{\exp\left\{-\frac{1}{2}U^t G^{-1} U\right\}}{(\sqrt{2\pi})^q \sqrt{|G|}}d\beta dU}$$

Since the denominator of this posterior is not a function of β or U (they have been integrated out), maximizing $p(\beta,U|y,G)$ is identical to maximizing $L(\beta,U|y,G)$. Thus, the method of maximizing the joint quasi-likelihood assuming known G is identical to the empirical Bayesian method, also assuming known *G*. The fact that the two approaches agree was comforting, as was the fact that it had been previously invented, tested, generalized, and alternately motivated by Stiratelli, et al (1984), Schall (1991), and Breslow and Clayton (1993). This was the approach adopted for this study.

The problem in applying this method is that *G* must also be estimated. If *U* were known then it would be possible to estimate G by maximizing $L(G|U)=f(U|G)$. This can be motivated through the maximum likelihood perspective or through the empirical Bayes perspective by placing an uninformative prior on *G*. Thus, there is a slight conundrum. Given the true value of *G*, it is possible to reasonably estimate β and *U*. On the other hand, given the true value of the unobservable *U*, it is possible to reasonably estimate *G*. If one knows neither *U* nor *G*, then it is difficult to know how to estimate either of them or β. Following, the work of the authors mentioned above, this conundrum was resolved through the use of an iterative technique.

In this iterative technique, there is an initial guess at a reasonable value for *G*. Using this guess, β and *U* are estimated by maximizing $L(\beta,U|y,G)$. The resulting estimate of *U* is used to estimate *G* through maximization of $L(G|U)$. This procedure cycles back and forth until convergence is obtained.

A fully Bayesian approach would have been to put a joint prior distribution on β, *U*, and *G*. The Bayesian would then find the posterior means β 55^, *U*, and *G* and use these to estimate λ and hence π. Until recently, the Bayesian approach involved even more intense integrations than did the maximum likelihood approach of the model-based frequentists. Now, with new Monte Carlo Markov Chain methods such as Gibbs Sampling, this approach may be computationally tractable. However, since these methods are new, reported to converge very slowly, and require the specification of prior distributions (a step with which SAMHSA was not very comfortable), this approach was not pursued.

As discussed in Section 1.4, the empirical Bayesian approach was modified slightly for this study to use the survey weights. The intent of this modification was to make the method more robust to model misspecification and to make the method design-consistent.

The modification to the method involves the insertion of the survey weights into the posterior distribution for *y* and *U*. Rather than maximizing

$$L(\beta, U|y, G) = \left\{ \prod_{i=1}^{n} \left[\pi_i^{y_i}(1-\pi_i)^{1-y_i} \right] \right\} \frac{\exp\left\{ -\frac{1}{2} U^t G^{-1} U \right\}}{\sqrt{2\pi}^{\,q} \sqrt{|G|}} \quad ,$$

the expression to be maximized is

$$L_w(\beta, U|y, G) = \left\{ \prod_{i=1}^{n} \left[\pi_i^{w_i y_i}(1-\pi_i)^{w_i(1-y_i)} \right] \right\} \frac{\exp\left\{ -\frac{1}{2} U^t G^{-1} U \right\}}{\sqrt{2\pi}^{\,q} \sqrt{|G|}} \quad ,$$

where w_i is the sampling weight for the *i*-th person, scaled so that the sum of the weights is equal to the total sample size. The procedure for estimating *G* is the same although the use of the weights does change the estimate of *G*.

If all persons in the 1991-93 NHSDA had the same sampling weight, then the estimators resulting from maximizing *L* and L_w would be the same. Also, if the background variables reflected in *X* totally explained the variation in sampling weights, the estimators would be the same. However, the sampling weights vary according to a more complex pattern than is reflected in *X*. As discussed in Section 1.1, the NHSDA oversampled by age and race and undersampled persons in large households. Also the sampling rates in six large metro areas were markedly higher than the national sample rates. Finally, block groups with high percentages of Hispanics and Blacks were oversampled to reduce the cost of overrepresenting Hispanics and Blacks. There is further unplanned variation caused by undersampling of blocks that were much larger at the time of listing than had been anticipated based on Decennial Census information and by adjustments for nonresponse and undercoverage. Since most of the weight variation due to age and race is reflected in the *X* matrix, and since separate models were fit for the collection of 6 oversampled metro areas, the most important differences between *L* and L_w concern the planned oversampling of concentrated Hispanic and Black neighborhoods and the undersampling of large households and unexpectedly large blocks.

The frequentist properties of the survey-weighted empirical Bayes estimates are not well understood. Since the only prior distributions used were noninformative (also called improper or vague), it is theoretically possible to derive the frequentist properties of these estimates, but this is quite difficult and has not yet been rigorously resolved. Breslow and Clayton (1993) point out that for estimating β and predicting *U*, this method is approximately equivalent to a maximum likelihood approach. They also point out structural similarities to estimators proposed by Harville (1977) for normal *y* variables. They also report on two small simulation studies, the results of which are moderately encouraging. As discussed in Section 1.5, larger, more varied, and more realistic simulation studies would be useful for evaluating the survey-weighted empirical Bayes model fitting procedure. This idea is discussed further in section 3.4.

2.4 Need for Multiple Models

The ideal modeling procedure would have been to fit one national model for each of the 11 behaviors of interest. For technical reasons, however, it was necessary to fit 8 models for each of the 11 behaviors for a total of 88 models. Separate models were fit for four age groups in each of two oversampling strata. The four age groups were 12-17, 18-25, 26-34, and over 34. The oversampling strata were defined by the set of metropolitan areas that were given sufficiently large sample sizes to allow design-based estimates with good precision. There were 6 such metropolitan areas. The combination of these 6 metropolitan areas constituted one of the strata for which separate models were fit. The other stratum consisted of the rest of the nation. The models for the collection of 6 targeted MSAs are referred to as the “Big City Models.” The models for the balance of the country are referred to as the “Remainder Models.”

There was an initial attempt to fit just one national model for each behavior for several reasons. First, a single national model would have yielded more stable estimates of the fixed effects. Second, the estimates of the fixed effects would not vary by age or stratum as they do when 8 different models are fit. Third, the estimates of mean square error of the small area estimates would have been more accurate. Lastly, it would have been somewhat easier to organize the file handling aspects of the task.

The reason for having to fit a different set of models for the oversampled MSAs was that the average sampling weight in the oversampled MSAs was much smaller than the average sampling weight in the balance of the country. The large disparity in weights caused a critical matrix in the model-fitting procedure to become ill-conditioned, meaning that it could not be inverted. (See Appendix C for a more detailed explanation of this point.) Larger and faster computers probably would not have solved this problem. It appears that although the method can make the small area estimates more robust through reflection of the sampling weights, if the variation in the sampling weights is too extreme, then the analysis must fail.

The reason for having to fit a different set of models for each age group was related to computer size and speed. Since age is the most important predictor of the behaviors of interest, it was desired to include many interactions of age with other variables. The X matrix of predictor variables became too large to read in at acceptable speeds. It was thus decided that the best way to shorten the X matrix without giving up very much predictive power was to fit separate models by age group. With a larger and faster machine, there would not have been a need to fit separate models by age group. Another solution might have been to use fewer fixed effects. At the time, there was a reluctance to give up prediction power by using fewer fixed effects, but since the evaluation in Chapter 3 seems to indicate that the fixed part of the model was overfit (too rich, meaning too many fixed effects), this might be the best solution to the problem in future applications.

The big city models had a simpler structure than the remainder models. The structure given in Section 2.1 is for the remainder models. For the big city models, the terms for the PSU-level random effects (U_{ij}) were forced equal to zero. Although it might seem natural to fit a model with 6 state-level random effects corresponding to the 6 MSAs, it was felt that the estimates of σ_1^2 would be very unstable if they were based upon just 6 realizations of U_i. Accordingly, each of the six oversampled MSAs were divided in two. Although it would have been possible to split each MSA by county, it was felt that it would be more useful to split each MSA into two strata of block groups, one with higher SES (socio-economic status) than the other. SES was measured by rent levels and housing values. The cut points on rent and housing values were set so that about 30 percent of the population in each MSA was in the low SES stratum. Thus, for the big city models, m=q=12.

After the sample in the oversampled MSAs was dropped out from the remainder models, those models converged nicely. However, there was still too much variation in the weights within the 6 oversampled MSAs. The reduced models did not converge initially. The problem was that Denver and Miami are much smaller MSAs than New York, Los Angeles, Chicago, and Washington. In order to obtain the desired sample sizes for Denver and Miami, these two MSAs had to be oversampled at rates considerably higher than in the other four MSAs. To resolve this problem, the weights were standardized within each of the 12 random-effect groups defined by MSA and by SES stratum. This means, that within each of these groups, the sampling weights were forced to add up to the nominal sample size. This was done by simply dividing every weight by the sum of all weights for the group. This adjustment preserved the variation of weights within each group but clearly changed the variation across groups considerably.

This adjustment to the weights means that the big city estimates are no longer nearly as design consistent as the remainder estimates are, but comparisons of the model-based estimates and the design-consistent estimates (unpublished) still show good agreement between the two, leading to the belief that the effect of these weight adjustments was fairly minor.

2.5 Form of the Estimator Given a Fitted Model

The population was divided into 32 domains by age, sex and a combination of race and Hispanic origin. The four age groups were 12-17, 18-25, 26-34, and over 34. The four race/ethnicity groups were white but not Hispanic, black but not Hispanic, other but not Hispanic, and Hispanic. (The "other but not Hispanic" group consists primarily of American Indians, Eskimos, and Aleuts; and Asians and Pacific Islanders.) For each behavior of interest, there were eight models covering different age groups and parts of the country as described in Section 2.4. After concatenating the predictions from the eight models, an estimated propensity to engage in a behavior of interest was available for each of 32 domains inside every one of the block groups in the 50 states and the District of Columbia. (There were about 230,000 block groups defined for the 1990 Decennial Census.)

In this report, the age domains are indexed by *a*, the sex-race/ethnicity subdomains with each age domain by *d*, the states and targeted metropolitan areas by *i*, and the block groups within each state and targeted metropolitan area by *b*. The list of targeted states and metropolitan areas is shown in ***Exhibit 1.2***. In order to apply the fitted model, the *X* and *Z* matrices originally defined in Section 2.3 had to be redefined. When fitting the model, both of these matrices have one row per sample person. When applying the model, they both have one row for every domain defined by age, sex and race/ethnicity in every block group in every state of interest, whether or not the block group is in a sample PSU. Where the meaning is clear, *X* can mean either the matrix with one row per sample person or the matrix with one row per domain per block group. Where both matrices appear in the same equation or there is a desire to stress the choice of matrix, ***X* 45*?** is used for the matrix with one row per domain per block group. The same convention is followed for the matrix *Z*. The rows of *X* 45*? are denoted byX_{ib}^{ad}. Similarly, the rows of ***Z* 45*?** are denoted by $\boldsymbol{Z_{ib}}$, dropping the *a* and the *d* subscripts since all the rows of ***Z* 45*?** in the same block group are identical. The row vector $\boldsymbol{X_{ib}^{ad}}$ consists of flags for the age group and domain, census data about the county, tract, and block group, and administrative data about the county as described in Section 2.2. For the remainder models, the row vector Z_{ib} consists mostly of zeros with two 1's in each row to indicate the state and PSU to which the block group belongs. For the big city models, the row vectorZ_{ib} has just one 1 to indicate the city and SES stratum of the block group.

Let $\hat{\beta}_a$ be the column vector of estimated fixed effects for the *a*-th age group and let $\hat{U}_a$ be the column vector of estimated state-level and PSU-level random effects for the *a*-th age group. Here, it is important to note that the PSU effect was taken to be zero if the PSU was not in sample. Also, to avoid conflicting predictions for the big cities, the state level random effects for states containing oversampled MSAs were actually defined as balance-of-state random effects. Finally, in order to be able to make national estimates, all the states in a region not targeted for separate estimates were combined to form a pseudo state. Each pseudo state was assigned a separate state-level random effect. Give this notation, the propensity for persons of the indicated age, sex, race and ethnicity in the indicated block group to engage in the behavior of interest was calculated as:

$$\hat{\pi}_{ib}^{ad} = \frac{1}{1+e^{-(X_{ib}^{ad}\hat{\beta}_a + Z_{ib}\hat{U}_a)}}.$$

As an example, if $\hat{\pi}_{ib}^{ad}=0.1$, this means that the model predicts that 10 percent of these people are likely to engage in the behavior of interest.

Using data from Claritas (a private firm specializing in current demographic data), the number of people in the specified area who have the indicated demographic characteristics was estimated. Let N_{ib}^{ad} be that estimated population size. Then the estimated number of people in that block group who engaged in the behavior of interest at the midpoint of the 1991-1993 period is $N_{ib}^{ad}\hat{\pi}_{ib}^{ad}$.

By summing this number up across all the block groups in the *i*-th area, a state/local estimate of the number of people in the age-sex-race-ethnicity domain *ad* who engaged in the behavior of interest at that time is obtained. By summing further on *i*, regional and national estimates are obtained for the domain *ad*. Summing alternatively on *b*, *d*, and *a*, one can obtain small area estimates for all ages, sexes, races, and ethnic groups combined.

Using formulas, the average rate of engaging in the behavior of interest in the *i*th small area is estimated to be

$$\hat{\pi}_i^W = \frac{\sum_a\sum_d\sum_b N_{ib}^{ad}\hat{\pi}_{ib}^{ad}}{\sum_a\sum_d\sum_b N_{ib}^{ad}}.$$

Similarly, the average rate of engaging in the behavior of interest in the *d*-th subdomain of the *a*-th age group is

$$\hat{\pi}^{W}_{ad} = \frac{\sum_i \sum_b N^{ad}_{ib} \hat{\pi}^{ad}_{ib}}{\sum_i \sum_b N^{ad}_{ib}} \quad .$$

2.6 Estimating Variance

The estimated prevalence rate for the *i*-th state or other small area is

Error!

where *a* indexes the four age domains for which separate models were run, *d* indexes the eight race/ethnicity by gender domains which were modeled by fixed effects in $\hat{\beta}_a$, *b* indexes block groups within the area, N^{ad}_{ib} is the domain-*ad* population of the indicated block group, X^{ad}_{ib} is the row vector of characteristics of the fixed predictor variables for the block group, Z_{ib} is the row vector of flag variables for the block group that indicated which random effects apply, and $\hat{\beta}_a$ and U_a are the vectors of fixed and random effects peculiar to the model for age group *a*.

This is clearly a nonlinear estimator for which one can only hope to approximate the variance. First the types of errors that need to be reflected must be conceptualized. As discussed in Section 1.4, one could be interested in variability and bias across replications of the survey, across possible manifestations of the population, etc. In the empirical Bayes approach, there are three random events that determine each person's propensity to engage in a behavior of interest. The first is a random event at the state level that determines the U_i component of the person's propensity. The second is a random event at the PSU level that determines the U_{ij} component of the person's propensity. The third is the flip of the biased coin that decides whether or not the person will actually engage in the behavior, where the probability of "heads" on the coin is equal to the person's realized propensity given their PSU, state, and background variables.

If π_{ib}^{ad} is interpreted as the uniform propensity for persons in domain *ad*, block group *b*, and state *i* to engage in the behavior, interest focuses on how closely the estimate π_i^W is expected to track the average of these propensities

$$\pi_i = \frac{\sum_a \sum_d \sum_b N_{ib}^{ad} \pi_{ib}^{ad}}{\sum_a \sum_d \sum_b N_{ib}^{ad}}$$

across all possible outcomes of the three variables across all members of the population; that is, it was decided to estimate $E(\pi_i^W - \pi_i)^2$, where both π_i^W and π_i are random variables and the expectation is with respect to the posterior distributions of π_i^W and π_i given the observed data.

It was decided not to try to reflect any extra variation in $(\pi_i^W - \pi_i)$ across all possible samples of a fixed population caused by the segment and household-level clustering in the sample design. That is, of course, of some concern for a hybrid method that incorporates design-based weights in the Bayesian approach, but there would be severe technical difficulties in trying to reflect that extra variation.

Also for reasons of technical difficulty, it was decided not to reflect the extra variation in $(\pi_i^W - \pi_i)$ caused by the estimation of σ_1^2 and σ_2^2, the components of variance. Instead, the mean was calculated as if $\sigma_1^2 = \sigma_1^2$ and $\sigma_2^2 = \sigma_2^2$ were fixed and known. This is not the desired procedure from a Bayesian viewpoint but seems to be unavoidable in the empirical Bayes approach. The magnitude of the error in estimating $E(\pi_i^W - \pi_i)^2$ caused by treating σ_1^2 and σ_2^2 as fixed is unknown, but some discussion in Appendix G indicates that the underestimation of $\sqrt{(\pi_i^W - \pi_i)^2}$ could be in the range of 5 to 20 percent.

Another technical difficulty prevented reflection of the fact that $Cov(\pi_{ib}^{ad}, \pi_{ib}^{a'd}) \neq 0$ for $a \neq a'$. Computing constraints required that this covariance be treated as zero. Since these covariances will generally be positive, this problem also led to the underestimation of $E(\pi_i^W - \pi_i)^2$. Again, the magnitude of the underestimation is unknown.

Perhaps most seriously, the overfitting discussed in Chapter 3 probably led to underestimation of the variance. The link between overfitting a model and underestimation of model error can perhaps best be understood with a simple extreme example. Consider the linear regression of a single continuous variable on *n* independent binary variables with a total sample size of *n* cases. Suppose further that all the predictor variables are independent of the continuous variable. In this case, the model fit will be perfect even though the predictor variables have no predictive value. Estimated model error will be zero. Actual prediction error will equal the variance of the continuous variable. Nothing that extreme was done on this project, but there were hundreds of predictor variables in some of the models. The sizes of the random effects were almost certainly underestimated due to the large number of predictor variables. This then probably led to underestimation of the error in the model.

Finally, there was an error in the variance formula that affected models with two stages. Corrected variances have been calculated for some of the estimates but are not shown in this report. The impact of the correction on standard errors was small (0-10%) for past month cigarette use, past month cocaine use, past month use of any illicit substances other than marijuana, and past year alcohol treatment. The impact of the correction on standard errors was moderate (0-30%) for past month alcohol use and past month use of any illicit substance. Appendix I shows the original confidence intervals prior to correction of the error. For a further discussion of these issues and a detailed description of the variance estimation procedure, see Appendix G.

3. Evaluation

This chapter discusses the work that was done to evaluate the quality of the small area estimates. This work can best be thought of as a partial evaluation since it is difficult to evaluate the quality of estimates based upon a mixture of models and data when the phenomenon being described is never actually measured. The purpose of small area estimation methodology is to have narrow confidence intervals without paying for extensive data collection. But it is precisely that kind of data that is needed to evaluate the methodology. If the data required for a stringent evaluation were available, then there would be no need for the small area estimation program. Despite this central dilemma, three types of partial evaluation were carried out.

The first type was to compare the model-based estimates with alternate estimates for the small areas of interest. This is a popular method with a fairly long history, but it was only marginally useful for this application. The second type of partial evaluation was to compare the model-based estimates with alternate estimates for artificial domains where theoretical considerations lead one to believe that agreement between the methods should be good. This method yielded more information, particularly for the more common behaviors of interest such as alcohol use, but is still subject to serious caveats. The third type of partial evaluation was to compare the model-based estimates with data from other federal data systems for the small areas of interest. Comparisons of this type are, of course, interesting, but methodological differences between the data systems make these comparisons difficult to interpret.

None of the three types of partial evaluation were very conclusive by themselves. The summary report for this study (SAMHSA, 1996) found that the joint preponderance of evidence favored the survey-weighted empirical Bayes approach over the alternatives and that for most of the estimates developed, there was evidence that the estimates adequately reflect the prevalence of substance abuse characteristics for States and MSAs. Upon further review of the evaluation methodology, it appears that some of the evaluation findings in the summary report should perhaps have been somewhat less definitive. The basic conclusions have not changed, but they are slightly more tentative in this report.

Another result of the methodology review has been some discussion of what could have been learned from a well crafted simulation study. In favor of the simulation study approach, it might be argued that the central reasons for using the survey-weighted empirical Bayes procedure were that it was thought to provide: (1) estimates with smaller errors than the design-based methods, (2) better error estimates than methods based only on fixed effect models, (3) better consistency with national estimates than could be achieved through the use of mixed models without survey weights, and (4) computational savings relative to a fully Bayesian or maximum likelihood approach. The evaluation work presented in this report does not address any of these points. At this time, it is strongly suspected but unproven that the first two assertions are true. The third assertion appears to be true on theoretical grounds alone, but the magnitude of the improvement associated with using the weights is unmeasured. The fourth assertion appears to be subject to change as the methodology for each approach improves with additional research.

The central questions from the users' point of view are whether the prediction intervals really are nearly certain to include the truth and whether the intervals are short enough to be useful. Only the user can determine whether the intervals are short enough to be useful to them, but the quality of the coverage of the prediction intervals is an open question that needs additional research. A computer simulation study would address this question directly. This point is discussed further in Section 3.4.

3.1 Comparisons with Alternate Estimates for the Small Areas of Interest

Estimates based on different models were compared to see whether theories about the properties of the various models are borne out. Specifically, there was interest to see (1) if the survey-weighted empirical Bayes estimates for California would agree better with the design-based estimates for that state and (2) how the dispersion of the state estimates varied by method. California was of interest because the very large sample size of 12,000 people in California makes the design-based estimate the most believable estimate for that state.

Regarding dispersion, it was expected that estimates based upon the leanest fixed-effect model would have the most compressed dispersion. Estimates based upon a mixed model were expected to have broader dispersion but not as broad as the design-based estimates. If there had been a fixed effect model with broader dispersion of state estimates from the national average than what was observed with a mixed model or even with the design-based estimates, then this would have constituted strong evidence that the fixed effect model had been overfit.

The estimates from the different models were not compared with the idea of deciding which set was best in some sense. Other workers in the field of small area estimation have used comparisons of this type to try to decide which set is best, but there are some serious problems with that approach that are discussed in detail in Appendix H. Very briefly, the standard definition of the best small area estimator is usually stated in terms of design-based mean square error. By definition, the design-based mean square error of an arbitrary estimator π_i is $\mathrm{MSE}_D(\pi_i) = \mathrm{Var}_D(\pi_i) + \mathrm{Bias}^2_D(\pi_i)$. The design-based estimates have large variance and almost no bias. Estimates based upon a lean fixed-effects model have very small variance and large bias. Composite estimates have average variance and bias. Estimates based upon mixed-effect models also have average variance and bias and can be applied to areas with no sample in them. Furthermore, procedures exist for estimating errors based upon mixed models that appear to have better validity than those based upon fixed effect models. Design-based variances can be estimated for each of the methods, but estimation of the design-based bias is not possible for composite estimates nor for estimates based upon mixed effect models. With no way to estimate the bias of these estimators, it is impossible to decide which of them is best. It was for this reason that the various estimates were compared across the small domains of interest primarily to see whether the expected dispersion patterns were realized.

Results

Exhibit 3.1 contrasts different estimators of past month alcohol use. ***Exhibits 3.2, 3.3*** and ***3.4*** do the same for past month cigarette use, past year drug treatment, and past year arrest. The first column of estimates are the design-based estimates for the 26 states with nontrivial NHSDA samples in them. Recall from Section 1.3 that these estimates are unstable and vary too much across the states. The second column of estimates shows the results of using mixed-effect models with survey weighting. As desired, the mixed models compressed the dispersion of the state estimates. The range, the standard deviation and the interquartile range are all smaller for all four behaviors examined in the four exhibits.

If the random effects are removed but all the same fixed effects are kept as in the full mixed model, then the state estimates are equal to the statistics shown in the third column of estimates. For most behaviors and measures of dispersion, using a fixed effect model instead of a mixed effect model results in greater compression of the dispersion across the states, as expected.

The next three columns of estimates show the results of applying three different fixed effect models. The fixed effect model with only county and demographic effects does not contain any tract level and block group level summary variables, but does include some additional

Exhibit 3.1 Relationship of Survey-Weighted Empirical Bayes Estimates of Past Month Alcohol Consumption to Design-Based Estimates and Estimates Based on Fixed Effect Models

	Design-Based NHSDA 1991-1993	Survey-Weighted Empirical Bayes (Mixed Model)	Estimators Based on Fixed Models			
			Same Fixed Effects as in Mixed Model	County and Demographic Effects Only	Separate Demographic Effects for Big City & Remainder	Demographic Effects Only
Total United States	53.01	53.46	53.43	53.01	53.41	53.40
California	57.69	56.67	58.07	56.05	53.08	52.55
Florida	49.67	48.45	48.52	49.75	52.43	52.78
Georgia	48.78	48.57	48.21	47.48	52.28	52.78
Illinois	55.73	54.43	55.26	55.48	55.18	53.27
Indiana	44.95	47.70	52.40	52.36	54.07	54.65
Kansas	60.82	56.51	54.33	56.40	54.10	54.60
Kentucky	32.03	41.18	44.26	40.97	54.18	54.79
Louisiana	56.62	49.40	44.13	44.66	51.53	52.02
Michigan	58.26	56.08	55.23	54.44	53.33	53.86
Minnesota	64.96	63.32	57.06	56.56	54.74	55.29
Missouri	44.10	54.04	52.22	54.71	53.49	54.07
New Jersey	61.10	59.94	62.03	60.32	52.62	52.91
New Mexico	56.21	53.98	51.57	57.40	52.41	51.88
New York	56.96	57.04	56.60	56.51	53.60	52.52
North Carolina	43.04	46.73	48.17	46.44	52.49	53.01
Ohio	50.45	52.24	53.18	55.30	53.56	54.14
Oklahoma	36.50	39.81	44.02	44.22	52.84	53.16
Oregon	59.72	55.95	54.83	59.81	53.97	54.45
Pennsylvania	52.70	55.82	53.75	55.90	53.39	53.96
South Carolina	47.03	46.84	46.67	41.34	51.83	52.36
Tennessee	35.76	40.70	45.10	39.07	53.06	53.65
Texas	55.23	52.88	48.80	50.07	53.10	53.06
Virginia	48.16	51.21	52.28	47.34	54.18	53.42
Washington	59.55	58.33	57.32	59.97	54.05	54.43
West Virginia	38.41	38.61	39.48	38.58	53.99	54.63
Wisconsin	67.92	59.15	55.95	56.31	54.38	54.94
Range	35.9	24.7	22.6	21.7	3.7	3.4
Standard Deviation	9.4	6.6	5.4	6.8	0.9	1.0
Interquartile Range	12.6	8.5	7.1	9.7	1.4	1.6
Correlation with Design-Based	--	0.917	0.769	0.769	0.172	0.020
Rank Correlation with Design-Based	--	0.916	0.804	0.835	0.212	0.063

Exhibit 3.2 Relationship of Survey-Weighted Empirical Bayes Estimates of Past Month Cigarette Use to Design-Based Estimates and Estimates Based on Fixed Effect Models

			Estimators Based on Fixed Models			
	Design-Based NHSDA 1991-1993	Survey-Weighted Empirical Bayes (Mixed Model)	Same Fixed Effects as in Mixed Model	County and Demographic Effects Only	Separate Demographic Effects for Big City & Remainder	Demographic Effects Only
Total United States	27.66	27.16	27.68	27.73	27.49	27.43
California	25.52	24.35	24.97	25.81	26.35	26.70
Florida	26.34	25.75	27.29	26.62	27.15	27.03
Georgia	28.41	28.55	28.81	29.18	28.59	28.14
Illinois	27.40	27.87	27.01	27.15	26.15	27.54
Indiana	24.65	26.04	28.20	28.13	28.14	27.77
Kansas	24.55	25.78	27.69	27.17	27.95	27.54
Kentucky	34.97	33.74	31.17	29.55	28.15	27.78
Louisiana	24.21	27.90	30.75	29.35	28.55	28.09
Michigan	29.45	28.78	29.89	29.41	28.26	27.84
Minnesota	25.42	24.16	26.30	28.99	28.11	27.70
Missouri	27.53	26.76	28.29	28.25	28.08	27.68
New Jersey	25.43	26.08	26.24	26.37	27.75	27.25
New Mexico	28.56	30.67	27.29	28.02	26.61	26.02
New York	24.18	25.13	26.57	26.06	26.10	27.24
North Carolina	30.49	28.26	28.69	28.59	28.43	27.97
Ohio	31.80	31.18	29.26	29.10	28.16	27.77
Oklahoma	25.85	29.00	31.09	29.36	27.87	27.20
Oregon	25.20	27.11	28.59	27.42	27.65	27.20
Pennsylvania	30.06	28.56	28.47	28.24	27.92	27.53
South Carolina	30.38	31.00	30.61	29.76	28.61	28.17
Tennessee	31.39	31.44	30.33	30.44	28.22	27.82
Texas	27.51	28.38	27.16	26.74	27.54	27.11
Virginia	28.14	26.49	27.16	27.81	27.47	27.98
Washington	25.00	25.16	27.33	28.27	27.92	27.37
West Virginia	34.39	32.67	31.58	29.50	27.75	27.40
Wisconsin	27.98	24.94	27.13	28.21	28.03	27.63
Range	10.8	9.6	6.6	4.6	2.5	2.2
Standard Deviation	3.0	2.6	1.8	1.2	0.7	0.5
Interquartile Range	4.5	3.1	2.6	2.1	0.6	0.6
Correlation with Design-Based	--	0.844	0.584	0.583	0.287	0.244
Rank Correlation with Design-Based	--	0.762	0.508	0.575	0.340	0.329

Exhibit 3.3 Relationship of Survey-Weighted Empirical Bayes Estimates of Past Year Drug Treatment to Design-Based Estimates and Estimates Based on Fixed Effect Models

			Estimators Based on Fixed Models			
	Design-Based NHSDA 1991-1993	Survey-Weighted Empirical Bayes (Mixed Model)	Same Fixed Effects as in Mixed Model	County and Demographic Effects Only	Separate Demographic Effects for Big City & Remainder	Demographic Effects Only
Total United States	0.62	0.70	0.70	0.68	0.65	0.64
California	1.04	0.97	0.92	0.77	0.65	0.65
Florida	0.47	0.69	0.73	0.69	0.60	0.60
Georgia	0.75	0.71	0.71	0.70	0.70	0.70
Illinois	0.46	0.54	0.59	0.53	0.63	0.65
Indiana	0.53	0.52	0.53	0.63	0.64	0.64
Kansas	0.48	0.66	0.66	0.63	0.63	0.63
Kentucky	0.41	0.57	0.57	0.66	0.63	0.63
Louisiana	0.35	0.64	0.66	0.61	0.70	0.70
Michigan	0.85	0.76	0.77	1.01	0.66	0.65
Minnesota	0.59	0.84	0.82	0.54	0.64	0.63
Missouri	0.78	0.70	0.64	0.79	0.63	0.63
New Jersey	0.42	0.68	0.68	0.67	0.64	0.63
New Mexico	0.25	0.77	0.80	0.76	0.64	0.61
New York	0.68	0.64	0.60	0.67	0.62	0.64
North Carolina	0.58	0.64	0.66	0.85	0.68	0.68
Ohio	0.34	0.70	0.70	0.66	0.64	0.64
Oklahoma	0.60	0.86	0.88	0.71	0.66	0.63
Oregon	0.69	0.90	0.88	0.80	0.60	0.59
Pennsylvania	0.45	0.56	0.60	0.62	0.62	0.62
South Carolina	0.39	0.53	0.54	0.79	0.69	0.70
Tennessee	0.27	0.60	0.60	0.62	0.65	0.65
Texas	0.65	0.61	0.66	0.68	0.66	0.65
Virginia	0.47	0.65	0.66	0.63	0.67	0.69
Washington	0.61	0.85	0.89	0.73	0.65	0.63
West Virginia	0.49	0.49	0.47	0.42	0.59	0.59
Wisconsin	0.19	0.61	0.63	0.50	0.63	0.62
Range	0.9	0.5	0.5	0.6	0.1	0.1
Standard Deviation	0.2	0.1	0.1	0.1	0.0	0.0
Interquartile Range	0.2	0.1	0.2	0.1	0.0	0.0
Correlation with Design-Based	--	0.536	0.446	0.509	0.065	0.070
Rank Correlation with Design-Based	--	0.472	0.391	0.509	0.037	0.131

Exhibit 3.4 Relationship of Survey-Weighted Empirical Bayes Estimates of Past Year Arrest to Design-Based Estimates and Estimates Based on Fixed Effect Models

	Design-Based NHSDA 1991-1993	Survey-Weighted Empirical Bayes (Mixed Model)	Estimators Based on Fixed Models			
			Same Fixed Effects as in Mixed Model	County and Demographic Effects Only	Separate Demographic Effects for Big City & Remainder	Demographic Effects Only
Total United States	1.57	1.64	1.62	1.66	1.61	1.60
California	1.90	1.90	1.68	1.74	1.58	1.65
Florida	1.66	1.54	1.48	1.72	1.50	1.51
Georgia	2.50	2.57	2.04	1.99	1.91	1.80
Illinois	1.35	1.58	1.46	1.56	1.38	1.62
Indiana	2.68	2.29	1.88	1.73	1.64	1.57
Kansas	2.46	2.37	1.84	1.73	1.62	1.55
Kentucky	1.31	1.59	1.69	1.77	1.63	1.57
Louisiana	1.47	2.26	2.14	2.06	1.89	1.78
Michigan	1.41	1.89	1.97	1.76	1.69	1.62
Minnesota	1.75	1.50	1.70	2.01	1.58	1.52
Missouri	1.70	1.83	1.91	1.69	1.61	1.54
New Jersey	1.18	1.18	1.23	1.26	1.63	1.55
New Mexico	1.79	2.43	2.02	1.87	1.65	1.54
New York	0.66	1.18	1.16	1.05	1.40	1.59
North Carolina	1.46	1.70	1.95	2.21	1.82	1.73
Ohio	2.03	2.07	1.96	1.81	1.63	1.56
Oklahoma	2.07	1.27	1.52	1.79	1.54	1.47
Oregon	1.67	1.70	1.66	1.55	1.46	1.41
Pennsylvania	1.07	1.21	1.26	1.19	1.56	1.50
South Carolina	1.15	1.71	1.90	1.66	1.91	1.80
Tennessee	1.10	2.07	1.60	1.54	1.69	1.61
Texas	1.92	1.77	1.68	2.03	1.85	1.73
Virginia	1.31	1.40	1.60	1.77	1.67	1.73
Washington	1.21	1.51	1.77	1.60	1.56	1.50
West Virginia	1.75	1.17	1.35	1.74	1.47	1.42
Wisconsin	1.91	1.29	1.74	1.95	1.60	1.54
Range	2.0	1.4	1.0	1.2	0.5	0.4
Standard Deviation	0.5	0.4	0.3	0.3	0.1	0.1
Interquartile Range	0.6	0.6	0.4	0.2	0.1	0.1
Correlation with Design-Based	--	0.560	0.468	0.545	0.178	-0.043
Rank Correlation with Design-Based	--	0.445	0.413	0.587	0.080	-0.079

county-level variables, particularly those associated with NDATUS and UCR (as described in Section 2.2). This model is interesting since it is less work to construct models with only county-level variables. Looking across the four tables, it appears that this model sometimes produced greater compression than the full fixed model and sometimes less. On balance, it does not appear that dropping tract and block group variables causes much, if any, stronger compression of the dispersion across the states than using the full fixed model (although of course the compression is stronger than using the full mixed model). This leads to the idea that perhaps for future applications, it would be satisfactory to use a mixed model where the only geographic fixed-effects were at the county level.

The fifth column shows the results from using a fixed effect model that includes only demographic effects except for one set of effects for the six oversampled metropolitan areas and a second set of effects for the remaining areas. This model resulted in very strong compression of the range and a rank correlation with the design-based estimates approaching zero. The simplest model with only demographic fixed effects was even worse.

Another estimator that would have been interesting to include in the tables is the design-consistent logistic regression estimator. The survey-weighted empirical Bayes estimates are asymptotically approximate composites of the estimates from the full fixed effect model with design-consistent logistic regression estimators. Thus, it is expected that the survey-weighted empirical Bayes estimates would lie between the other two. It is noted that while the survey-weighted empirical Bayes estimates frequently lie between the estimates from the full fixed effect model and the design consistent expansion estimators (simple Horwitz-Thompson estimates with some post-stratification), this is not always the case.

It is obvious from these tables that the survey-weighted empirical Bayes estimator performs at least some of the required compression of the dispersion across the states, but does it more lightly than is possible to do with a fixed-effect model. The correlations and rank correlations across the states indicate that using the modeling did change the ordering compared to using the design-based approach. The strength of the impact varies considerably across the behaviors examined. While it is hoped that the rank ordering was improved through the use of modeling, these tables do not contain adequate information to make such a judgment.

One could intently study the complete set of individual state statistics and try to make sense of them individually, but there is no statistical basis to support such analysis. For example, estimates of past-month alcohol use in Tennessee vary from 35.76% to 53.65%. Which is best? Even if individual estimates were determined to be bad in some sense, there would still be the possibility that the bad value was a fluke and that on average, the method does provide better estimates than the alternatives.

The only state where it makes sense to look at the individual values is California. The California sample size was over 12,000 people, so the design-based estimates are quite reliable. ***Exhibit 3.5*** highlights the comparisons of the California estimates. Note that an RSE is shown only for the design-based estimates. While all the columns are subject to sampling error, and the columns are correlated with each other (making testing of differences for statistical significance difficult), the RSE at least helps place the relative deviations of the model-based estimates in perspective. The full mixed model does best for arrest and drug treatment. It also is almost the best for alcohol use. It is just as bad for cigarette use as the simple fixed model with only demographic effects. The reason for this is not clear. On average, the full mixed model provides the closest agreement with the design-based estimates for California, as expected given the large sample size in California.

Comparisons of the national estimates are highlighted in ***Exhibit 3.6.*** Although the agreement between the survey-weighted empirical Bayes estimates and the design-based estimates at the national level are generally satisfactory, the agreement is not as close as was hoped would be the case, particularly for drug treatment in the past year. Even though the survey-weighted mixed model gives better consistency with design-based estimates at the national level than could be obtained with an unweighted mixed model, the consistency is not as good as can be obtained with a weighted fixed model. (The fixed models were weighted, so they are design consistent.) The problem is that even though the number of sample people is large, the number of sample PSUs is fairly small.

3.2 Comparison with Alternate Estimates for Artificial Domains

The second type of partial evaluation that was done was to compare the estimates from the survey-weighted empirical Bayesian method with alternative estimates on artificial domains where good agreement was expected. The idea behind this approach was that all good methods should produce similar estimates for large homogenous domains, where a homogenous domain is defined to be a group of people all of whom have the same propensity to engage in the behavior of interest. However, the study did not attempt to identify naturally occurring large homogenous domains upon which this theory could be tested. Instead, an idea due to Lemeshow and Hosmer (1982) was adapted to this application to create artificial domains that were large and expected to be fairly homogenous.

Exhibit 3.5 California Estimates

	Design-Based NHSDA 1991-1993	RSE on Design-Based	Relative Deviation from Design Based: Survey-Weighted Empirical Bayes (Mixed Model)	Estimators Based on Fixed Effect Models: Same Fixed Effects as in Mixed Model	County and Demographic Effects Only	Separate Demographic Effects for Big City & Remainder	Demographic Effects Only
Alcohol use	57.69	1%	-2%	1%	-3%	-8%	-9%
Cigarette use	25.52	2%	-5%	-2%	1%	3%	5%
Any illicit	8.43	3%	-2%	-7%	-14%	-31%	-32%
Any illicit but marijuana	3.96	4%	-1%	-13%	-8%	-36%	-36%
Drug treatment	1.04	9%	-7%	-12%	-26%	-38%	-38%
Arrest	1.90	7%	0%	-12%	-8%	-17%	-13%
Average relative deviation			-3%	-7%	-10%	-21%	-21%

Exhibit 3.6 U.S. Estimates

	Design-Based NHSDA 1991-1993	RSE on Design-Based	Relative Deviation from Design Based: Survey-Weighted Empirical Bayes (Mixed Model)	Estimators Based on Fixed Effect Models: Same Fixed Effects as in Mixed Model	County and Demographic Effects Only	Separate Demographic Effects for Big City & Remainder	Demographic Effects Only
Alcohol use	53.01	1%	1%	1%	0%	1%	1%
Cigarette use	27.66	1%	-2%	0%	0%	-1%	-1%
Any illicit	5.86	4%	-1%	2%	1%	0%	0%
Any illicit but marijuana	2.49	6%	-2%	-1%	1%	-1%	-1%
Drug treatment	0.62	11%	13%	13%	10%	5%	3%
Arrest	1.57	7%	4%	3%	6%	3%	2%
Average relative			2	3	3	1	1

deviation				

This methodology exploits the fact that as a result of fitting logistic regressions, each sample person in the three years of the NHSDA that were used to fit the models has a predicted propensity to exhibit the behavior of interest (e.g., smoke cigarettes, need treatment for substance abuse, etc.). By sorting the sample persons by the predicted propensity as in ***Exhibit 3.7***, and then cutting the list into a fairly small number of contiguous pieces of equal size, it should be possible to develop a partition that meets the desired criteria of large groups that are internally homogenous. Note that these groups don't represent any specific area or demographic domain; rather, each consists of people who, according to a model, have similar propensities to exhibit the characteristic of interest.

This methodology was used to partition the sample population into L mutually exclusive and exhaustive groups of equal size, each with a distinct average estimated propensity to exhibit the behavior of interest. According to the model used to create the partition, it is impossible to partition the sample into equal sized groups in a manner that would result in greater homogeneity within each group. The groups were ordered by increasing estimated propensity and indexed by d.

It was not expected that the various methods would agree exactly on each group due to variance on the estimates. A test was developed to determine whether the observed differences between the design-based estimates and each of the alternative estimators were significant. Two measures of agreement were also developed -- a correlation coefficient and a ratio of ranges. The construction of the test and measures of agreement is discussed in Section 3.2.1. The statistical properties of these statistics on the full sample are discussed in Section 3.2.2. Their properties on cross-validation samples are discussed in Section 3.2.3. The results of applying this method are presented in Section 3.2.4.

3.2.1 Construction of the Test and Measures of Agreement

Let π_d^F, π_d^M, π_d^W, and π_d^D be fixed-effect model estimates, mixed effect model estimates, survey-weighted empirical Bayes estimates, and design consistent estimates for the d-th large homogenous group. These were calculated as

$$\pi_d^F = \frac{\sum_{i \in d} w_i \pi_i^F}{\sum_{i \in d} w_i} \qquad \pi_d^D = \frac{\sum_{i \in d} w_i y_i}{\sum_{i \in d} w_i},$$

where π_i^F is the propensity to engage in the behavior of interest predicted by a fixed model for the i-th sample person and y_i is a binary flag indicating whether or not the person reported engaging in the behavior of interest. The average group propensities π_d^M andπ_d^W were calculated analogously toπ_d^F. Let π 50~}^{F} be the vector of the L estimated average propensities from the fixed model. The propensity vectors π 50~}^{M}, π 50~}^{W}, andπ 50~}^{D} were defined similarly.

The test for agreement between the estimates from the fixed effect model with the design-based estimates was calculated as

T=(π 50~}^{F}-π 50~}^{D})^{t}Ψ^{-1}(π 50~}^{F}-π 50~}^{D}),

where

Ψ=Var(π 50~}^{F}-π 50~}^{D})=E(π 50~}^{F}-π 50~}^{D})^t(π 50~}^{F}-π 50~}^{D}).

The tests for comparing the survey-weighted empirical Bayes estimates with the design-based estimates were constructed similarly. Tests of this form are called Wald tests. The details of how the variance-covariance matrix was calculated are given in Appendix G, Section G.4. Under assumptions that appear fairly reasonable, T has a chi-square distribution with about L degrees of freedom under the null hypothesis that the two sets of estimates are in good agreement with each other. The power of this test is only fair, with about 40 degrees of freedom for alcohol use and cigarette use and about 20 degrees of freedom for the other behaviors. Also, there were many replications of the test for different behaviors and age groups. Accordingly, attention did not focus on just whether any p-values below 0.05 were found. While a p-value below 0.05 is certainly strong evidence that the model does not fit the data, a preponderance of p-values below 0.5 would also indicate problems with the model. This can be seen by noting that if the two sets of estimates were in good agreement, p-values for the test should be randomly spread across the interval from zero to one with an expected value of 0.50. Small p-values would indicate that the two estimators do not have the same expected values even for large homogenous domains, a finding that would be troubling. However, there are some difficulties in the interpretation of T as are discussed further in the next section.

The correlation of the estimates from the fixed model with the design-based estimates was calculated as

$$\rho^F = \frac{\sum_{d=1}^{L}\left(\pi_d^F - \pi_d^D\right)^2}{\sqrt{\sum_{d=1}^{L}\left(\pi_d^F - \bar{\pi}_{.}^{F}\right)^2}\sqrt{\sum_{d=1}^{L}\left(\pi_d^D - \bar{\pi}_{.}^{D}\right)^2}}.$$

Similar definitions were used for ρ^M and ρ^W. The initial idea was that large correlations would indicate that the model fit the data well and that smaller correlations would indicate a less satisfactory fit. This correlation coefficient is a descriptive measure of model fit in the spirit of the multiple R^2 of linear regression. Since the design-based estimates are subject to more sampling error than the model-based alternatives, one would expect the correlations to be less than one even when the model fits. For this reason, there is uncertainty regarding how small the correlation should be to suggest serious lack of fit.

Let R^F equal the observed range of π_1^F to π_L^F. The ranges R^M, R^W, and R^D were calculated comparably. The range ratio for the fixed model relative to the design-based estimates was defined to be R^F/R^D. The range ratios for the mixed model and for the survey weighted empirical Bayes method were defined similarly. The initial idea was that range ratios close to one would indicate that the model fit the data well and that smaller or larger range ratios would indicate a less satisfactory fit. However, further review of this methodology has raised some questions about the interpretation of small range ratios when the variance of the characteristic under consideration across the domains is not large relative to the variance within the domains. (Since the domains were formed to be fairly internally homogenous, one would expect strong heterogeneity across the domains, but the degree of heterogeneity across the domains depends on the strength of the predictors in the model.) When the heterogeneity across groups is not large relative to the within-domain sampling variance, the design-based estimates will be overdispersed, resulting in a value of R^D that will tend to be larger than the true range. Since estimates based on a good model are designed to correct the overdispersion of the design-based estimates, it is not clear how much less than one the range ratio can be before serious lack of fit is suggested. These issues are discussed in the next section.

3.2.2 Statistical Properties of the Test and Measures of Agreement

In this section, there are some slight contradictions with the interpretation of evaluation results presented in the summary report (SAMHSA, 1996). Further review of the test and measures of agreement has established interpretation difficulties that can probably be ignored for common characteristics such as smoking and consumption of alcohol but that may be more serious for rare characteristics such as cocaine use. The Wald test is discussed first. The measures of agreement are then discussed under the simplifying assumption that the partition is fixed. Then there is further discussion of the consequence of the random nature of the partition when there is model overfitting.

Wald Test

The Wald test that was used in this study derives from a standard result in statistics. If *y* is a random vector of dimension *L* with a multivariate normal distribution with mean 0 and variance-covariance matrix Ψ, then

$$T = y^t \Psi^{-1} y \sim \chi^2(L).$$

A statistic with this form is called a Wald statistic. Wald statistics are a generalization of t-statistics to allow the simultaneous testing of whether a vector of (possibly correlated) random variables all have zero means. If the observed value of *T* is too large relative to what is expected for a statistic with a chi-square distribution, then it is concluded that one of the variables in the vector must have a nonzero mean. Even if *y* is not normally distributed and if Ψ is not known but has to be estimated, there are conditions under which *T* is still approximately distributed as a chi-squared variable with *L* degrees of freedom (Stroud, 1971). The quality of the approximation will depend on the severity of the nonnormality of the components of *y*, on the stability of the estimated variance-covariance matrix, and on the correlation between *y* and Ψ 65^?. The approximation is better when *y* is close to normal and Ψ 65^? is stable and independent of *y*.

For this application, *y* is replaced by either π 50~}^{F}-π 50~}^{D} or π 50~}^{W}-π 50~}^{D}, depending upon whether the test is being applied to estimated propensity vectors from a fixed-effect model or to survey-weighted empirical Bayes estimates. Thus, instead of testing whether E*y*=0, the test is for whether E(π 50~}^{F}-π 50~}^{D})=0 or E(π 50~}^{W}-π 50~}^{D})=0.

The assumption that π 50~}^{F}-π 50~}^{D} and π 50~}^{W}-π 50~}^{D} each has a multivariate normal distribution is reasonable since the groups were formed large enough for the central limit theorem to indicate that the weighted mean of the predicted or observed propensities in the group has an approximately normal distribution. The more serious problem was how to define the meaning of the expected value in the definition of Ψ. This is also an issue in defining what is meant by saying that the two estimators have the same expected values. There are four sets of random events with respect to which the expected value could be defined. The first is the random selection of persons for the NHSDA sample. The second consists of the set of state-level random effects. The third consists of the set of PSU-level random effects. The fourth consists of the person level random outcomes. In addition, it is possible to define the expectation as a conditional expectation, conditioning on various statistics and/or outcomes. Ideally, one should condition the expectation on as little as possible. It would have been most satisfying to have defined the expectation as over all possible samples, over the distribution of the state and PSU random effects, and over the distribution of person level random outcomes. However, the contractor was not able to derive a method of doing this within the time and budget available. Accordingly, something simpler was done. The variances were conditioned on the state and PSU random effects, on the estimated β vector, on the estimated set of random effects, and on the partition induced by the fitted model. This means that the expectation was just over all possible samples and over the conditional distribution of the person level random events given that state and PSU random effects are fixed **and** equal to the estimated state and PSU random effects and that the true β vector is equal to the estimated β vector.

A less conditional definition of the expectation in the equation for Ψ might have resulted in either larger or smaller variances. Focusing on the variance of π 50~}^{W}-π 50~}^{D}, less conditioning would have resulted in larger variances on π 50~}^{W} and on π 50~}^{D}, which would, in turn, have led to less significant test results since T would have tended to be smaller. On the other hand, less conditioning would also have resulted in the estimation of a positive correlation between π 50~}^{W} and π 50~}^{D}. A positive correlation would have reduced the variance on the difference between the two, thereby leading to more significant results since T would have tended to be larger. Since the results in section 3.1.4 do show strong correlations, the p-values for the Wald tests are probably not as small as they should be. This means that there is probably stronger evidence of lack of fit than what the tests showed.

Although it would probably have been possible with more work to derive the variance not conditioning on the random effects or on the estimated parameters, there was another difficulty. Since the partition is random, the distribution of π 50~}^{W}-π 50~}^{D} depends on the order statistics of the entire set of person-level predicted propensities. The effects of the random partition can be quite strong, particularly for rare characteristics if there was any overfitting of the model. It is doubtful that it would be possible to derive the variance of π 50~}^{W}-π 50~}^{D} without conditioning on the partition induced by the model.

Measures of Agreement

To understand the statistical properties of the measures of agreement across the *L* artificial domains, it is useful to have a model for the domain propensities. Indeed, without a model, it is impossible to make any inferences from those measures. However, the method for constructing the domains makes it difficult to conceive of an adequate model. As a first step, a model was developed assuming that the partition was fixed prior to fitting the model or even collecting the NHSDA data.

Let the *d*-th domain be labeled π_d. The model is that there is an underlying random process that leads to the propensity for each domain. This process is assumed to have mean π and variance φ, where $0<\varphi<\pi(1-\pi)$. This means that π_d is a random variable constrained to lie between 0 and 1, according to the model. Referring back to Section 1.4, this is a Bayesian interpretation of the propensity π_d. The model is rationalized by accepting that since it is not known why the propensity is different in each domain, it is best to treat the propensity in each domain as a random variable with some overall mean and variance. The parameter π can be thought of as the national average propensity to engage in the behavior of interest. The parameter φ quantifies the expected dispersion of the domain propensities from the national average. Since φ quantifies the variability in the process among the domains, it is referred to as the process variance. In a survey, there is interest in both estimating the national propensity and how much that propensity varies across domains. For example, there may be interest in identifying the domains with the highest and lowest realized propensities or in comparing the propensities for two specific domains.

The importance of the process variance depends on its size relative to the overall propensity. Clearly, a large process variance for either a rare or a nearly universal characteristic is more important than the same level of process variance for a characteristic that is neither rare nor nearly universal. This relationship is quantified through the intraclass correlation, defined as

$$\delta=\frac{\varphi}{\pi(1-\pi)} .$$

If the intraclass correlation is large, then the underlying process leads to sharp differences in group propensities. If the intraclass correlation is small, then the underlying process leads to only mild differences in group propensities. Since the idea of the partition is that propensity is uniform within each group, if the intraclass correlation is small, that says that everyone in the nation has about the same propensity to engage in the behavior of interest.

Generally, the desirability of close agreement between model-based and design-based estimates depends strongly on the intraclass correlation and on the sample size per group. Close agreement between the two sets of estimates is desirable only when the intraclass correlation is high and the sample size per group is large. This can be demonstrated by noting that the dispersion of the design-based estimates is too large unless both conditions are met. After demonstrating this assertion, implications are drawn for the interpretation of both the range ratio and the correlation coefficient.

The expected dispersion in the design-based domain estimates can be derived fairly easily. The dispersion of π_1^D through π_L^D depends on two sources of random variability. There is the random variation that caused π_1 through π_L to be different from π and there is sampling variance on each of the estimates given the true propensity for the domain. The first type of random variation depends only on φ while the second type depends on π, φ, and n_d, the direct sample size for the domain. Let E_M denote expectation with respect to the model and Var_M denote variance with respect to the model. Similarly, let E_D denote expectation with respect to the design and Var_D denote variance with respect to the design. Then the unconditional expected squared deviation of π_d^D from π over all possible samples of a given manifestation and over all possible manifestations of the superpopulation is

$$E_M E_D \left\{ \frac{1}{L} \sum_{d=1}^{L} \left(\pi_d^D - \pi \right)^2 \right\}$$

$$= \frac{1}{L} \sum_{d=1}^{L} \mathrm{Var}\left(\pi_d^D \right)$$

$$= \frac{1}{L} \sum_{d=1}^{L} \left\{ E_M \mathrm{Var}_D\left(\pi_d^D \mid \pi_d \right) + \mathrm{Var}_M E_D\left(\pi_d^D \mid \pi_d \right) \right\} \quad \text{by a basic statistical theorem}^{15}$$

$$= \frac{1}{L} \sum_{d=1}^{L} E_M \left[\frac{\pi_d\left(1-\pi_d\right)}{n_d} \right] + \frac{1}{L} \sum_{d=1}^{L} \mathrm{Var}_M\left(\pi_d \right)$$

$$= \frac{1}{L} \sum_{d=1}^{L} \left[\frac{\pi(1-\pi)-\varphi}{n_d} \right] + \varphi$$

$$= \frac{\left[\pi(1-\pi)-\varphi \right]}{n_d} + \varphi.$$

[15]See for example, Hansen, Hurwitz, and Madow, Volume 2, Section 6, Theorem 15.

The left hand side of the final line is the measurement variance due to sampling, while the right hand side is the process variance. Note that the sum of the two must be larger than φ by itself. This is the same as saying that the expected dispersion of π_1^D through π_L^D is greater than the expected dispersion of π_1 through π_L. The difference between the two expected dispersions is measurement variance. In other words, it is a mistake to think that true propensities for the domains are as spread out as the design-based estimates of those propensities unless measurement variance is negligible. ***Exhibit 3.8*** illustrates this with some numbers. For this example, the left most column shows invented propensities for four domains. Six simple random samples of 1,062 were then drawn from each domain. (This was a common domain sample size actually used in this study.) As the table shows, the standard deviation among the estimated domain propensities can be smaller than the true standard deviation but is more often larger -- sometimes far larger. Furthermore, when the estimated standard deviation is too large, then the range is also too large. If the estimated range of the design-based estimates is too large, then the ideal model-based procedure will produce a range ratio less than one.

By means of both the derivation of the total dispersion and the illustration, it is now clear that a range ratio of one is ideal only when measurement error is negligible. This can be achieved by having a large intraclass correlation and a large sample size per domain. In light of this discussion, it is evident that the best model is not necessarily the one that most closely tracks the values of the design-based statistics across the domains. If the variation in the process is large relative to the measurement variance, then close tracking to the design-based statistics is ideal, but otherwise the goal is to obtain a tighter set of statistics with a smaller population variance across the domains and a smaller range across the domains.

Exhibit 3.8 Illustration of Extra Variation in Domain Estimates Due to Sampling Error

Actual domain propensity	Estimated domain propensities based on simple random samples of 1,062 each					
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
3.91	4.32	3.87	4.45	4.35	5.64	2.83
3.54	3.89	3.72	3.82	3.21	2.77	3.97
3.31	4.21	3.85	2.87	2.40	2.93	3.47
3.20	3.70	2.98	3.52	2.94	2.45	3.42
Averages						
3.49	4.03	3.61	3.67	3.22	3.44	3.42

Standard deviations						
0.31	0.29	0.42	0.66	0.82	1.47	0.46

The effect of measurement error on the correlation coefficient is more difficult to determine. More theoretical work or perhaps simulations would be useful. Note that the Wald test accounts for measurement error. Thus, the overdispersion of the design-based estimates does not mean that the ideal model should produce estimates that demonstrate lack of fit.

Given the effect of measurement variance on the range ratio, the question then becomes, what is the ideal value for the range ratio? This research did not establish the ideal value. However, ***Exhibit 3.9*** does show the relative overdispersion of the design-based estimates for illustrative values of the intraclass correlation and for the group sample size. For example, if the intraclass correlation is $\delta=0.1$ (a large value) and the group sample size is 1,062, then the overdispersion in the design-based estimates is trivial. For such a situation, the ideal range ratio is equal to one, the ideal correlation coefficient is one, and the best model-based estimator would not shows signs of a lack of fit. On the other hand, when the intraclass correlation is just 0.01 and the sample size per group is just 266, then the overdispersion is 37 percent, meaning that the ideal model-based estimator will have 37 percent smaller dispersion across the domains than the design based estimator. Unfortunately, no measures of intraclass correlation were calculated for this study. Given a value for the observed variance in the design-based estimates across the groups and an overall mean, it would be fairly easy to compute the intraclass correlation, but this idea was only conceived as the final report was being written. At that point, there was inadequate time and funding to conduct more analysis.

Exhibit 3.9 Relative Overdispersion in Design Based Estimates by Intraclass Correlation and Group Sample Size

	Sample Size per Domain		
Intraclass correlation	1,062	531	266
0.100	0.8%	1.7%	3.4%
0.050	1.8%	3.6%	7.1%
0.025	3.7%	7.3%	14.7%
0.010	9.3%	18.6%	37.2%
0.005	18.7%	37.5%	74.8%
0.001	94.1%	188.1%	375.6%

However, it is noted here that intraclass correlation where the classes are defined to be states was estimated to be 3% for alcohol use, 0.4% for cigarette smoking, and less than 0.1% for drug treatment and arrest. When the model fits, the procedure used to form the artificial domains should maximize the intraclass correlation for a given number of the domains. Also, the number of artificial domains formed for alcohol and cigarette usage were 40 each per age domain. Only 20 artificial domains were formed per age domain for the other behaviors. Since the average sample size per age domain was 21,244, the sample size per artificial domain was 531 for alcohol and cigarettes and 1,062 for the other behaviors. (The column with 266 per domain reflects the situation when 20 groups are formed per age group for a one-quarter sample as was used for the cross-validation discussed in Section 2.2.3 below.) It is hoped that the intraclass correlations for the artificial domains should be at least 5-10%, indicating that range ratios for good models should be only slightly below one.

Model Overfitting and Random Partitions

Recall that all of the discussion of the measures of agreement so far have been preconditioned on the existence of a partition that was fixed prior to analyzing the data. The interpretation becomes more difficult when one considers the random nature of the partition. For a fixed partition, it is expected that the model-based estimators will have less variation across the groups than the design-based estimator. For a model-dependent partition, the opposite can easily be true due to model overfitting.

When the model contains inappropriate (nonsignificant) predictors, or predictors with poorly estimated coefficients, the model will tend to produce predictions for individual cases that are outliers; that is, cases with inordinately big or small propensity predictions. Such models are referred to as “overfit” since they include too many predictors. If there are only a small number of potential predictor variables, this problem does not arise since rigorous techniques exist to decide which of the predictor variables should be retained in the model. When, however, there are hundreds of potential predictor variables, as was the case on this project, those procedures tend to keep too many variables in the model.

Choosing prediction variables is an aspect of the statistician’s task that was discussed in Chapter 1. Recall that the goal of any modeling exercise is to find internally homogenous groups. In this study, all the potential predictor variables were categorical variables. Picking predictor variables for the logistic regression was therefore equivalent to finding groups of sample persons that were internally homogeneous with respect to the propensity to engage in the behavior of interest. By the same logic, finding the largest set of significant predictor variables was equivalent to searching for the smallest set of groups such that it was impossible to break any of the groups into subgroups with different true propensities. However, all that was observed for each person was a yes or a no (a 1 or a 0). As different ways to break a tentative group into subgroups were examined to see if their propensities were different, their observed rates of engaging in the behavior of interest were calculated. If the observed rates were different enough, then it was concluded that the true propensities for the subgroups must be different. Each time that this test is conducted, there is only a small probability that two subgroups with identical propensities will be mistakenly classified as having truly different propensities. However, when the test is repeated hundreds of times, the number of errors that are made of this type can become nontrivial. This problem can probably be ameliorated by having more stringent variable selection rules that minimize overfitting. For this study, predictors were kept in the models if their p-values were at least 0.05 using SUDAAN as described in Section 2.2.

Recall that the artificial partition was created by sorting on the predictions. When the model is overfit, all the outliers get grouped together -- the low outliers in the first group and the high outliers in the last group. As a result, on a partition created from an overfit model, the design-based estimator will tend to vary less across the groups than the model-based estimator that was used to create the partition. Note that this is the opposite relationship from what is expected for a fixed partition. As a result, the statistical properties of the range ratio on a model-dependent partition are sharply different from those on a fixed partition.

Let the true range R of propensities for a partition be defined as the range of π_1 through π_L for a particular manifestation of the superpopulation. Let R^F be defined as the range of π_1^F through π_L^F. Let R^M, R^W, and R^D be defined similarly. As discussed above, for a fixed partition,

$ER^D > R$ because of measurement error. For a fixed partition, a range ratio less than one ($R^F < R^D$) could thus mean either that the model is not rich enough or that R^F is less biased than R^D. On the other hand, a range ratio greater than one ($R^F > R^D$) is a strong sign that the model is too rich (i.e., overfit).

When the partition depends upon a model, it is expected that the positive bias in R^D will increase since the same data that led to the model also get used to calculate the π_d^D. The behavior of the range of the model-based estimates is more complex. The behavior of the range will depend strongly upon whether the partition is induced by the same model that yielded the model-based estimates and upon the extent of overfitting in the model that induced the partition. If the partition is induced by the same model that is being evaluated and that model is overfit, then the range statistic is likely to be inflated due to the impact of outliers. With worse overfitting, the inflation will also become worse. It is difficult to quantify the relationship between overfitting and inflation of the range, but range ratios greater than one are likely to be signals of overfitting.

If a very rich model is used to create a partition, and the model is capturing real signal rather than just generating noise, and a lean model is used to generate the estimated propensities, then the range ratio for the lean model-based estimators will be substantially less than one.

A correlation substantially less than 1 could signal a nonlinear monotone relationship between the model-based and design-based statistics but is more likely to signal either model lack-of-fit (either underfitting or overfitting) or large measurement variance in the design-based estimators. An examination of a graph can rule out a nonlinear monotone relationship, but it will generally be impossible to tell whether a value less than one means model lack-of-fit (underfitting or overfitting) or unstable design-based estimators. At any rate, values substantially less than one are not expected for a model-dependent partition since the data that led to the model also influence the design-based estimators, tending to create relatively high correlations.

3.2.3 Cross-Validation

Methods to assess the degree of overfitting in models do exist. One way is to divide the sample into groups prior to fitting the model, develop the model based on the sample in one group, and evaluate the fit using the other group. This method is quite rigorous with properties that are well understood. However, this method does not produce the best possible models since there is less sample available for identifying significant predictors and for estimating model parameters. For this study, it was decided to use cross-validation after the model fitting by dividing the sample into groups, refitting the model on one group and evaluating the refit model on the other group.

Choosing the group sizes was difficult since it was desired to have a large subsample both for the model refitting and for the evaluation of the refitted model. The subsample for the model refitting had to be large since the final models were too rich (i.e., contained too many variables) to fit on samples much smaller than the original sample. The subsample for the evaluation of the refitted model also had to be large because of the problems for the range ratio discussed above when the artificial domains are small and because the Wald test requires a large sample size to be valid and to have respectable power to detect lack of fit. Accordingly, it was decided to divide the sample into four equal parts, where each part contained a uniform slice of every state and sample PSU. For each model that was evaluated using this **cross-validation** procedure, the parameters were estimated for the final model using three of the four subsamples pooled together; then the fit was evaluated using the remaining subsample. This operation was repeated four times, so that all possible combinations of three out of the four subsamples were used for parameter estimation and each of the four subsamples were used for model evaluation. The four results were then averaged.

Given the already high computational burden of the model-fitting procedure, it was decided that it was not feasible to refit all 88 models. To keep the cost down, the cross-validation was done for only 12 of the models. These consisted of models for six of the eleven characteristics on the 26-34 age group, both inside and outside of the oversampled MSAs.

Note that the structure of a model was not changed when it was cross-validated, meaning that the set of fixed and random effects remained the same. Since the models were refit on a sample 25 percent smaller than the full sample, the refit parameters are less stable than the original parameters. This means that there are likely to be more outliers at the person level, meaning that range ratios will tend to be inflated. This is further exacerbated by the fact that only one-fourth of the original sample was available for the validation. Taken together, these two factors led to the expectation that the range ratios would be much larger for the cross-validated sample than for the original sample.

One of the improvements that cross validation offers when combined with the model-dependent partitions discussed in 3.2.2 is that some of the difficulties in interpreting range ratios and correlation coefficients are ameliorated. Specifically, the bias in R^D when computed on the reserved one-quarter sample is no worse than its bias on a fixed partition. It is still a biased estimator of *R* due to the measurement variance in the π_d^D, but because the partition is independent of the data used to compute R^D, the extra bias discussed above disappears. Offsetting that to some degree is the fact that the smaller sample sizes worsen the bias due to measurement variance. Which effect will be stronger depends on the situation. The range for the model-based methods, however, does not become any better behaved since the partition is still based on the predicted propensities for the units in the quarter sample. Therefore, the range of the model-based predictions is still expected to be biased upwards by the effect of outliers. The fact that the model parameters were fit on a separate data set (the other three quarters of the full data set) does not change this fact. It is still the case that model-based estimators will look quite different on partitions induced by themselves versus other models.

Another benefit of the cross-validation is that the estimated variance-covariance matrix Ψ 65^? has better validity. Recall that for the full sample, this variance was estimated as a conditional variance given the random effects, the fixed effects, and on the partition induced by the fitted model. This conditional definition meant that the covariance between π 50~}^{W} and π 50~}^{D} was estimated to be zero, even though it is clear that there is a substantial unconditional covariance between them due the fact that both use information in the *y* vector for each domain. In the cross-validation, the *y* vector for the one-quarter evaluation sample was not used in the estimation of π 50~}^{W}, meaning that the unconditional covariance is zero. Since the lack of a positive correlation between the two will tend to increase the size of π 50~}^{W}-π 50~}^{D}, this means that cross-validated Wald statistics are expected to be larger than the ordinary Wald statistics, thereby leading to small (more significant) p-values.

3.2.4 Results

Exhibit 3.10 shows the results of comparing the survey-weighted empirical Bayes estimates with the design-based estimates for the artificial domains described in Section 3.2.1. Each domain is predicted by the model to be fairly internally homogenous with respect to the propensity to engage in the behavior of interest. The correlations are high for most of the behaviors and age groups, as expected with a model-dependent partition of the population. A few of them are below 0.9, a possible sign of poor model fit. The p-values tend to range fairly

Exhibit 3.10 Tracking of Survey-Weighted Empirical Bayes Estimates with Design-Based Estimates Across Model-Homogenous Subgroups

Behavior	Statistic	Age Group 12-17	18-25	26-34	35 Plus	All Ages
Licit Drugs						
Past Month Cigarette Use	Correlation	0.917	0.952	0.952	0.946	0.978
	χ^2 probability	0.038	0.054	0.136	0.846	0.788
	Range Ratio	0.870	0.934	1.016	1.027	0.985
Past Month Alcohol Use	Correlation	0.871	0.960	0.972	0.981	0.990
	χ^2 probability	0.056	0.066	0.049	0.984	0.759
	Range Ratio	0.878	0.872	0.894	0.968	0.948
Illicit Drugs						
Past Month Any Illicit Drug Use	Correlation	0.939	0.973	0.962	0.942	0.990
	χ^2 probability	0.340	0.639	0.384	0.474	0.450
	Range Ratio	0.905	0.827	0.880	0.881	0.868
Past Month Any Illicit But Marijuana Use	Correlation	0.916	0.926	0.941	0.845	0.973
	χ^2 probability	0.611	0.089	0.774	0.566	0.609
	Range Ratio	0.923	0.801	1.028	0.821	0.879
Past Month Cocaine Use	Correlation	0.824	0.878	0.920	0.903	0.970
	χ^2 probability	0.864	0.094	0.803	0.611	0.607
	Range Ratio	0.856	0.740	0.880	0.889	0.849
Dependence						
Past Year Dependence On Illicit Drugs	Correlation	0.912	0.868	0.827	0.723	0.966
	χ^2 probability	0.871	0.778	0.625	0.670	0.537
	Range Ratio	1.031	0.882	0.894	1.147	1.000
Past Year Dependence On Alcohol	Correlation	0.787	0.898	0.943	0.927	0.978
	χ^2 probability	0.221	0.007	0.688	0.523	0.020
	Range Ratio	0.892	0.716	0.899	0.878	0.807
Treatment						
Past Year Treatment For Illicit Drugs	Correlation	0.908	0.772	0.918	0.884	0.962
	χ^2 probability	0.489	0.437	0.539	0.354	0.184
	Range Ratio	0.879	0.920	0.923	0.984	0.944
Past Year Treatment For Alcohol	Correlation	0.776	0.842	0.763	0.878	0.948
	χ^2 probability	0.428	0.680	0.612	0.037	0.159
	Range Ratio	0.978	0.748	0.899	0.814	0.826
Needing Treatment In Past Year	Correlation	0.899	0.910	0.923	0.934	0.980
	χ^2 probability	0.112	0.255	0.355	0.859	0.403
	Range Ratio	0.892	0.851	0.937	0.823	0.872
Arrest						
Past Year Arrested	Correlation	0.961	0.883	0.911	0.878	0.977
	χ^2 probability	0.623	0.225	0.902	0.589	0.416
	Range Ratio	0.778	0.883	0.952	1.123	0.950

*Probability of observing the calculated difference in the predicted and direct estimates across evaluation subgroups given that there is no difference.

uniformly over the interval (0,1), a good sign for acceptance of the model as discussed in the prior section. The actual median of the values in the table is 0.467. The range ratios are close to 1. Taken together with the high correlations, this is a sign that the two sets of estimates track closely over the artificial domains.

A somewhat different pattern emerges from a table of cross-validated statistics as discussed in Section 3.1.4. ***Exhibit 3.11*** shows the cross validation results for the Survey-weighted Empirical Bayes method and for several fixed effect models. Note that the correlations are sharply smaller than in ***Exhibit 3.10***, the p-values are usually slightly more significant, and range ratios are larger, all of which indicates a lack of agreement between the model-based and design-based estimators.

As discussed in Section 3.2.3, the attenuation of the cross-validated correlation coefficients and the increases in the cross-validated range ratios (both relative to the ordinary versions in ***Exhibit 3.10***) are expected consequences of the reduced domain sample sizes available for the evaluation. These smaller sample sizes lead to higher measurement error on the design-based estimates for the domains and also more instability in the random partition since the first and last groups will have fewer observations in them. For these descriptive statistics, there does not appear to be a good way of judging whether the cross-validated results are more dramatic than one would expect and are therefore symptomatic of lack of fit. Nonetheless, the large range ratios substantially greater than one are certainly suggestive of overfit models.

There are good theoretical reasons, on the other hand, to place more confidence in the cross-validated p-values from the Wald statistics than on the original p-values, as discussed in Section 3.2.3. Note that the p-values in ***Exhibit 3.11*** are mostly below the corresponding values for 26- to 34-year olds in ***Exhibit 3.10*** despite the lesser power of the cross-validated test given the smaller sample sizes for the cross-validation. This tendency illustrates the danger of excessive conditioning in the definition of Ψ. Since there are other random events upon which there is still conditioning and since not conditioning upon those random events would tend to increase variances, these p-values may be a little too small. Nonetheless, the fact that all the p-values in ***Exhibit 3.11*** are below 0.5 and some are substantially below 0.5 constitutes reasonably strong evidence of lack of fit. This is seen in all the columns, including the column for the mixed models. Evidence of lack of fit is disturbing since it impinges on the validity of the prediction intervals, but the impact on prediction interval coverage was not quantified in this study.

Exhibit 3.11 Cross-Validating the Survey-Weighted Empirical Bayes Estimates and Several Fixed Effect Models

Behavior	Statistic	Survey-Weighted Empirical Bayes (Mixed Model)	Estimators Based on Fixed Effect Models			
			Same Fixed Effects as in Mixed Model	County and Demographic Effects Only	Separate Demographic Effects for Big City & Remainders	Demographic Effects Only
Past Month Cigarette Use	Correlation χ^2 probability Range Ratio	0.765 0.015 1.095	0.737 0.007 1.084	0.622 0.008 0.956	0.514 0.000 0.440	0.584 0.000 0.272
Past Month Alcohol Use	Correlation χ^2 probability Range Ratio	0.866 0.280 0.969	0.858 0.231 0.841	0.832 0.108 0.966	0.824 0.023 0.685	0.839 0.001 0.524
Past Month Any Illicit Drug Use	Correlation χ^2 probability Range Ratio	0.728 0.043 1.618	0.704 0.036 1.500	0.659 0.132 1.038	0.573 0.080 0.510	0.637 0.109 0.310
Past Month Any Illicit Drug Use But Marijuana	Correlation χ^2 probability Range Ratio	0.636 0.412 1.450	0.615 0.249 1.451	0.392 0.242 1.213	0.212 0.131 0.256	0.297 0.131 0.153
Past Year Treatment For Illicit Drugs	Correlation χ^2 probability Range Ratio	0.588 0.287 1.420	0.481 0.275 1.509	0.407 0.386 1.348	0.532 0.294 0.278	0.561 0.297 0.219
Past Year Arrested	Correlation χ^2 probability Range Ratio	0.641 0.418 1.552	0.662 0.399 1.537	0.543 0.324 1.430	0.639 0.278 0.373	0.635 0.165 0.425
MEAN	Correlation χ^2 probability Range Ratio	0.704 0.243 1.351	0.676 0.199 1.321	0.576 0.200 1.159	0.549 0.134 0.424	0.592 0.117 0.317

Note: These tests were restricted to the 26- to 34-year-old age group due to the cost of computations.
*Probability of observing the calculated difference in the predicted and direct estimates across evaluation subgroups given that there is no difference.

An important methodological issue in the table concerns how the random partition of the data set was created for each column. For the first three columns, the same model was used to generate the partition as was being evaluated. For the last two columns, the full weighted mixed effect model was used to generate the partition. This accounts for most of the dramatic difference between the range ratios in the last two columns and the other columns. If the simple models in the last two columns had been evaluated with respect to the partitions induced by themselves, the range ratios would have been near one.

When a model is cross-validated with respect to the partition induced by itself, the range ratio is almost always above one. This is true for the full mixed effects model, for the model with the random effects removed but all the fixed effects retained, and for the model with the random effects and all of the tract and block group level fixed effects removed. This model with only demographic and county-level fixed effects is leaner than the full fixed model and yet there are still range ratios above one. This is probably good evidence that the models are overfit—even the models with just demographic and county-level variables. Note that the overfitting appears to be worse for the rarer behaviors.

Another interesting observation about the table concerns the low correlations, small p-values and small range ratios in the last two columns. Since the partition for these columns was induced by the full mixed model, this demonstrates that the full mixed model is capturing important information that the simple models cannot. The design-based estimator does vary significantly across the cells of the partition, and the simple models are not sensitive enough to mirror that variation.

3.3 Comparisons with External Data Sources

This approach is useful and important for any study where a census is occasionally conducted of the population of interest with respect to the behavior of interest. For example, this approach is often used to judge the quality of small area estimators for labor force statistics since the Decennial Census has several labor-related questions. Even in that application, however, there are difficulties since the questions are worded slightly differently than in the Current Population Survey and in other surveys that measure employment and unemployment. Systematic measurement biases tend to make evaluation of comparisons difficult.

This is all the more true with a topic as sensitive as substance abuse. It is known that reporting of substance abuse is extremely sensitive, not just to question wording, but also to the setting of the interview, the procedures that the interviewer uses, the assurances of confidentiality and so on. Furthermore, a substantial portion of the population that engages in the behaviors of interest (such as getting arrested) may not be accessible through a household survey since they reside in prisons at the time of interview or have such a tenuous relationship with any one household that they are not included in any household roster. Administrative databases are also subject to systemic biases. Also, the data collectors for administrative systems are often quite autonomous and thus may develop reporting idiosyncracies that compromise the comparability of the data across states and MSAs.

Despite the concern for the potential for differential nonsampling biases, comparisons were made between the NHSDA estimates and estimates from other sources. Data from three external sources were used: the Behavioral Risk Factor Surveillance System (BRFSS), the National Drug and Alcoholism Treatment Unit Survey (NDATUS), and the FBI's Uniform Crime Reports (UCR). Each of these external sources are first described. Then comparisons are made.

Behavioral Risk Factor Surveillance System (BRFSS): The BRFSS is a telephone survey conducted in all 50 States under cooperative agreements with the Centers for Disease Control and Prevention with no representation of persons living in households without telephones. State sample sizes are larger for many states than in the NHSDA, but different survey data collection agents are used in every State, making comparability across the states a difficult issue. However, definitions used for alcohol consumption and cigarette use are comparable to NHSDA definitions, since the BRFSS estimates reflect past month use. BRFSS estimates of alcohol consumption and cigarette use were compared to the survey-weighted empirical Bayes estimates. Studies have shown that reporting of substance use behaviors may be lower in telephone surveys than in face-to-face surveys, particularly for illicit drugs. The BRFSS State estimates are simple averages over the three years 1991 through 1993.

National Drug and Alcoholism Treatment Unit Survey (NDATUS): NDATUS is an inventory of all specialty substance abuse treatment facilities in the U.S. Based on reporting by State substance abuse agencies, it provides estimates (including adjustments for nonresponse) of the number of clients in treatment at a given point in time. NDATUS estimates of drug treatment volumes were compared to the survey-weighted empirical Bayes estimates. To develop an estimate of persons treated during a year, the NDATUS client counts (including drug only and combined drug and alcohol clients) were multiplied by the reciprocals of average lengths of stay, and adjusted to account for multiple treatment episodes in a year by the same individual. Estimates of length of stay and multiple episodes were obtained from the Drug Services Research Survey, conducted in 1990. These calculations were done within categories of treatment modality and applied separately to each State. No adjustment was made to account for the inclusion in the NHSDA estimates of persons reporting treatment through self-help groups, private physicians, or emergency rooms, none of which are counted in NDATUS. The State estimates are averages over only 1992 and 1993 since the 1991 estimates could not be adequately adjusted for nonresponse.

Uniform Crime Reports (UCR): The UCR compiles data from local jurisdictions on the number of arrests. UCR estimates of past year arrests were compared to the survey-weighted empirical Bayes estimates. For comparison with the NHSDA estimates, an adjustment to the UCR data was made to account for persons arrested more than once during a year, so the adjusted UCR estimates reflect number of persons arrested at least once. This adjustment was made within the four Census regions, using data on multiple arrests reported by arrestees in the NHSDA sample. The State estimates are simple averages over the three years 1991 through 1993.

Exhibit 3.12 summarizes the results of the comparisons across the 26 states for which NHSDA estimates were prepared. One positive finding is that the ratio of the range of the design-based NHSDA estimates across the 26 examined states was almost always smaller than the range in the external estimates across the same states for three out of the four characteristics. Similarly, relative standard error across the states is smaller with the design based estimates than with external estimates for two out of the four characteristics. Given the earlier discussion in this chapter about the impact of measurement error on the ranges and dispersion, the expectation had been that the design-based estimates would be more dispersed than the external estimates. The fact that this is not true constitutes reasonable evidence that there are important differences in nonsampling errors between the data systems.

Turning attention to the other range ratios in ***Exhibit 3.12***, one observes the same sorts of patterns as in ***Exhibits 3.1*** through ***3.4***. The survey-weighted empirical Bayes estimates always had a tighter range across the states than the design-based NHSDA estimates, as desired. Whether the reduction in range is too much or too little is impossible to say, but it is clear that all the estimators based on fixed models almost always produce greater shrinkage than the mixed effect model, as expected. The simple demographic model, with or without a separate set of effects for oversampled metropolitan areas, seems to produce clearly undesirable levels of

Exhibit 3.12 Comparison of Alternative Small Area Estimators to Estimates from External Sources

Comment [COMMENT1]: 4/12 --- tables corrected per Mike Witt's request and faxed to client

State	Estimate from External Source	Survey-Weighted Empirical Bayes (Mixed Models)		Estimators Based on Fixed Effect Models				
		With Full NHSDA Sample	Big City Sub-Sampled Model	Same Fixed Effects as in Mixed Model	County and Demographic Effects Only	Separate Demographic Effects for Big City & Remainder	Demographic Effects Only	Design-Based 91-93 NHSDA Estimates
Alcohol Use	BRFSS							
National Estimate	50.90	53.46	53.59	53.43	53.01	53.41	53.40	53.01
Rank Correlation[1]	.	0.861	0.854	0.841	0.765	0.278	0.097	0.807
Range Ratio[2]	1.000	0.597	0.598	0.545	0.525	0.088	0.082	0.867
RSE[3]	21	12	13	10	13	2	2	18
Cigarette Use	BRFSS							
National Estimate	23.10	27.16	27.27	27.68	27.73	27.49	27.43	27.66
Rank Correlation	.	0.491	0.610	0.631	0.670	0.485	0.472	0.473
Range Ratio	1.000	0.930	0.786	0.642	0.449	0.244	0.208	1.048
RSE	10	10	8	6	4	3	3	11
Drug Treatment	(NDATUS)							
National Estimate	0.85	0.70	0.71	0.70	0.68	0.65	0.64	0.62
Rank Correlation	.	0.375	0.523	0.406	0.289	-0.160	-0.240	0.063
Range Ratio	1.000	0.349	0.326	0.324	0.432	0.079	0.080	0.615
RSE	38	18	16	17	18	4	5	31
Arrest	(UCR)							
National Estimate	4.10	1.64	1.69	1.62	1.66	1.61	1.60	1.57
Rank Correlation	.	0.350	0.351	0.389	0.356	0.510	0.450	-0.066
Range Ratio	1.000	0.361	0.415	0.252	0.298	0.136	0.100	0.518
RSE	27	26	24	16	16	9	7	30

[1] Correlation calculated by first ranking states and then calculating the correlation of the ranks.
[2] Ratio of the range of the predicted values to the range of data from the external source.
[3] Relative standard error among state estimates as a percentage.

Note: Estimates of prevalence rates have been multiplied by 100.

Source: SAMHSA, Office of Applied Studies, National Household Survey on Drug Abuse. Model-based estimates using 1991-1993 NHSDA data.

shrinkage. The range ratios for the county demographic model for drug treatment and arrest are higher than for the full fixed model. This is a sign that it is possible to induce as much variation in state estimates of drug treatment and arrest with just county-level variables as with a combination of county, tract, and block-group level variables. Whether the extra induced variation is good or not constitutes an open question.

The data on rank correlations are difficult to interpret. First, since it is not known which set of estimates is best, it is hard to say whether a high rank correlation is desirable. Second, the patterns are not consistent across the behaviors. The full mixed model resulted in the highest rank correlation with the external data source only for alcohol. The very lean fixed effect models (basic demographic and basic demographic with big city effects) generally produced the lowest rank correlations, but by some fluke, produced the highest rank correlations for arrests.

It is interesting that the design-based estimates for drug treatment and arrests show essentially no correlation with the NDATUS and UCR data. This would appear to indicate that the two systems are measuring different phenomena while attaching the same name to those phenomena. The fact that the models with county level predictors have substantial rank correlations with NDATUS and UCR is probably due to the fact that variables from these databases were in those models.[16] It is not clear though how those variables got picked for inclusion in the model, when at the state level there is no correlation. It may be that these variables are part of the overfitting problem. The fact that the NDATUS and UCR statistics vary more across the states than the design-based estimates despite the fact that they are not subject to sampling error and the fact that there is no rank correlation between them and the design-based estimates raises some concerns about the appropriateness of variables from these databases as predictors in the models. Although these variables may have better face validity than census variables for some data consumers, the potential that they are not consistently collected across the states means that their presence in the models could be undermining the purpose of creating a set of state estimates that will fairly rank the states.

The column for a big city subsampled model describes the results when the survey-weighted empirical Bayes procedure is applied to a subset of the 1991-1993 NHSDA. This subsample was drawn by dropping the oversample of the six targeted metropolitan areas. Dropping the oversamples (so that the areas were sampled at national rates) allowed the fitting of

[16]The included predictor variables were not, however, identical to the outcome variables of interest here. Otherwise, it would have been possible to get a perfect fit and have a rank correlation of 1.

a single mixed model for each age group and behavior instead of having to fit separate models for the big cities. The fact that the two survey-weighted empirical Bayes columns are so similar indicates that the oversample in the six MSAs had little effect on most state estimates.

3.4 Ideas for Future Evaluations

The evaluation focused on trying to determine whether the model-based estimates were more accurate than the design-based estimates. This effort was partially successful, and it is hard to see what more could be done. While it would be interesting to compare the variances of the model-based estimates with those of the design-based estimates, the variances are almost certainly smaller with the model-based approach. The reason that this comparison was not done is that the estimates of the variances on the design-based estimates for states and MSAs would be highly unstable. It would be difficult to make the comparisons. Also, there are questions about underestimation of the variances for the model-based estimates.

One area, however, where it does appear to be possible to do more concerns the coverage of the prediction intervals, both when the model is true and when the model is not true. Two ideas are presented in this section on how the coverage could be evaluated. These methods were not used in the study since they cannot be used to evaluate the accuracy of the small area estimates themselves, but they could be useful for assessing coverage properties.

Since the labor force items are fairly consistent between the Decennial Census and the Current Population Survey, a good test of the methodology for estimating the variances for the model-based estimates would be to prepare model-based estimates and their estimated variances using variables from the CPS from April of 1990 as outcome variables and variables from the 1980 Census as predictor variables. (It is possible to link the 1990 CPS back to the 1980 Census.) By using ten-year-old predictors, it would be possible to simulate the effects of having imperfect predictors. Another advantage of such an approach is that it would have an evaluation data set completely independent of the data set used to create the predictions. Such a study would be useful for assessing the coverage properties of the prediction intervals. The actual mean squared error reductions that would be achieved for labor force statistics could, however, be quite different from the reductions expected for NHSDA statistics.

If the logistics of matching 1980 Decennial Census data to 1990 CPS data are too difficult,[17] the coverage of the prediction intervals could also be assessed through a simulation study. For such a study, it would be necessary to specify a model for the U.S. population and then to create a simulated population of very large size. Various models could be utilized to simulate the population. The methods could then be tested using either the same model structure that was used to create the simulated population or some other model. Although it might seem to be a weak test of the methodology to evaluate it assuming that the modeler knows the correct model structure, it would be of interest for the NHSDA project since approximations were used in fitting the model and in estimating the variances for the estimates. Without simulation studies, the effect of these approximations remains unknown.

Using a different model structure in the modeling than in the population simulation constitutes a much more stringent test of a methodology, but such an approach requires more judgment to design and interpret. If the model used to simulate the population is sufficiently pathological, then no reasonable method can be designed. This is even true of design-based methods since these methods assume that domain totals are normally distributed across all possible samples. It is quite difficult to decide how severe a pathology to use in testing.

One possibility to avoid having to create the population is to assemble more years of old NHSDA sample and to treat the conglomeration as a population from which samples are drawn. The problem with this approach would be the population sizes of the clusters. Fairly large clusters are required for representative testing.

A related idea would be to rerun the model on only a subsample of the 1991-1993 NHSDA, dropping out large states like California from the modeling in order to see how well the model works in predicting the behavior rates for those states when no sample was available. Although this idea has the appeal of being simple, it seems like a statistically unstable decision method. By this, it is meant that luck of the draw could easily result in the method either being accepted or rejected. It is sort of similar to making a major decision based on a single coin toss. A more complex simulation program would be more costly but would also yield much more information for the decision.

[17]It is likely that the Census Bureau would insist on performing the match itself and on adding noise to the prediction variables to protect confidentiality.

3.5 Summary of the Evaluations

As noted previously, all of the evaluation measures that were carried out have limitations. However, some useful findings do arise from the evaluation. First, the California estimates agree well with the design-based estimates. The method was designed to provide such agreement for domains with large sample sizes, so even though this says nothing about the quality of estimates for domains with small sample sizes, it is at least comforting that the method functioned as intended. National estimates agreed fairly well with the design-based estimates but not as well as hoped. Estimates based on simple fixed models actually fit better at the national level for several of the behaviors of interest. This is because weights were used in the fixed models, making estimates based on this design consistent, and because the number of sample PSUs was not large enough to make the national aggregates from the mixed models converge to the design-based estimates.

Second, the survey-weighted empirical Bayes estimates displayed less dispersion across the states than the design-based estimates and more dispersion than the estimates based only upon fixed effects. The method was designed to provide this sort of compression of the dispersion, so this outcome is not surprising, but it is again comforting that the method functioned as intended.

Third, there is some indication that the fixed parts of the models were overfit. Since overfitting results in negatively biased variance estimates, it might have been better to set a higher hurdle for fixed effects to enter the model. On the other hand, there is also evidence that there was genuine predictive value in at least some of the county, tract and block group level summary data, so it would have been a mistake to set the hurdle so high that only demographic variables could have entered the model. For this study, the maximum p-value for a variable to be retained was set at 0.05. It might have been better set at 0.01.

Fourth, estimates from BRFSS, NDATUS and UCR for alcohol use, cigarette use, drug treatment, and arrest vary as much or more than the NHSDA design-based estimates across the states. Since theory indicates that the NHSDA design-based estimates vary too much across the states, it might be reasonable to conclude that the estimates from the other sources also vary too much across the states. If this is true, then it might have been better not to use NDATUS and UCR county-level variables in the modeling. Variables from the decennial census may not appear as relevant and can be updated only once every ten years, but at least they are collected consistently across the states.

These findings verify that the methodology is basically functioning as intended. That is reassuring in the sense that it appears to rule out major conceptual or programming errors. The findings also indicate small ways in which the methodology can be improved. However, open questions remain about the reduction in variance achieved through the use of modeling and whether the prediction intervals have the claimed coverage levels. More research is needed to prove that these objectives were actually attained.

4. Thoughts for Future Applications

There are a number of important areas for further research. Most critically, work needs to be done on improving the variance estimates and then using simulation studies to demonstrate that the revised prediction intervals have good coverage properties, thereby establishing the validity of the methodology. Once more accurate variance estimates are obtained, it would also be important for planning purposes to determine the relationship between the variances on survey- weighted empirical Bayes estimates and design-based variances. This information could then be used to predict the width of model-based prediction intervals prior to actually preparing the estimates. Additionally, there is room for improvement in the point estimates themselves. Finally, it is important to consider the quality of the change estimates that would be obtained from model-based estimates. Each of these issues is discussed in turn below.

4.1 Improvement of Estimated Variances

There were several problems in the estimation of variance. First, a small component of variance due to sampling PSUs was inadvertently omitted. The previously published estimates reflect only the variance of the fixed part of the model in nonsample PSUs, neglecting the variance due to assuming a zero random effect for all nonsample PSU, as explained in detail in Appendix G, Section G.1. Revised variance estimates have been prepared that show a modest impact of the error on most estimates. This correction would need to be routinized for future applications.

Second, the variance due to estimating components of variance was intentionally omitted due to the lack of theory on how to properly reflect this component of variance. There are some ideas on how to proceed in this area, but progress will not be easy. As a temporary measure, one could approximate an upper bound on this effect and then use that upper bound, thereby providing prediction intervals with conservative coverage properties; i.e., one would deliberately overestimate the mean square error so that when 95% coverage is claimed, the actual coverage would be higher than 95%. For a better solution to this problem, it might be necessary to adopt a fully Bayesian approach to the small area estimation. This would require developing an entirely new software system and documentation. Since such development would be costly, it would only be recommended if simulation studies indicated that neglecting the variance on the estimated variance components was seriously affecting the coverage of the prediction intervals and that a simple upper bound did not remedy those coverage problems.

Third, estimated variances for domains that span the age groups are too low because the methodology ignored the positive covariance between the age groups across areas. There was a plan to approximate this effect, but it could not be implemented due to budget constraints. If one could fit a single model for all the age groups, this problem would disappear, although one would need to rewrite sections of the software to maintain separate random effects by age group.

Fourth, there is some evidence that the models were overfit. This means that too many predictor variables were used. If this is true, this also resulted in underestimation of the variances. The problem is that when the model is overfit, all of the estimated random effects tend to be too small. Since the variance estimate involves a term with the sum of the expected squared random effects, the overfitting thereby results in variance estimates that are too small. This problem can be fixed by keeping the models leaner.

4.2 Validation

Validation is an intrinsically difficult task with small area estimation unless there is an occasional census on the variables of interest using the same measurement techniques. Short of a census, simulation studies are the best that can be done to assess the coverage properties of the prediction intervals. The internal measures that were done for this study were innovative and interesting, but don't provide sufficiently strong evidence of validity. Comparisons with other surveys are of limited value because methodologies across surveys differ and because each has sampling error.

Various simulation studies could be designed. These can be grouped into two broad categories. The first category includes studies in which the simulated population is created according to the same model that will be used to analyze it. The second category includes studies in which pathologies are introduced into the population so that the model applies only approximately. Although the first category might appear to be an easy test of a method, it would be of interest for the NHSDA project since approximations were used in fitting the model and in estimating the variance for the estimates. Without simulation studies, the effect of these approximations is unknown.

The second category of simulation study is a much more stringent test of a methodology, but requires more judgment to design and interpret. If the pathologies are severe enough, then no reasonable method can be designed to cope with them. This is even true of design-based methods. It is quite difficult to decide how severe a pathology to use in testing.

Ideally, both sorts of simulation studies would be conducted. The second category should, of course, be delayed until the results of the first category can be studied. There is no reason to test the method on pathological pseudo populations until it has passed the test on model-conforming pseudo populations.

4.3 Reduction in Variance Due to Modeling

Before accurate predictions can be made of the variance reduction achievable through the SAE methodology, it is necessary to improve the estimates of variance as discussed above. The correction of the problem regarding a neglected bias-squared term is particularly important. After making that correction, however, more theoretical work is required since the variance of the model-based estimates is not strictly inversely proportional to the sample size. As the sample size increases, the reduction in variance due to modeling is likely to decrease. A reasonable guess is that the reduction is somewhere in the range of 10 to 75 percent. More time and study would be required to make more precise projections.

4.4 Improvement of Point Estimates

The point estimates appear to be of good quality. However, the quality could probably be improved by fitting leaner models. The overfitting mentioned earlier has several consequences for the quality of the point estimates. First, the small area estimates are not shrunk quite as close together as they should be. Another outcome of overfitting is that resulting computer time problems force one to run separate models by age. If a single model could be run including all ages, then the parameters for effects that do not interact with age would be estimated more accurately. A third outcome is that the ordering of the states may not be as stable as it could be. For all three reasons, the models should not be allowed to be as rich on a repetition of the project.

One of the ways that this could be achieved is to reverse the order of elimination for county-level summaries vis-à-vis tract or block-group level summaries. Before, the county-level summaries were only admitted to the model if they were significant after accepting the tract and block-group level summaries for the same characteristics. It may be better to let the county-level variables in first and then to accept the tract and block-group level summaries if they are marginally significant. It may also be better to require higher significance for each variable prior to admitting it to the model, perhaps $\alpha = .01$ or even smaller instead of $\alpha = .05$.

4.5 Validity of Change Estimates

Estimates of change can be made by using the survey-weighted empirical Bayes procedure to prepare estimates for each time period and then forming the differences between the two sets. Estimates of change made in this manner will be more stable than estimates of change made using purely design-based methods. However, the model-based estimates of change for particular states reflect a blend of changes at the state and national level. Particularly for states with small sample sizes, the estimates are likely to mostly reflect nation changes. The validity of these change estimates will depend on the extent to which the national changes are uniform within the cells defined by the fixed effects in the model.

4.6 Sample Design

In this study, the methodology for small area estimation was developed after the sample had already been designed and the survey had been conducted. The 1991-1993 NHSDA was designed to be nearly optimal for national estimates by race-ethnicity domain by age group. If the objective of state specific estimates is fixed in advanced, it is possible to develop a sample design that is better suited to the objective. Specifically, if state estimates were made an objective, the number of sample PSUs should be increased, the PSUs should be stratified by state, and the sample of people should be more equally allocated across the states instead of being massed in the large states. Since all of these design features would be less than optimal for national estimates by race, ethnicity and age, deciding whether to adopt these features is a difficult question.

References

Breslow, N. E. and Clayton, D. C. (1993). "Approximate Inference in Generalized Linear Mixed Models." Journal of the America Statistical Association, 88, 9-25.

Breslow, N. E. and Day, N. E. (1980). Statistical methods in cancer research; Volume 1 -- The Analysis of Case-Control Studies. Lyon: International Agency for Research on Cancer.

Cox, D. R. And Reid, N. (1987). "Parameter orthogonality and approximate conditional inference." Journal of the Royal Statistical Society, Series B, 49, 1-39.

Cox, D. R. and Snell, E. J. (1989). Analysis of Binary Data, 2nd edition. London: Chapman & Hall.

Cramér, H. (1946). Mathematical Methods of Statistics. Princeton: Princeton University Press.

Diggle, P. J., Liang, K. Y., and Zeger, S. L. (1994). Analysis of Longitudinal Data. Oxford: Clarendon Press.

Fellner, W. H. (1986). "Robust estimation of variance components." Technometrics, 28, 51-60.

Fellner, W. H. (1987). "Sparse matrices, and the estimation of variance components by likelihood methods." Communications in Statistics B, 16, 439-463.

Ghosh, M. and Rao, J. N. K. (1994). "Small Area Estimation: An Appraisal." Statistical Science, 9, 55-93.

Gonzales, M.E. (1973). "Use and evaluation of synthetic estimates. Proceedings of the Social Statistics Section. American Statistical Association, 33-36.

Graybill, F. A. (1969). Introduction of Matrices with Applications in Statistics. Belmont, CA: Wadsworth.

Green, P. J. (1987). "Penalized likelihood for general semi-parametric regression models." International Statistical Review, 55, 255-277.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1993). Sample Survey Methods and Theory, Volume II: Theory. New York: Wiley Classics.

Harville, D. A. (1974). "Bayesian inference for variance components using only error contrasts." Biometrika, 61, 383-5.

Harville, D. A. (1977). "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems." Journal of the America Statistical Association, 72, 320-340.

Holt, D., Smith, T.M.F, and Winter, P.D. (1980). Regression analysis of data from complex surveys. Journal of the Royal Statistical Society, A143, 474-87.

Lemeshow, S. and Hosmer, D. W. (1982). "The use of goodness-of-fit statistics in the development of logistic regression models." American Journal of Epidemiology, 115, 92-106.

McCullagh, P. and Nelder, J. A. (1989). Generalized Linear Models, 2nd edition. London: Chapman & Hall.

McGibbon, B. And Tomerlin, T. J. (1989). "Small area estimation of proportions via empirical Bayes techniques." Survey Methodology, 15, 237-252.

Patterson, H. D. And Thompson, R. (1971). "Recovery of interblock information when block sizes are unequal." Biometrika, 58, 545-554.

Rao, J.N.K. and Scott, A.J. (1981). "The Analysis of Categorical Data from Complex Sample Surveys: Chi-Squared Tests for Goodness of Fit and Independence in Two-Way Tables." Journal of the American Statistical Association, 76, 221-230.

Robinson, G. K. (1991). "That BLUP is a Good Thing: The Estimation of Random Effects." Statistical Science, 6, 15-51.

SAMHSA (1996). Substance Abuse in States and Metropolitan Areas: Model Based Estimates from the 1991-1993 National Household Surveys on Drug Abuse Summary Report. Rockville, MD: SAMHSA, Office of Applied Studies.

Särndal, C-E, Swensson, B., and Wretman, J. (1992) Model Assisted Survey Sampling, New York: Springer-Verlag.

Schaible, W. (Ed) (1996). Indirect Estimators in U. S. Federal programs. New York: Springer-Verlag.

Schall, R. (1991). "Estimation in generalized linear models with random effects." Biometrika, 78, 719-727.

Smith, T. M. F. (1994). "Sample Surveys 1975-1990; An Age of Reconciliation?" International Statistical Review, 62, 3-34.

Stiratelli, R., Laird, N., and Ware, J. (1984). "Random effects models for serial observations with binary response." Biometrics, 40, 961-971.

Stroud, T.W.F. (1971). "On obtaining large-sample tests from asymptotically normal estimators." Annals of Mathematica Statistics, 42, 1412-1424.

Appendix A: Classification of Drug and Alcohol Dependence

CLASSIFICATION OF DRUG AND ALCOHOL DEPENDENCE

The Diagnostic and Statistical Manual of Mental Disorders, Revised Third Addition (DSM-III-R), published by the American Psychiatric Association is designed to be used by clinicians as well as researchers for making diagnosis of psychiatric disorders. Substance abuse and dependence are considered psychiatric disorders under DSM-III-R. DSM-III-R defines a person as dependent on a substance if they meet 3 out of 9 criteria for that substance. The method for estimating dependence from the NHSDA is based on the NHSDA questionnaire items that are assumed to approximate five of the nine DSM-III-R criteria. These five criteria include: unable to cut down on use; reduced social, occupational, or recreational activities; continued use despite knowledge of having a problem; tolerance; withdrawal. In the NHSDA algorithm, respondents are defined as dependent on a substance if they responded affirmatively to at least two of the five NHSDA criteria for that substance.

To evaluate this method for estimating dependence, NHSDA dependence estimates were compared to estimates from the National Comorbidity Survey (NCS).[1,2] Conducted in 1991, the NCS was designed to provide nationally representative estimates of 14 psychiatric disorders, including substance abuse and dependence, using a structured diagnostic interview. It employed a multistage area probability sample of 8,098 respondents in the household population.

Based on this analysis it was concluded that the approximation to the DSM-III-R definition of drug dependence produces dependence estimates which are comparable to the NCS. The table below indicates estimates of dependence from the NHSDA from 1991 through 1993 and estimates of dependence from the NCS.

Twelve Month Estimates of Dependence from the NCS and NHSDA
(Old and New Methods) for Persons 15-54 Years of Age (1991-1994)

DRUG	NCS	NHSDA		
	1991	1991	1992	1993
Any Illicit	1.8%	1.8%	1.4%	1.5%
Marijuana	0.9	1.2	0.9	0.9
Cocaine	0.4	0.5	0.4	0.3
Alcohol	4.4	5.7	4.7	5.2

[1]This is included in a paper by Joan Epstein and Joe Gfroerer, A Method for Estimating Substance Abuse Treatment Need from a National Household Survey, that was presented at the 37th International Congress on Alcohol and Drug Dependence, August 20-25, 1995.

[2]Lifetime and 12-Month Prevalence of DSM-III-R Psychiatric Disorders in the United States. Results from the National Comorbidity Survey. R.C. Kessler, PhD; K.A.McGonagle, PhD; S. Zhao, PhD; C.B. Nelson, MPH; M. Hughes, PhD; S. Eshleman, MA; H.U. Wittchen, PhD; K.S.Kendler, MD. Arch Gen Psychiatry, Vol.51, P.P. 8-19, January 1994.

Appendix B: Definition of 25 MSAs Selected for Small Area Estimation

Appendix B. Definition of 25 MSAs Selected for Small Area Estimation

	1992 Population Projection
1. Anaheim-Santa Ana, CA	**1,996,343**
California, Orange County	1,996,343
2. Atlanta, GA	**2,424,882**
Georgia, Barrow County	25,179
Georgia, Butts County	12,970
Georgia, Cherokee County	77,733
Georgia, Clayton County	151,319
Georgia, Cobb County	391,285
Georgia, Coweta County	45,782
Georgia, De Kalb County	463,515
Georgia, Douglas County	59,987
Georgia, Fayette County	55,556
Georgia, Forsyth County	38,659
Georgia, Fulton County	546,808
Georgia, Gwinnett County	310,368
Georgia, Henry County	51,116
Georgia, Newton County	35,030
Georgia, Paulding County	35,385
Georgia, Rockdale County	46,428
Georgia, Spalding County	45,367
Georgia, Walton County	32,395
3. Baltimore, MD	**1,996,133**
Maryland, Anne Arundel County	357,569
Maryland, Baltimore County	588,006
Maryland, Carroll County	104,516
Maryland, Harford County	153,125
Maryland, Howard County	163,767
Maryland, Queen Annes County	29,465
Maryland, Baltimore City	599,685
4. Boston, MA	**3,144,871**
Massachusetts, Essex County	551,793
Massachusetts, Middlesex County	1,170,799
Massachusetts, Norfolk County	516,160
Massachusetts, Plymouth County	355,370
Massachusetts, Suffolk County	550,750

Appendix B. Definition Of 25 MSAs Selected For Small Area Estimation (continued)

	1992 Population Projection
5. Chicago, IL	**4,981,375**
Illinois, Cook County	4,175,359
Illinois, Du Page County	654,320
Illinois, Mchenry County	151,696
6. Dallas, TX	**2,120,056**
Texas, Collin County	227,616
Texas, Dallas County	1,518,374
Texas, Denton County	237,971
Texas, Ellis County	70,322
Texas, Kaufman County	43,729
Texas, Rockwall County	22,044
7. Denver, CO	**1,345,668**
Colorado, Adams County	211,897
Colorado, Arapahoe County	329,437
Colorado, Denver County	381,917
Colorado, Douglas County	53,022
Colorado, Jefferson County	369,394
8. Detroit, MI	**3,592,682**
Michigan, Lapeer County	60,050
Michigan, Livingston County	96,340
Michigan, Macomb County	603,448
Michigan, Monroe County	107,869
Michigan, Oakland County	909,637
Michigan, St Clair County	119,365
Michigan, Wayne County	1,695,974
9. El Paso, TX	**456,359**
Texas, El Paso County	456,359
10. Houston, TX	**2,661,181**
Texas, Fort Bend County	184,640
Texas, Harris County	2,261,297
Texas, Liberty County	42,903
Texas, Montgomery County	152,654
Texas, Waller County	19,687
11. Los Angeles, CA	**7,127,293**
California, Los Angeles County	7,127,293
12. Miami-Hialeah, FL	**1,600,103**
Florida, Dade County	1,600,103

Appendix B. Definition Of 25 MSAs Selected For Small Area Estimation (continued)

	1992 Population Projection
13. Minneapolis-St. Paul	**2,035,347**
Minnesota, Anoka County	198,093
Minnesota, Carver County	39,020
Minnesota, Chisago County	24,780
Minnesota, Dakota County	228,159
Minnesota, Hennepin County	865,489
Minnesota, Isanti County	20,829
Minnesota, Ramsey County	398,328
Minnesota, Scott County	47,001
Minnesota, Washington County	118,214
Minnesota, Wright County	54,470
Wisconsin, St Croix County	40,963
14. Nassau-Suffolk, NY	**2,177,956**
New York, Nassau County	1,085,945
New York, Suffolk County	1,092,011
15. New York, NY	**7,085,772**
New York, Bronx County	938,595
New York, Kings County	1,873,534
New York, New York County	1,283,134
New York, Putnam County	69,880
New York, Queens County	1,648,200
New York, Richmond County	316,399
New York, Rockland County	218,687
New York, Westchester County	737,342
16. Newark, NJ	**1,499,843**
New Jersey, Essex County	628,943
New Jersey, Morris County	353,820
New Jersey, Sussex County	106,854
New Jersey, Union County	410,225
17. Oakland, CA	**1,727,459**
California, Alameda County	1,054,766
California, Contra Costa County	672,693

Appendix B. Definition Of 25 MSAs Selected For Small Area Estimation (continued)

	1992 Population Projection
18. Philadelphia, PA-NJ	**4,036,961**
New Jersey, Burlington County	329,783
New Jersey, Camden County	407,164
New Jersey, Gloucester County	190,673
Pennsylvania, Bucks County	456,745
Pennsylvania, Chester County	320,513
Pennsylvania, Delaware County	454,906
Pennsylvania, Montgomery County	573,220
Pennsylvania, Philadelphia County	1,303,956
19. Phoenix, AZ	**1,770,449**
Arizona, Maricopa County	1,770,449
20. San Antonio, TX	**1,040,318**
Texas, Bexar County	941,916
Texas, Comal County	44,601
Texas, Guadalupe County	53,801
21. San Bernardino, CA	**2,122,022**
California, Riverside County	978,288
California, San Bernardino County	1,143,734
22. San Diego, CA	**2,089,322**
California, San Diego County	2,089,322
23. St. Louis, MO-IL	**2,006,385**
Illinois, Clinton County	28,060
Illinois, Jersey County	16,844
Illinois, Madison County	206,040
Illinois, Monroe County	18,700
Illinois, St Clair County	212,612
Missouri, Franklin County	66,062
Missouri, Jefferson County	139,312
Missouri, St Charles County	177,308
Missouri, St Louis County	826,259
Missouri, St Louis City	315,187
24. Tampa-St. Petersburg, FL	**1,822,154**
Florida, Hernando County	97,648
Florida, Hillsborough County	708,882
Florida, Pasco County	258,574
Florida, Pinellas County	757,050

Appendix B. Definition Of 25 MSAs Selected For Small Area Estimation (continued)

	1992 Population Projection
25. Washington, DC	**3,345,080**
District of Columbia	513,842
Maryland, Calvert County	43,912
Maryland, Charles County	84,579
Maryland, Frederick County	127,118
Maryland, Montgomery County	651,746
Maryland, Prince Georges County	611,812
Virginia, Arlington County	151,317
Virginia, Fairfax County	706,226
Virginia, Loudoun County	73,193
Virginia, Prince William County	178,043
Virginia, Stafford County	51,577
Virginia, Alexandria City	98,061
Virginia, Fairfax City	16,699
Virginia, Falls Church City	8,157
Virginia, Manassas City	23,622
Virginia, Manassas Park City	5,176

Appendix C: Fitting the Models Given Variable Selection

Appendix C. Fitting the Models Given Variable Selection

As discussed in Section 2.4, the model fitting procedure was an iterative series of iterative subprocedures. The first subprocedure is based on knowing G, the variance-covariance matrix for the random effects. The objective of this subprocedure is to find β 55^ and U 50^ that maximize $L_w(\beta,U|y,G)$.

C.1 Estimating β and *U* Given *G*

Because the natural log function is monotone increasing, finding β 55^ and U 50^ that maximize $L_w(\beta,U|y,G)$ is identical to finding β 55^ and U 50^ that maximize $\ln[L_w(\beta,U|y,G)]$. Taking the log will simplify further steps considerably. Standard results from calculus tell us that if a function of one variable has a maximum at an interior point of its domain, then the first derivative will be zero at that point and the second derivative will be negative. Similar results are available for functions of many variables. Recall that if f is a real-valued function of $x_1,\ldots,x_n$ then the transpose of the Jacobian matrix[1] of f is an n x 1 matrix defined as

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_1,\ldots,x_n) \\ \mathit{DOTSVERT} \\ \frac{\partial f}{\partial x_n}(x_1,\ldots,x_n) \end{bmatrix},$$

and the Hessian of f is an n x n matrix defined as

$$\frac{\partial^2 f(x)}{\partial x^2} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_1,\ldots,x_n) & \mathit{DOTSAXIS} & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_1,\ldots,x_n) \\ \mathit{DOTSVERT} & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x_1,\ldots,x_n) & \mathit{DOTSAXIS} & \frac{\partial^2 f}{\partial x_n^2}(x_1,\ldots,x_n) \end{bmatrix},$$

For f to have a local maximum at an interior point x^*, it is necessary that

$$\frac{\partial f(x^*)}{\partial x} = \begin{bmatrix} 0 \\ \mathit{DOTSVERT} \\ 0 \end{bmatrix}$$

[1] Sometimes the determinate of this matrix is called "the Jacobian," sometimes the matrix itself is labeled "the Jacobian."

and sufficient that this condition hold and that

$$\frac{\partial^2 f(x^*)}{\partial x^2}$$

is a negative definite matrix, by which it is meant that

$$x' \frac{\partial^2 f(x^*)}{\partial x^2} x < 0 \quad all\ x \neq 0.$$

The transpose of the Jacobian matrix of the log of the likelihood function (with respect to the parameters — not the data) is often called the efficient score of the parameters for the data or simply the score function. In our application, we write

$$S_w(\beta, U | y, G) = \frac{\partial \ln[L_w(\beta, U | y, G)]}{\partial(\beta, U)} .$$

We now derive this score function

Note that

$$\ln L_w = \sum_{i=1}^{nleft} [w_i y_i \ln(\pi_i) + w_i(1-y_i)\ln(1-\pi_i)$$

Recall from our model that logit (π_i) = $X_i\beta + Z_iU$, where X_i is the $^{i\text{-th}}$ row of X and Z_i is the $^{i\text{-th}}$ row of Z. By inverting the logit transform, we have that

$$\pi_i = \frac{e^{X_i\beta + Z_iU}}{1+e^{X_i\beta + Z_iU}}$$

From this, it easily follows that

$$1 - \pi_i = \frac{1}{1+e^{X_i\beta + Z_iU}} .$$

Thus, we may write

$$\ln L_w = \sum_{i=1}^{n} \left\{ w_i y_i \left[(X_i\beta + Z_iU) - \ln\left(1+e^{X_i\beta+Z_iU}\right) \right] - w_i(1-y_i)\ln\left(1+e^{X_i\beta+Z_iU}\right) \right\}$$

$$-\frac{1}{2}U^tG^{-1}U - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln|G|$$

$$= \sum_{i=1}^{n} \left\{ w_i y_i (X_i\beta + Z_iU) - w_i \ln\left(1+e^{X_i\beta+Z_iU}\right) \right\} - \frac{1}{2}U^tG^{-1}U - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln|G| .$$

Taking first the derivative with respect to β_1, we have that

$$\frac{\partial}{\partial \beta_1}\ln(L_w) = \sum_{i=1}^{n}\left[w_i y_i X_{i1} - \frac{w_i X_{i1} e^{X_i\beta + Z_i U}}{1+e^{X_i\beta + Z_i U}}\right]$$

$$= \sum_{i=1}^{n} X_{i1} w_i (y_i - \pi_i) \ ,$$

where X_{i1} is the first component of the row vector X_i.

Similarly,

$$\frac{\partial \ln}{\partial \beta_p}(L_w) = \sum_{i=1}^{n} X_{ip} w_i (y_i - \pi_i).$$

Let W be an n x n diagonal matrix with the sampling weights on the diagonal:

$$W = \begin{bmatrix} w_1 & & O \\ & \ddots & \\ O & & w_n \end{bmatrix},$$

where it is important to note that the n weights were standardized to sum to n. Otherwise, the procedure behaves as if we have a sample as large as the U.S. population. Then

$$\frac{\partial \ln}{\partial \beta}(L_w) = \begin{bmatrix} \sum_{i=1}^{n} X_{i1} w_i (y_i - \pi_i) \\ \vdots \\ \sum_{i=1}^{n} X_{ip} w_i (y_i - \pi_i) \end{bmatrix} = X^t W(y - \pi).$$

Note that $W(y-\pi)$ is an n x 1 matrix and that X^t is p x n so that $X^t W(y-\pi)$ is a p x 1 matrix, as required. Having found the Jacobian of $\ln L_w$ with respect to β, we now need to find the Jacobian with respect to U. Here it is useful to note that

$$G^{-1} = \begin{bmatrix} \frac{I_m}{\sigma_1^2} & O \\ O & \frac{I_r}{\sigma_2^2} \end{bmatrix}$$

So that $U^t G^{-1} U$ is simply

$$\frac{1}{\sigma_1^2}\sum_{i=1}^{m} U_i^2 + \frac{1}{\sigma_2^2}\sum_{i=m+1}^{m+r} U_i^2 \ ,$$

where m is the number of states and r is the number of sample PSUs. Taking first the derivative with respect to U_1, we have that

$$\frac{\partial}{\partial U_1}\ln(L_w) = \sum_{i=1}^{n}\left[w_i y_i Z_{i1} - \frac{w_i Z_{i1} e^{X_i\beta + Z_i U}}{1+e^{X_i\beta+Z_i U}}\right] - \frac{\partial}{\partial U_1}\frac{1}{2}\left[\frac{1}{\sigma_1^2}\sum_{i=1}^{m} U_i^2 + \frac{1}{\sigma_2^2}\sum_{i=m+1}^{m+r} U_i^2\right]$$

$$= \sum_{i=1}^{n} Z_{i1} w_i (y_i - \pi_i) - \frac{U_1}{\sigma_1^2}$$

Similarly, for the random effect corresponding to the m[th] state,

$$\frac{\partial}{\partial U_m}\ln(L_w) = \sum_{i=1}^{n} Z_{im} w_i (y_i - \pi_i) - \frac{U_m}{\sigma_1^2}$$

For the r random effects corresponding to the PSU-level random effects, we have

$$\frac{\partial}{\partial U_j}\ln(L_w) = \sum_{i=1}^{n} Z_{ij} w_i (y_i - \pi_i) - \frac{U_j}{\sigma_2^2} \quad m < j \le m+r.$$

We can thus write

$$\frac{\partial}{\partial U}\ln(L_w) = \begin{bmatrix} \sum_{i=1}^{n} Z_{i1} w_i (y_i - \pi_i) \\ \vdots \\ \sum_{i=1}^{n} Z_{im} w_i (y_i - \pi_i) \\ \sum_{i=1}^{n} Z_{i,m+1} w_i (y_i - \pi_i) \\ \vdots \\ \sum_{i=1}^{n} Z_{i,m+r} w_i (y_i - \pi_i) \end{bmatrix} - \begin{bmatrix} \frac{1}{\sigma_1^2} U_1 \\ \vdots \\ \frac{1}{\sigma_1^2} U_m \\ \frac{1}{\sigma_2^2} U_{m+1} \\ \vdots \\ \frac{1}{\sigma_2^2} U_{m+r} \end{bmatrix} = Z^t W (y - \pi) - G^{-1} U$$

By stacking $\frac{\partial}{\partial \beta}\ln(L_w)$ above $\frac{\partial}{\partial U}\ln(L_w)$, we may write

$$S_w(\beta, U \mid y, G) = \frac{\partial}{\partial(\beta, U)}\ln(L_w) = \begin{bmatrix} X^t W (y-\pi) \\ Z^t W (y-\pi) - G^{-1} U \end{bmatrix}$$

Note that this an (p + q) x 1 matrix as required. In order to maximize L_w for fixed y and G, we need to find $\hat{\beta}$ and $\hat{U}$ such that all p + g rows of S_w are identically equal to zero. Since this system of equations is nonlinear in β and U, we need to use numerical methods to solve it. The method we used is called Fisher's method of scoring (or simply "the method of scoring").

To describe this method, we need to define another term. The "Fisher information matrix" (or simply "the information matrix") is the expected value of the product of the score function with its transpose with respect to the distribution of y. We write this as

$$J_w(\beta, U \mid y, G) = E(S_w S_w^t),$$

where the expected value is with respect to the modeled distribution for y and U, ignoring the sample design and treating G and β as fixed.

In our application, note that J_w is a (p + q) x (p + q) matrix. Deriving it for our application is made simpler if we write S_w in terms of a transform of y-π.

$$\varepsilon_i = \frac{y_i - \pi_i}{\pi_i(1-\pi_i)}$$

Note that the expected value of ε_i given fixed π_i is $E(\varepsilon_i|\pi_i) = 0$ since $E(y_i \mid \pi_i) = \pi_i$. Also note that

$$Var(\varepsilon_i|\pi_i) = \frac{\pi_i(1-\pi_i)}{[\pi_i(1-\pi_i)]^2}$$

$$= \frac{1}{\pi_i(1-\pi_i)}$$

and that $Cov(\varepsilon_i, \varepsilon_j|\pi_i, \pi_j) = 0$ since the individual outcomes are assumed to be independent once the random effects have been fixed. If we use C to denote the n x n diagonal matrix with $\pi_i(1-\pi_i)$ sequenced down the main diagonal as

$$C = \begin{bmatrix} \pi_{11}(1-\pi_1) & O \\ \text{DOTSDIAG} & \\ O & \pi_n(1-\pi_n) \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \text{DOTSVERT} \\ \varepsilon_n \end{bmatrix}$$

then we may write

$$\varepsilon = C^{-1}(y-\pi)$$

the covariance matrix for ε as $\mathrm{Cov}(\varepsilon) = C^{-1}$, and the score function as

$$S_w(\beta, U|\varepsilon, G) = \begin{bmatrix} X^t W C \varepsilon \\ Z^t W C \varepsilon - G^{-1} U \end{bmatrix}.$$

The Fisher information matrix for β and U given ε and G is then

$$\mathcal{J}_w(\beta,U|\varepsilon,G) = E\left\{ \begin{bmatrix} X^tWC\varepsilon \\ Z^tWC\varepsilon - G^{-1}U \end{bmatrix} [\varepsilon^tCWX \quad \varepsilon^tCWZ - U^tG^{-1}] \right\},$$

where we note that C, W, and G^{-1} are all symmetric. Expanding the product, we have

Error!

We now note that

$$E(\varepsilon_i U_i) = E_{U_i}\left[E(\varepsilon_i|U_i)\right] = E_{U_i}\left[E(\varepsilon_i|\pi_i)\right] = E_{U_i}(0) = 0$$

so that $E(\varepsilon U^t) = E(U\varepsilon^t) = 0$. We also note that by definition, $E(\varepsilon\varepsilon^t) = C^{-1}$ and $E(UU^t) = G$. Furthermore, since we are treating G as fixed and known, we can move the expected value operator inside to just the factors involving U and ε. We thus have

$$\mathcal{J}_w = \begin{bmatrix} X^tWCWX & X^tWCWZ \\ Z^tWCWX & Z^tWCWZ + G^{-1} \end{bmatrix}.$$

Returning to the definition of Fisher's method of scoring, this is an iterative method that starts with a guess at a solution and then improves upon the guess. The improvement equation in our application is

$$\begin{bmatrix} \hat\beta^{(a+1)} \\ \hat U^{(a+1)} \end{bmatrix} = \begin{bmatrix} \hat\beta^{(a)} \\ \hat U^{(a)} \end{bmatrix} + [J_w^{(a)}]^{-1} S_w^{(a)},$$

where the a superscript indexes iterations of the cycle.

As the algorithm converges, $\hat\beta^{(a+1)} - \hat\beta^{(a)} \to 0$, $\hat U^{(a=1)} - \hat U^{(a)} \to 0$, $S_w^{(a)} \to 0$, and $\mathcal{J}_w^{(a)} \to J_w^{(\infty)}$. We note that the matrices X, W, Z and G^{-1} are all constant during the convergence. Only the matrix C changes in J_w. Note that if any of the π_i are close to 0 or to 1, then C will have a diagonal entry close to zero, thereby making J_w ill-conditioned (i.e., difficult to invert). In such a case, convergence may be a problem. Otherwise, it may be proven that this process will converge. It is our understanding, however, that it has not been proven that there is a single solution to $S_w = 0$. If there are multiple solutions, it may be best to start the system with several different initial guesses. See Section C.3 for more discussion of this point.

Although we have completely specified the procedure for finding $\hat\beta$ and $\hat U$ given G, it is useful for discussion purposes to describe the algorithm in an equivalent but different form. To this end, we note that the score function can be rewritten as
where

$$S_w = \begin{bmatrix} X^tW(y-\pi) \\ Z^tW(y-\pi) - G^{-1}U \end{bmatrix} = \begin{bmatrix} X^tWCW(\zeta - X\beta - ZU) \\ Z^tWCW(\zeta - X\beta - ZU) - G^{-1}U \end{bmatrix},$$

$$\zeta = \begin{bmatrix} X_1\beta+Z_1U+\dfrac{y_1-\pi_1}{w_1\pi_1(1-\pi_1)} \\ DOTSVERT \\ X_n\beta+Z_nU+\dfrac{y_n-\pi_n}{w_n\pi_n(1-\pi_n)} \end{bmatrix} = X\beta+ZU+W^{-1}C^{-1}(y-\pi) = X\beta+ZU+W^{-1}\varepsilon$$

Continuing to transform S_w, we have

$$S_w = \begin{bmatrix} X^tWCW\zeta \\ Z^tWCW\zeta \end{bmatrix} - \begin{bmatrix} X^tWCWX\beta \\ Z^tWCWX\beta \end{bmatrix} - \begin{bmatrix} X^tWCWZU \\ Z^tWCWZU+G^{-1}U \end{bmatrix}$$

$$= \begin{bmatrix} X^tWCW\zeta \\ Z^tWCW\zeta \end{bmatrix} - \begin{bmatrix} X^tWCWX \\ Z^tWCWX \end{bmatrix}\beta - \begin{bmatrix} X^tWCWZ \\ Z^tWCWZ+G^{-1} \end{bmatrix}U$$

$$= \begin{bmatrix} X^tWCW\zeta \\ X^tWCW\zeta \end{bmatrix} - \begin{bmatrix} X^tWCWX & X^tWCWZ \\ Z^tWCWX & Z^tWCWZ+G^{-1} \end{bmatrix} \begin{bmatrix} \beta \\ U \end{bmatrix}$$

$$= \begin{bmatrix} X^tWCW\zeta \\ Z^tWCW\zeta \end{bmatrix} - J_w \begin{bmatrix} \beta \\ U \end{bmatrix}$$

So if we had the final version of C available, we could solve S_w=0 by just setting

This is a nice form since it corresponds to the form that would be used if ζ were the observed

$$\begin{bmatrix} \beta \\ U \end{bmatrix} = J_w^{-1} \begin{bmatrix} X^tWCW\zeta \\ Z^tWCW\zeta \end{bmatrix} .$$

random variable instead of y and if ζ were a normal random variable. For this reason, we have sometimes referred to ζ as the working linear variable since it allows us to express the solution for β and U as a linear function of the "data." Of course, iteration is still required for the solution in this form since one must have preliminary values of β and U in order to calculate ζ , C, and J_w.

We give the form of J_W^{-1} below and prove that the form is correct, but some additional notation and a lemma are first required. Let $R=(WCW)^{-1}$ and $V=R+ZGZ^t$. Also, let $P=V^{-1}-V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}$. This matrix has several useful properties that we prove as needed throughout the appendix. The properties that we need to find J_w^{-1} are that P is symmetric and that P is orthogonal to X, meaning that PX = 0.

Lemma C.1: PX=0

Proof:

$$
\begin{aligned}
PX &= [V^{-1}-V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}]X \\
&= V^{-1}[X-X(X^tV^{-1}X)^{-1}X^tV^{-1}X] \\
&= V^{-1}[X-XI] \\
&= V^{-1}0 \\
&= 0 .
\end{aligned}
$$

We are now ready to prove that

$$
J_w^{-1} = \begin{bmatrix} (X^tV^{-1}X)^{-1} & -(X^tV^{-1}X)^{-1}(X^tV^{-1}Z)G \\ -G(Z^tV^{-1}X)(X^tV^{-1}X)^{-1} & G-GZ^tPZG \end{bmatrix}
$$

We will prove this by demonstrating that $J_wJ_w^{-1} = I.$ A similar argument would demonstrate that $J_w^{-1}J_w = I,$ as required. There are four steps to the proof. Using the new notation, we have that

$$
J_w = \begin{bmatrix} X^tR^{-1}X & X^tR^{-1}Z \\ Z^tR^{-1}X & Z^tR^{-1}Z+G^{-1} \end{bmatrix}
$$

We must show that

$$
\begin{aligned}
X^tR^{-1}X(X^tV^{-1}X)^{-1} - X^tR^{-1}ZG(Z^tV^{-1}X)(X^tV^{-1}X)^{-1} &= I_{pxp}, \\
-X^tR^{-1}X(X^tV^{-1}X)^{-1}(X^tV^{-1}Z)G + X^tR^{-1}Z(G-GZ^tPZG) &= 0_{pxq}, \\
Z^tR^{-1}X(X^tV^{-1}X)^{-1} - (Z^tR^{-1}Z+G^{-1})G(Z^tV^{-1}X)(X^tV^{-1}X)^{-1} &= 0_{qxp}, \\
-Z^tR^{-1}X(X^tV^{-1}X)^{-1}(X^tV^{-1}Z)G + (Z^tR^{-1}Z+G^{-1})(G-GZ^tPZG) &= I_{qxq}.
\end{aligned}
$$

Proving the first condition, we have that

$$
\begin{aligned}
&X^tR^{-1}X(X^tV^{-1}X)^{-1} - X^tR^{-1}ZG(Z^tV^{-1}X)(X^tV^{-1}X)^{-1} \\
&= X^tR^{-1}[I-ZGZ^tV^{-1}]X(X^tV^{-1}X)^{-1} \\
&= X^tR^{-1}[V-ZGZ^t]V^{-1}X(X^tV^{-1}X)^{-1} \\
&= X^tR^{-1}RV^{-1}X(X^tV^{-1}X)^{-1} \\
&= X^tV^{-1}X(X^tV^{-1}X)^{-1} \\
&= I_{pxp},
\end{aligned}
$$

as required. Proving the second condition, we have that

$$-X^tR^{-1}X(X^tV^{-1}X)^{-1}(X^tV^{-1}Z)G+(X^tR^{-1}Z)(G-GZ^tPZG)$$
$$= X^tR^{-1}[-X(X^tV^{-1}X)^{-1}X^tV^{-1}Z+Z(I-GZ^tPZ)]G$$
$$= X^tR^{-1}[-X(X^tV^{-1}X)^{-1}X^tV^{-1}+(I-ZGZ^tP)]ZG$$
$$= X^tR^{-1}V[-V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}+V^{-1}(I-ZGZ^tP)]ZG$$
$$= X^tR^{-1}V[V^{-1}-V^{-1}X(X^tV^{-1}x)^{-1}X^tV^{-1}-V^{-1}ZGZ^tP]ZG$$
$$= X^tR^{-1}V[P-V^{-1}ZGZ^tP]ZG \qquad \textit{by definition of } P$$
$$= X^tR^{-1}V[I-V^{-1}ZGZ^t]PZG$$
$$= X^tR^{-1}[V-ZGZ^t]PZG$$
$$= X^tR^{-1}RPZG$$
$$= (X^tP)ZG$$
$$= 0ZG \qquad \textit{since } PX=0 \quad X^tP=(PX)^t=0^t=0$$
$$= 0_{pxq}$$

as required.

Proving the third condition, we have that

$$
\begin{aligned}
&Z^tR^{-1}X(X^tV^{-1}X)^{-1}-(Z^tR^{-1}Z+G^{-1})G(Z^tV^{-1}X)(X^tV^{-1}X)^{-1}\\
&= [Z^tR^{-1}X-(Z^tR^{-1}Z+G^{-1})GZ^tV^{-1}X](X^tV^{-1}X)^{-1}\\
&= [Z^tR^{-1}-(Z^tR^{-1}Z+G^{-1})GZ^tV^{-1}]X(X^tV^{-1}X)^{-1}\\
&= [Z^tR^{-1}V-(Z^tR^{-1}Z+G^{-1})GZ^t]V^{-1}X(X^tV^{-1}X)^{-1}\\
&= [Z^tR^{-1}V-(Z^tR^{-1}ZGZ^t+Z^t)]V^{-1}X(X^tV^{-1}X)^{-1}\\
&= [Z^tR^{-1}V-Z^tR^{-1}(ZGZ^t+R)]V^{-1}X(X^tV^{-1}X)^{-1}\\
&= [Z^tR^{-1}V-Z^tR^{-1}V]V^{-1}X(X^tV^{-1}X)^{-1}\\
&= 0V^{-1}X(X^tV^{-1}X)^{-1}\\
&= 0_{qxp},
\end{aligned}
$$

as required.

Proving the fourth condition, we have that

$$
\begin{aligned}
&-Z^tR^{-1}X(X^tV^{-1}X)^{-1}(X^tV^{-1}Z)G+(Z^tR^{-1}Z+G^{-1})(G-GZ^tPZG)\\
&= [-Z^tR^{-1}VV^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}Z+(Z^tR^{-1}Z+G^{-1})(I-GZ^tPZ)]G\\
&= [-Z^tR^{-1}V(V^{-1}-P)Z+(Z^tR^{-1}Z+G^{-1})(I-GZ^tPZ)]G \qquad \text{by definition of } P\\
&= [-Z^tR^{-1}Z+Z^tR^{-1}VPZ+Z^tR^{-1}Z-Z^tR^{-1}ZGZ^tPZ+G^{-1}-Z^tPZ]G\\
&= [Z^tR^{-1}VPZ-Z^tR^{-1}(V-R)PZ+G^{-1}-Z^tPZ]G \qquad \text{by definition of } V\\
&= [Z^tR^{-1}VPZ-Z^tR^{-1}VPZ+Z^tPZ+G^{-1}-Z^tPZ]G\\
&= G^{-1}G\\
&= I_{qxq},
\end{aligned}
$$

as required. This completes the proof that J_w^{-1} has the form claimed. Going back to the solution for (β 55^, U 50^) in terms of J_w^{-1} and ζ, we have that

$$\begin{bmatrix}\hat{\beta}\\ \hat{U}\end{bmatrix} = J_w^{-1}\begin{bmatrix}X^tWCW\zeta\\ Z^tWCW\zeta\end{bmatrix} = J_w^{-1}\begin{bmatrix}X^tR^{-1}\zeta\\ Z^tR^{-1}\zeta\end{bmatrix}$$

$$= \begin{bmatrix}(X^tV^{-1}X)^{-1} & -(X^tV^{-1}X)^{-1}(X^tV^{-1}Z)G\\ -G(Z^tV^{-1}X)(X^tV^{-1}X)^{-1} & G-GZ^tPZG\end{bmatrix}\begin{bmatrix}X^tR^{-1}\zeta\\ Z^tR^{-1}\zeta\end{bmatrix}$$

$$= \begin{bmatrix}(X^tV^{-1}X)^{-1}X^t & -(X^tV^{-1}X)^{-1}(X^tV^{-1}Z)GZ^t\\ -G(Z^tV^{-1}X)(X^tV^{-1}X)^{-1}X^t & (G-GZ^tPZG)Z^t\end{bmatrix}\begin{bmatrix}R^{-1}\zeta\\ R^{-1}\zeta\end{bmatrix}$$

$$= \begin{bmatrix}(X^tV^{-1}X)^{-1}(X^t-X^tV^{-1}ZGZ^t)\\ GZ^t[-V^{-1}X(X^tV^{-1}X)^{-1}X^t+I-PZGZ^t]\end{bmatrix}R^{-1}\zeta$$

$$= \begin{bmatrix}(X^tV^{-1}X)^{-1}X^tV^{-1}(V-ZGZ^t)\\ GZ^t[-V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}+V^{-1}-PZGZ^tV^{-1}]V\end{bmatrix}R^{-1}\zeta$$

$$= \begin{bmatrix}(X^tV^{-1}X)^{-1}X^tV^{-1}R\\ GZ^t[P-PZGZ^tV^{-1}]V\end{bmatrix}R^{-1}\zeta$$

$$= \begin{bmatrix}(X^tV^{-1}X)^{-1}X^tV^{-1}\\ GZ^t[P-PZGZ^tV^{-1}]VR^{-1}\end{bmatrix}\zeta$$

$$= \begin{bmatrix}(X^tV^{-1}X)^{-1}X^tV^{-1}\\ GZ^tP(V-ZGZ^t)R^{-1}\end{bmatrix}\zeta$$

$$= \begin{bmatrix}(X^tV^{-1}X)^{-1}X^tV^{-1}\\ GZ^tP\end{bmatrix}\zeta$$

In other words, given preliminary estimates of π, we may compute ζ, C, R, V, and P and then obtained improved estimates of β and U as

$$\hat{\beta} = (X^tV^{-1}X)^{-1}X^tV^{-1}\zeta$$

and

$$\hat{U} = GZ^tP\zeta.$$

Given improved estimates of β and U, we can get an improved estimate of π by

$$\pi\ 50\wedge = \frac{e^{X\beta\ 50\wedge + ZU\ 50\wedge}}{1+e^{X\beta\ 50\wedge + ZU\ 50\wedge}}.$$

We iterate back and forth, first improving $\beta\ 55\wedge$and $U\ 50\wedge$ and then improving $\pi\ 50\wedge$ until the process converges.

Recall, however, that this is all conditional on knowing *G*, which we don't. In the next section, we describe the estimation process for *G*.

C.2 Estimating G Given $\beta\, U$

Staying within the Bayes approach, the ideal procedure would be to derive the posterior distribution of G given y, β, and U, and then to find the mean of that posterior distribution or perhaps its mode. Unfortunately, this approach is not tractable. This would require integration of L_W not just with respect to U, but also with respect to β. As discussed in Section C.1, the integration of L_W with respect to just U was deemed too numerically intensive. Integrating with respect to the *p* components of β would raise the amount of required computer time to even higher levels.

In order to get some sort of estimate of *G*, we followed the approach used by Stiratelli et al (1984), Schall (1991), and Breslow and Clayton (1993). This approach works by adapting a method that has favorable properties when the outcome variable *y* has a multivariate normal distribution.

Specifically, consider the linear model

$$\zeta = X\beta + ZU + e,$$

where *X, β, Z,* and *U* are defined the same as in our model, Cov*(U,e) = 0,* and e ~ $N_n(0,R)$. For this linear model, there are several alternate methods for estimating *G*. The maximum likelihood method finds the maximum of the likelihood of *G* given ζ. This means finding the value of *G* that maximizes

$$lnL_1^*(G\,|\,\beta,\zeta) = -\frac{1}{2} ln\,|\,R+ZGZ^t\,| - \frac{1}{2}(\zeta - X\beta)^t\ (R+ZGZ^t)^{-1}\ (\zeta - X\beta).$$

This method is favored by some statisticians, but others are displeased by the fact that G obtained by this method is different from the classical estimates for balanced ANOVAs and that the maximum likelihood estimate of G is seriously negatively biased when the number of fixed effects is large relative to the number of observations.

Because of these concerns, an alternative method of estimating G in linear models was developed. This method is called restricted maximum likelihood (REML). For this method, G is estimated by finding the value of G that maximizes

$$\ln L_1 \ (G \mid \beta, \zeta) = \frac{1}{2} \ln | R + ZGZ^t | - \frac{1}{2} \ln | X^T (R + ZGZ^t)^{-1} X | + \\ - \frac{1}{2} (\zeta - X\beta)^t (R + ZGZ^t)^{-1} (\zeta - X\beta),$$

where it is assumed that $X^t X$ is of rank p so that the determinate in the second term is nonzero. The REML estimate G does agree with the classical G for balanced designs and it is not as biased when p is large relative to n. Further arguing in favor of the REML estimate of G, Harville (1977) noted that in Harville (1974) he had demonstrated that with a noninformative prior on (β, G), the REML estimate of G maximizes the marginal posterior distribution of G, while the ML estimate of G is merely the value of G that together with some value of β maximizes the joint posterior distribution of β and G. For more information on REML methods, see Patterson and Thompson (1971) and Cox and Reid (1987).

We chose the REML estimator for this project because we believed that it might be more robust to the nonnormality of ζ in our application and because Breslow and Clayton conducted some simulations of it which appeared encouraging.

Further justifying the use of this approach, we note that although ζ as defined in Section C.1 is far from multivariate normal, it does have the correct mean and covariance matrix. Furthermore, if we equate $W^{-1}\varepsilon$ from our model with e in the linear model, we have that $\mathrm{Cov}(U, W^{-1}\varepsilon) = 0$ as required and that $\mathrm{Cov}(W^{-1}\varepsilon) = R = (WCW)^{-1}$.

Without further reference to nonnormality of $W^{-1}\varepsilon$, we describe the algorithm we used to obtain the REML estimate of G. The method was iterative, as for the estimation of (β, V). In fact, we again used Fisher's method of scoring as defined in Section C.1.

The score function for G using the REML approach is

$$S_1(G \mid \beta, \zeta) = \frac{\partial}{\partial G} \ln L_1(G \mid \beta, \zeta).$$

Here we note that G has a very simple structure as specified in Section C.1. So we will define the score function as the derivative of $\ln L_1$, with respect to (σ_1, σ_2). By finding the values of σ_1 and σ_2 (positive or negative) that maximize L_1, we will be able to estimate σ_1^2 by $(\sigma_1)^2$ and σ_2^2 by $(\sigma_2)^2$, thereby avoiding difficulties with negative estimates of variance components.

Starting with the first partial derivative with respect to σ_1, we have

$$S(\sigma_1) = \frac{\partial}{\partial \sigma_1} \ln L_1$$

$$= \frac{\partial}{\partial \sigma_1} \left\{ -\frac{1}{2} \ln |R + ZGZ^t| - \frac{1}{2} \ln |X^t (R + ZGZ^t)^{-1} X| - \frac{1}{2} (\zeta - X\beta)^t (R + ZGZ^t)(\zeta - X\beta) \right\}$$

$$= \frac{\partial}{\partial \sigma_1} \left\{ -\frac{1}{2} \ln |V| - \frac{1}{2} \ln |X^t V^{-1} X| - \frac{1}{2} (\zeta - X\beta)^t V^{-1} (\zeta - X\beta) \right\}$$

To derive this partial derivative, we need some results from stochastic linear algebra.

Lemma C.2: Let A be a symmetric nonsingular n x n matrix whose elements are functions of variables $x_1, \ldots, x_t$. Then

$$\frac{\partial}{\partial x_i} \ln |A| = \text{trace}\left(A^{-1} \frac{\partial A}{\partial x_i} \right),$$

where

$$\frac{\partial A}{\partial x_i} = \begin{pmatrix} \frac{\partial a_{11}}{\partial x_i} \ DOTSLOW \ \frac{\partial a_{1n}}{\partial x_i} \\ DOTSVERT \\ \frac{a_{n1}}{\partial x_i} \ DOTSLOW \ \frac{\partial a_{nn}}{\partial x_i} \end{pmatrix} ,$$

recalling that the trace of a matrix is the sum of the elements on its main diagonal. For the proof of Lemma C.2, see Graybill (1969), section 10.8.

Lemma C.3:

Let A and B be conformable matrices so that AB is defined. Then

$$\frac{\partial AB}{\partial \theta} = \frac{\partial A}{\partial \theta} B + A \frac{\partial B}{\partial \theta} .$$

Again, see Graybill for a proof.

Lemma C.4:

Let A be a square nonsingular matrix. Then $\frac{\partial A^{-1}}{\partial \theta} = - A^{-1} \frac{\partial A}{\partial \theta} A^{-1}$.

Proof:

Since $AA^{-1} = I$ and $\frac{\partial I}{\partial \theta} = 0$, we have

$$0 = \frac{\partial A}{\partial \theta} A^{-1} + A \frac{\partial A^{-1}}{\partial \theta}$$, by Lemma C.4

leading to

$$-\frac{\partial A}{\partial \theta} A^{-1} = A \frac{\partial A^{-1}}{\partial \theta}$$ and

$$- A^{-1} \frac{\partial A}{\partial \theta} A^{-1} = \frac{\partial A^{-1}}{\partial \theta},$$

as required.

Lemma C.5:

Let A and B be conformable matrices so that AB and BA are defined. Then $tr\,(AB) = tr\,(BA)$. (See any linear algebra textbook for proof.)

This means that even though matrix multiplication is not commutative, it is possible to commute the order of multiplication inside the trace operator.

Lemma C.6:

Let x be a matrix such that $x^t A x$ is defined and x does not depend on θ. Then

$$\frac{\partial (x^t A x)}{\partial \theta} = x^t \frac{\partial A}{\partial \theta} x .$$

(Again, see Graybill for proof).

Lemma C.7:

$$V^{-1} (\zeta - X\beta) = P\zeta$$

Proof:

$$\begin{aligned} V^{-1} (\zeta - X\beta) &= V^{-1} (\zeta - X(X^t V^{-1} X)^{-1} X^t V^{-1})\zeta) \\ &= V^{-1} (I - X (X^t V^{-1} X)^{-1} X^t V^{-1}) \zeta \\ &= P\zeta \end{aligned}$$

Lemma C.8: If y is a random vector such that E$y = 0$ with covariance matrix V and A is a fixed square conformable matrix, then

$$E\,(y^t A y) = \text{trace}\,(AV)$$

Since this lemma is simple to prove, the proof is not given here.

Lemma C.9:

$$\frac{\partial P}{\partial \sigma_i} = -P\frac{\partial V}{\partial \sigma_i}P$$

Proof:

$$\frac{\partial P}{\partial \sigma_i} = \frac{\partial}{\partial \sigma_i}(V^{-1} - V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1})$$

$$= \frac{\partial V^{-1}}{\partial \sigma_i}(I - X(X^tV^{-1}X)^{-1}X^tV^{-1}) + V^{-1}\frac{\partial}{\partial \sigma_i}(I - X(X^tV^{-1}X)^{-1}X^tV^{-1})$$

$$= -V^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)V^{-1}(I - X(X^tV^{-1}X)^{-1}X^tV^{-1}) - V^{-1}\frac{\partial}{\partial \sigma_i}(X(X^tV^{-1}X)^{-1}X^tV^{-1})$$

$$= -V^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)P - V^{-1}\left\{X\left[\frac{\partial}{\partial \sigma_i}(X^tV^{-1}X)^{-1}\right]X^tV^{-1} + X(X^tV^{-1}X)^{-1}X^t\left(\frac{\partial V^{-1}}{\partial \sigma_i}\right)\right]$$

$$= -V^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)P +$$

$$-V^{-1}\left\{-X(X^tV^{-1}X)^{-1}\frac{\partial(X^tV^{-1}X)}{\partial \sigma_i}(X^tV^{-1}X)^{-1}X^tV^{-1} - X(X^tV^{-1}X)^{-1}X^tV^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)V^{-1}\right\}$$

$$= -V^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)P +$$

$$-V^{-1}\left\{ + X(X^tV^{-1}X)^{-1}X^tV^{-1}\left(\frac{\partial V^{-1}}{\partial \sigma_i}\right)V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1} - X(X^tV^{-1}X)^{-1}X^tV^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)V^{-1}\right\}$$

$$= -V^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)P - V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)[V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1} - V^{-1}]$$

$$= -V^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)P + (P - V^{-1})\left(\frac{\partial V}{\partial \sigma_i}\right)(-P)$$

$$= -V^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)P - P\left(\frac{\partial V}{\partial \sigma_i}\right)P + V^{-1}\left(\frac{\partial V}{\partial \sigma_i}\right)P$$

$$= -P\left(\frac{\partial V}{\partial \sigma_i}\right)P$$

Returning now to the derivation of $S(\sigma_1)$, we have by Lemma C.2 that

$$\frac{\partial}{\partial \sigma_1}\ln|V| = \text{trace}\left[V^{-1}\frac{\partial V}{\partial \sigma_1}\right].$$

For the second term of $S(\sigma_1)$, we have that

$$\frac{\partial \ln}{\partial \sigma_1}|X^tV^{-1}X|$$

$$= \text{trace } [\ (X^t\ V^{-1} X)^{-1}\ \frac{\partial (X^t V^{-1} X)}{\partial \sigma_1}\ ] \qquad \text{by Lemma C.2}$$

$$= \text{trace } [\ (X^t\ V^{-1} X)^{-1}\ X^t\ \frac{\partial V^{-1}}{\partial \sigma_1}\ X\ ] \qquad \text{by Lemma C.6}$$

$$= -\text{trace } [\ (X^t\ V^{-1} X)^{-1}\ X^t V^{-1}\ \frac{\partial V}{\partial \sigma_1}\ V^{-1} X\ ] \qquad \text{by Lemma C.4}$$

$$= -\text{trace } [\ V^{-1} X (X^t V^{-1} X)^{-1}\ X^t V^{-1}\ \frac{\partial V}{\partial \sigma_1}\ ] \qquad \text{by Lemma C.5}$$

Adding this result to the result for $\frac{\partial}{\partial \sigma_1} \ln|V|$, we have that

$$\frac{\partial}{\partial \sigma_1} [\ \ln|V| + \ln|X^t V^{-1} X|\ ] = \text{trace} \left[V^{-1} \frac{\partial V}{\partial \sigma_1} \right] - \text{trace} \left[V^{-1} X (X^t V^{-1} X)^{-1} X^t V^{-1} \frac{\partial V}{\partial \sigma_1} \right]$$

$$= \text{trace} \left[V^{-1} \frac{\partial V}{\partial \sigma_1} - V^{-1} X (X^t V^{-1} X)^{-1} X^t V^{-1} \frac{\partial V}{\partial \sigma_1} \right]$$

$$= \text{trace } \left\{ [\ V^{-1}\ -\ V^{-1} X (X^t V^{-1} X)^{-1} X^t V^{-1}\ ] \frac{\partial V}{\partial \sigma_1} \right\}$$

$$= \text{trace} \left(P \frac{\partial V}{\partial \sigma_1} \right)$$

This is the general result given by Harville. Applying this general result to our particular application results in further simplification:

$$= \text{trace}\left[P \frac{\partial}{\partial \sigma_1} (R + ZGZ^t) \right] \quad \text{by definition of V}$$

$$= \text{trace}\left[P \; Z \frac{\partial G}{\partial \sigma_1} Z^t \right]$$

Error!

$$= \text{trace} \left\{ P Z \begin{bmatrix} 2\sigma_1 0 \\ 0 \;\; 0 \end{bmatrix} Z^t \right\}$$

$$= \text{trace} \left\{ P \begin{bmatrix} Z \; 50\sim_1 \\ Z \; 50\sim_2 \end{bmatrix} \begin{bmatrix} 2\sigma_1 I_m 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} Z_1 & 80\sim^t \\ Z_2 & 80\sim^t \end{bmatrix} \right\}$$

$$= 2\sigma_1 \text{trace} \left\{ P [Z \; 50\sim_1 \; Z \; 50\sim_2] \begin{bmatrix} Z \; 50\sim^t_1 \\ 0 \end{bmatrix} \right\}$$

$$= 2\sigma_1 \text{trace} \left(PZ \; 55\sim_1 Z \; 55\sim^t_1 \right)$$

$$= 2\sigma_1 \text{trace}\left(Z\ 50\sim_1^{t} PZ\ 50\sim_1\right)$$

where $Z\ 55\sim_1$ consists of the first *m* columns of *Z* and $Z\ 55\sim_2$ consists of the remaining *r* columns of *Z*.

Working on the third term of $S(\sigma_1)$, we have that

$$\frac{\partial}{\partial\sigma_1}[(\zeta - X\beta)^t V^{-1}(\zeta - X\,betahat)]$$

$$= \frac{\partial}{\partial\sigma_1}[\zeta - X\beta]^t P\zeta) \text{ by lemma C.7}$$

$$= \frac{\partial}{\partial\sigma_1}[\zeta^t P\zeta] \text{ since } X^t P = 0 \text{ by lemma C.1}$$

$$= \zeta^t \left(\frac{\partial P}{\partial\sigma_1}\right)\zeta$$

$$= -\zeta^t P\left(\frac{\partial V}{\partial\sigma_1}\right)P\zeta \text{ by lemma C.9.}$$

$$= -(\zeta - X\beta)^t V^{-1}\left(\frac{\partial V}{\partial\sigma_1}\right)V^{-1}(\zeta - X\beta) \text{ by lemma C.7}$$

This is the general result given by Harville. Applying this general result to our particular application results in further simplification.

As demonstrated in the derivation of the second term of $S(\sigma_1)$, we know that in this application,

$$\frac{\partial V}{\partial \sigma_1} = 2\sigma_1 Z\ 55\sim_1 Z\ 55\sim_1^t \ .$$

So

$$\frac{\partial}{\partial \sigma_1}[(\zeta - X\beta)^t V^{-1}(\zeta - X\beta)]$$

$$= -2\sigma_1(\zeta - X\beta)^t V^{-1} Z\ 55\sim_1 Z\ 55\sim_1^t V^{-1}(\zeta - X\beta)$$

$$= -2\sigma_1 \left[Z\ 55\sim_1^t V^{-1}(\zeta - X\beta)\right]^t \left[Z\ 55\sim_1^t V^{-1}(\zeta - X\beta)\right]$$

$$= -2\sigma_1 \xi_1^t \xi_1 \quad ,$$

where $\xi_1 = Z\ 55\sim_1^t V^{-1}(\zeta - X\beta)$

Finally, we can write out a complete expression for the score function for σ_1 using Harville's general form as

$$S(\sigma_1) = \frac{\partial \ln L_1}{\partial \sigma_1} = \frac{\partial}{\partial \sigma_1}\left\{-\frac{1}{2}\ln|V| - \frac{1}{2}\ln\left|X^t V^{-1} X\right| - \frac{1}{2}(\zeta - X\beta)^t V^{-1}(\zeta - X\beta)\right\}$$

$$= -\frac{1}{2}\,\mathrm{trace}\left(P\frac{\partial V}{\partial \sigma_1}\right) + \frac{1}{2}(\zeta - X\beta)^t V^{-1}\frac{\partial V}{\partial \sigma_1}V^{-1}(\zeta - X\beta)$$

Equivalently for our application, we may write

$$S(\sigma_1) = -\sigma_1\,\mathrm{trace}\,(Z\ 55\sim_1^t P Z\ 55\sim_1) + \sigma_1\ \xi\ 50\wedge_1^t \xi\ 50\wedge_1$$

Going through similar algebra, we get the score function for σ_2 as

$$S(\sigma_2) = \frac{\partial \ln L_1}{\partial \sigma_2} = -\sigma_2 \operatorname{trace}(\tilde{Z}_2^t P \tilde{Z}_2) + \sigma_2 \hat{\xi}_2^t \hat{\xi}_2$$

Just as in the estimation of (β, U), we want to find (σ_1, σ_2) such that simultaneously solves the system $S(\sigma_1) = 0$ and $S(\sigma_2) = 0$. Clearly, $(\hat{\sigma}_1, \hat{\sigma}_2) = (0,0)$ is a root of the score function that does not interest us. It is interesting to note the lemmas C.5, C.7 and C.8 can be used to demonstrate that $E(\hat{\xi}_i^t \hat{\xi}_i) = (\tilde{Z}_i^t P \tilde{Z}_i)$ for σ_1 and σ_2 that solve the system. This means that finding the REML estimates of σ_1 and σ_2 is equivalent to setting $\hat{\xi}_i^t \hat{\xi}_i$ equal to its own expected value.

Solving the system requires an iterative approach. We did this using Fisher's method of scoring again. However, rather than computing the Fisher information matrix as

$$E\left\{ \begin{bmatrix} S(\sigma_1) \\ S(\sigma_2) \end{bmatrix} \begin{bmatrix} S(\sigma_1) \\ S(\sigma_2) \end{bmatrix} \right\},$$ we used a theorem from advanced probability theory.

This theorem states that given suitable regularity conditions,[2] $E(SS^t) = -E\left[\frac{\partial S}{\partial \theta}\right]$, where

$$\frac{\partial S}{\partial \theta} = \begin{pmatrix} \frac{\partial^2 \ln f}{\partial \theta_1^2} & \cdots & \frac{\partial^2 \ln f}{\partial \theta_1 \partial \theta_n} \\ \vdots & & \vdots \\ \frac{\partial^2 \ln f}{\partial \theta_1 \partial \theta_n} & \cdots & \frac{\partial^2 \ln f}{\partial \theta_n^2} \end{pmatrix}$$

[2] See Cramer (1946) for a set of conditions and a proof.

is the Hessian of ln*f* with respect to θ.

For this particular application, it is easier to find the second partial derivatives of $\ln L_1$ and then to find the expected value of those second partial derivatives than to try to find the expected value of the product of the score function with its transpose. The reason for this is the difficulty of integrating products of traces. To find the second partial derivatives of $\ln L_1$, we go back to the equations for $S(\sigma_i)$ in terms of V to allow easier comparison with Harville (1977).

Focusing first on the first two components of $\ln L_1$, we have that

$$\frac{\partial^2}{\partial\sigma_i\partial\sigma_1}\left[\ln|V| + \ln|X^tV^{-1}X|\right] = \frac{\partial}{\partial\sigma_i}\operatorname{trace}\left(P\frac{\partial V}{\partial\sigma_j}\right)$$

$$= \operatorname{trace}\left[\frac{\partial}{\partial\sigma_i}\left(P\left(\frac{\partial V}{\partial\sigma_j}\right)\right)\right] \text{ since the trace is a linear function of the elements of a matrix}$$

$$= \operatorname{trace}\left[\left(\frac{\partial P}{\partial\sigma_i}\right)\left(\frac{\partial V}{\partial\sigma_j}\right) + P\left(\frac{\partial^2 V}{\partial\sigma_i\partial\sigma_j}\right)\right]$$

$$= \operatorname{trace}\left[-P\left(\frac{\partial V}{\partial\sigma_i}\right)P\left(\frac{\partial V}{\partial\sigma_j}\right) + P\left(\frac{\partial^2 V}{\partial\sigma_i\partial\sigma_j}\right)\right]$$

$$= \operatorname{trace}\left[P\left(\frac{\partial^2 V}{\partial\sigma_i\partial\sigma_j}\right) - P\left(\frac{\partial V}{\partial\sigma_i}\right)P\left(\frac{\partial V}{\partial\sigma_j}\right)\right]$$

$$= \operatorname{trace}\left[PZ\left(\frac{\partial^2 G}{\partial\sigma_i\partial\sigma_j}\right)Z^t - PZ\left(\frac{\partial G}{\partial\sigma_i}\right)Z^tPZ\left(\frac{\partial G}{\partial\sigma_j}\right)Z^t\right]$$

For $i = j$, this expression will equal

$$\text{trace}\left[\, 2\, P Z\ 55\sim_i\ Z\ 55\sim_i^t - 4\, sigmasubi^2\, P Z\ 55\sim_i Z\ 55\sim_i^t P Z\ 55\sim_i Z\ 55\sim_i^t \,\right]$$

For $i \neq j$, this expression will equal

$$-\text{trace}\left[\, 4\, \sigma_i \sigma_j P Z\ 55\sim_i Z\ 55\sim_i^t P Z\ 55\sim_j Z\ 55\sim_j^t \,\right]$$

since $\frac{\partial^2 G}{\partial\sigma_i \partial\sigma_j} = 0$ for $i \neq j$.

Focusing now on the third component of $\ln L_1$, we have that

$$\frac{\partial^2}{\partial\sigma_i \partial\sigma_j}(\zeta - X\beta)^t V^{-1}(\zeta - X\beta)$$

$$= -\frac{\partial}{\partial\sigma_i}\left\{(\zeta - X\beta)^t V^{-1}\left(\frac{\partial V}{\partial\sigma_j}\right)V^{-1}(\zeta - X\beta)\right\}$$

$$= -\frac{\partial}{\partial\sigma_i}\left[\zeta^t P\left(\frac{\partial V}{\partial\, sigmasubi}\right)P\zeta\right]$$

$$= -\zeta^t\left[\left(\frac{\partial P}{\partial\sigma_i}\right)\left(\frac{\partial V}{\partial\sigma_j}\right)P + P\left[\left(\frac{\partial^2 V}{\partial\sigma_i \partial\sigma_j}\right)P + \left(\frac{\partial V}{\partial\sigma_j}\right)\left(\frac{\partial P}{\partial\sigma_i}\right)\right]\right]\zeta$$

$$= -\zeta^t\left[-P\left(\frac{\partial V}{\partial\sigma_i}\right)P\left(\frac{\partial V}{\partial\sigma_j}\right)P + P\left(\frac{\partial^2 V}{\partial\sigma_i\partial\sigma_j}\right)P - P\left(\frac{\partial V}{\partial\sigma_j}\right)P\left(\frac{\partial V}{\partial\sigma_i}\right)P\right]\zeta$$

$$= -\zeta^t P\left[-\left(\frac{\partial V}{\partial\sigma_i}\right)P\left(\frac{\partial V}{\partial\sigma_j}\right) + \left(\frac{\partial^2 V}{\partial\sigma_i\partial\sigma_j}\right) - \left(\frac{\partial V}{\partial\sigma_j}\right)P\left(\frac{\partial V}{\partial\sigma_i}\right)\right]P\zeta$$

$$= -(\zeta^t - X\beta)V^{-1}\left[\frac{\partial^2 V}{\partial\sigma_i\partial\sigma_j} - 2\left(\frac{\partial V}{\partial\sigma_j}\right)P\left(\frac{\partial V}{\partial\sigma_i}\right)\right]V^{-1}(\zeta - X\beta)$$

(The last equality holds since $\frac{\partial^2 A}{\partial\sigma_j\partial\sigma_i}=\frac{\partial^2 A}{\partial\sigma_j\partial\sigma_i}$ for any twice differential matrix A.)

We have now reconfirmed the formula $\frac{\partial^2 lnL_1}{\partial\sigma_j\partial\sigma_i}$ given by Harville (1977):

Error!

Having found the Hessian of $\ln L_1$, we now need to find the expected value of this Hessian with respect to the distribution of ζ in order to get the Fisher information matrix for G. We first prove several lemmas.

Lemma C.10: Cov(ζ)=V

Proof:

Lemma C.11: *V* is a generalized inverse of *P*, meaning that *PVP=P*

$$\begin{aligned} Cov\,\zeta &= Cov(X\beta + ZU + W^{-1}C^{-1}(y-\pi)) \\ &= Cov(ZU) + Cov(W^{-1}\varepsilon) \qquad U \;\; \varepsilon \\ &= Z(Cov\,U)Z^t + W^{-1}1(Cov\,\varepsilon)W^{-1} \quad (W \textit{ is diagonal hence symmetric}) \\ &= ZGZ^t + W^{-1}C^{-1}W^{-1} \\ &= ZGZ^t + (WCW)^{-1} \\ &= ZGZ^t + R \\ &= V. \end{aligned}$$

Proof:

$$PVP = [V^{-1}-V^{-1}X(X^tV^{-1}X)^{-1}V^{-1}]\,V\,[V^{-1}-V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}]$$
$$= V^{-1}[I-X(X^tV^{-1}X)^{-1}X^tV^{-1}][I-X(X^tV^{-1}X)^{-1}X^tV^{-1}]$$
$$= V^{-1}[I-2X(X^tV^{-1}X)^{-1}X^tV^{-1}+X(X^tV^{-1}X)^{-1}X^tV^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}]$$
$$= V^{-1}[I-2X(X^tV^{-1}X)^{-1}X^tV^{-1}+X(X^tV^{-1}X)^{-1}X^tV^{-1}]$$
$$= V^{-1}[I-X(X^tV^{-1}X)^{-1}X^tV^{-1}]$$
$$= V^{-1}-V^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}$$
$$= P.$$

Lemma C.12: Cov(Ay) = A Cov(y)A^t

The proof for Lemma C.12 is not given here but is elementary.

Continuing with the task of finding the expected value of the Hessian of lnL_1, we note that the

first term of $\frac{\partial^2 \ln L_1}{\partial\sigma_j\partial\sigma_i}$ is a constant and that the second term may be written as $\frac{1}{2}b^t A b$,

where b = Pζ= $V^{-1}(\zeta - X\beta)$ is a

random variable with expected value 0, and

$$A = \frac{\partial^2 V}{\partial\sigma_j\partial\sigma_i} - 2\left(\frac{\partial V}{\partial\sigma_i}\right)P\left(\frac{\partial V}{\partial\sigma_j}\right).$$

Using lemma C.8, we have that

$$\mathrm{E}(b^t Ab) = \mathrm{trace}(A\mathrm{Cov}(b)) = \mathrm{trace}((\mathrm{Cov}(b))A)$$

The covariance matrix for b is

Cov b = Cov $(P\zeta)$
= P Cov (ζ) P by lemma C.12 since $P^t = P$
= PVP by lemma C.10
= P by lemma C.11.

Thus

$$
\begin{aligned}
&\mathrm{E}_{\zeta}\left(\frac{\partial^2 L_1}{\partial\sigma_i\,\partial\sigma_j}\right)\\
&= -\frac{1}{2}\,\mathrm{trace}\left\{P\left[\left(\frac{\partial^2 V}{\partial\sigma_i\,\partial\sigma_j}\right) - \left(\frac{\partial V}{\partial\sigma_i}\right)P\left(\frac{\partial V}{\partial\sigma_j}\right)\right]\right\}\\
&\quad + \frac{1}{2}\,\mathrm{trace}\left\{P\left[\left(\frac{\partial^2 V}{\partial\sigma_i\,\partial\sigma_j}\right) - 2\left(\frac{\partial V}{\partial\sigma_i}\right)P\left(\frac{\partial V}{\partial\sigma_j}\right)\right]\right\}
\end{aligned}
$$

$$
= -\frac{1}{2}\,\mathrm{trace}\left[P\left(\frac{\partial V}{\partial\sigma_i}\right)P\left(\frac{\partial V}{\partial\sigma_i}\right)\right],
$$

as claimed by Harville.

Applying this general formula to our application, we have that

$$
E_{\zeta}\left(\frac{\partial^2 L_1}{\partial\sigma_i\,\partial\sigma_j}\right) = -\frac{1}{2}\,\mathrm{trace}\left[P Z\frac{\partial G}{\partial\sigma_i}Z^t P Z\frac{\partial G}{\partial\sigma_j}Z^t\right]
$$

$$
= -2\sigma_i\sigma_j\,\mathrm{trace}\left[P Z\ 55\sim_i Z\ 55\sim_i^t P Z\ 55\sim_j Z\ 55\sim_j^t\right]
$$

Thus, Fisher’s information matrix is

Error!

Alternatively using lemma C.5, we can specify the (i,j)-th element of J_G as

$$J_G(i,j) = 2\sigma_i\sigma_j \left[(Z\ 55\sim_i^t PZ\ 55\sim_j)(Z\ 55\sim_i^t PZ\ 55\sim_j)^t \right] \quad .$$

This alternative formula is computationally useful since for an arbitrary matrix A, the trace of AA^t can be computed as the sum of the squared elements of A.

We now have all the information needed to write out our iterative cycle for estimating (σ_1, σ_2) using Fisher's method of scoring and the REML form of the log likelihood function.

Given initial estimates $\sigma_1^{(0)}\ \sigma_2^{(0)}$, we calculated an initial value for G and hence V and P. We already had values for β to calculate ξ 50^. We then used these to obtain an initial value for the score function

Error!

and an initial value for Fisher's information matrix

Error!

Improved estimates of σ_1 and σ_2 were then obtained as

$$\begin{pmatrix} \sigma_1^{(1)} \\ \\ \sigma_2^{(1)} \end{pmatrix} = \begin{pmatrix} \sigma_1^{(0)} \\ \\ \sigma_2^{(0)} \end{pmatrix} + \left(J_G^{(0)} \right)^{-1} S_G^{(0)}$$

This process is then allowed to iterate until convergence is obtained.

C.3 Joint Estimation of β, U, G

In Section C.1, an iterative method was described for estimating β and U, given G . In section C.2, another iterative method was described for estimating G given $\beta\, U$. These algorithms were run simultaneously with interchange of estimates back and forth to jointly estimate β, U, G.

We experimented with different ways of running the algorithms. The method that we chose as providing the fastest convergence was to iterate the first algorithm a single step, then to iterate the second algorithm a single step, then return to the first for one more step, then return to the second for one more step, and so on.

The main alternative we considered was to allow the first algorithm to converge, then to allow the second to converge, return to the first for another full convergence, and so on. This alternative proved to be slower. Other alternatives are possible such as allowing the second algorithm two steps for every one step of the first, but interchange after a single step of each algorithm appeared to be a good choice.

Even so, the computational burden was such that the full model (with state and PSU random effects) could not be run on a Pentium in a reasonable time. We ran the reduced models (with only state effects) on several Pentiums and the full models on a separate VAX server. For future applications, we would recommend the use of a dedicated VAX server or one or more fast UNIX workstations. For some states and variables, we could not get the full model to converge. When this occurred, we dropped back to the reduced model.

With an iterative procedure of this complexity, there is always the possibility of convergence to local maxima rather than global maxima. There is also the possibility of convergence to saddlepoints (where even though the score function is zero, the derivative of the score function is indeterminate, meaning that there are directions for change in the log likelihood that are both positive and negative). If the method were strict maximum likelihood, then there are theorems to indicate when the score function has a unique solution and that thus there were no saddlepoints and the only local maximum was also a global maximum. However, our method only approximates maximum likelihood and the method used to force $\sigma_1^2 > 0$ $\sigma_2^2 > 0$ does create saddlepoints since $S(\sigma_1, \sigma_2) = 0$ for $(\sigma_1^2, \sigma_1^2) = (0,0)$, which is clearly not an optimum estimate. Thus, the ideal procedure would have been to choose multiple starting points for the algorithm and to compare the results of starting it from those various points. However, the CPU time and VAX charges were higher per run than our time schedule and budget could support for multiple runs. We, therefore, ran each model only once, choosing the starting values carefully.

The starting values for β were taken from the ordinary logistic regressions for y or X without random effects. The starting values for U were taken to be zero, a reasonable procedure since the unconditional expectation of U is zero. The starting values for G were taken from by guessing at how much shrinkage should occur. See Section C.5 for a description of how shrinkage relates to the components of variance.

C.4 Asymptotic Properties of the Estimators

A sequence of random variables $\{Y_i\}$ is said to be consistent for a parameter θ if

$$\lim_{i\to\infty} Pr(|Y_i - \theta| > \varepsilon) = 0 \text{ for every } \varepsilon > 0.$$

Similarly, two sequences $\{Y_i\}$ and $\{X_i\}$ are said to be asymptotically equivalent if $\lim_{i\to\infty} Pr(|Y_i - X_i| > \varepsilon) = 0$ for every $\varepsilon > 0$. We claimed in Section 1.4.6 and again in Section 2.1.2
that

$$\lim_{n_i \to N_i} \frac{\pi\ 50\wedge^{w}_{i}}{\pi\ 50\wedge^{D}_{i}} = 1,$$

where i indexes the target small areas. By this rather loose statement, we meant that $\pi\ 50\wedge^{w}_{i}$ and $\pi\ 50\wedge^{D}_{i}$ are asymptotically equivalent as defined above. This claim has actually not yet been rigorously proven, but we do discuss our reasoning in this section for believing this assertion to be true. A rigorous proof might also require that N_i tend to infinity. Letting both n_i and N_i tend to infinity requires some care in specifying the stratification and clustering of the population as it grows larger that we do not attempt to address here.

We base our belief that $\pi\ 50\wedge^{w}_{i}$ and $\pi\ 50\wedge^{D}_{i}$ are asymptotically equivalent upon the observation that if the scaled residual error ε was normally distributed such that its covariance matrix did not depend on β, then the desired asymptotic equivalency can be proven. Recall from Section C.1 that we defined $\varepsilon = C^{-1}(y - \pi)$ and that we defined $\zeta = X\beta + ZU + W^{-1}\varepsilon$ with Cov(U) = G. If ε were multivariate with covariance matrix $C^{-1} = \sigma^2_e I$, then our score function for β and U would have the same form as in Section C.1, but since the conditional covariance matrix for ζ given U would not depend on β or U, it would be possible to get an explicit formula for the root of the score function instead of having to use an iterative method to find the root. Note that these conditions hold for the standard linear mixed model where $y = X\beta + ZU + \varepsilon$, Cov(U) = G, $\varepsilon \sim N_n(O, I\sigma^2_e)$, U and ε are independent and ζ is still defined as $\zeta = X\beta + ZU + W^{-1}\varepsilon$. In this case the conditional covariance matrix for ζ given U is $R = W^{-1} C^{-1} W^{-1}$.

Recall that the score function for β and U given C, G and ζ is

$$S(\beta, U|C,G,\zeta) = \begin{pmatrix} X^t WCW(\zeta - X\beta - ZU) \\ Z^t WCW(\zeta - X\beta - ZU) - G^{-1}U \end{pmatrix}.$$

Given a large enough sample size, we can assume that the method used to estimate *C* and *G* is accurate enough to treat $C = C$ and $G = G$ as fixed and known.

To keep discussion simple assume that there is only one level of random effects so that $G = \sigma_u^2 I_q$. (I_q is the *q*x*q* identity matrix, where m=q is the number of states in the model.)

Since $S(\beta$ 55^$, U$ 50^$|C,G,\zeta) = 0$, we know from looking at the bottom half of the score function that $Z^t WCW(\zeta - X\beta$ 55^$- ZU$ 50^$) = G^{-1}U$ 50^ so

$$Z^t \frac{WW}{\sigma_e^2}(\zeta - X\beta \text{ 55^} - ZU \text{ 55^}) = \frac{U \text{ 55^}}{\sigma_u^2} .$$

By the definitions of *y*, ζ and ε, we know that $\varepsilon = W(\zeta - X\beta$ 55^$- ZU$ 50^$)$ and that $\varepsilon = y - X\beta$ 55^$- ZU$ 50^. By equating these two expressions for ε, we have that

$$W(\zeta - X\beta \text{ 55^} - ZU \text{ 50^}) = y - X\beta \text{ 55^} - ZU \text{ 50^}.$$

Substituting this result into the earlier equation, we have that

$$\frac{Z^t W(y - X\beta \text{ 55^} - ZU \text{ 50^})}{\sigma_e^2} = \frac{U \text{ 50^}}{\sigma_u^2}.$$

Collecting all the terms involving U 50^, we have that

$$\left[\frac{I}{\sigma_u^2} + \frac{Z^t WZ}{\sigma_e^2}\right] U \text{ 55^} = \frac{Z^t W(y - X\beta \text{ 55^})}{\sigma_e^2}$$

So

$$U \text{ 55^} = \left[\frac{\sigma_e^2}{\sigma_e^2} I + Z^t WZ\text{right]}^{-1} Z^t W(y - X\beta \text{ 55^})$$

Note further that

$$Z^t WZ = \begin{bmatrix} n_1 w_1 & & 0 \\ & n_2 w_2 & \\ & \ddots & \\ 0 & & n_m w_m \end{bmatrix},$$

where n_i is the number of sample persons in the i-th state and w_i is the average sampling weight in the i-th state.

Thus,

$$\left[\frac{\sigma_e^2}{\sigma_u^2} I + Z^t WZ\right]^{-1} = \begin{bmatrix} \dfrac{1}{\dfrac{\sigma_e^2}{\sigma_u^2} + n_1 w_1} & & 0 \\ & \ddots & \\ 0 & & \dfrac{1}{\dfrac{\sigma_e^2}{\sigma_u^2} + n_m w_m} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\sigma_u^2/(n_1 w_1)}{\dfrac{\sigma_e^2}{n_1 w_1} + \sigma_u^2} & & 0 \\ & \ddots & \\ 0 & & \dfrac{\sigma_u^2/(n_m w_m)}{\dfrac{\sigma_e^2}{n_m w_m} + \sigma_u^2} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\gamma_1}{n_1 w_1} & & 0 \\ & \ddots & \\ 0 & & \dfrac{\gamma_m}{n_m w_m} \end{bmatrix},$$

where

$$\gamma_i = \frac{\sigma_u^2}{\sigma_u^2 + \frac{\sigma_e^2}{n_i w_i}}$$

is the shrinkage factor for the i-th state.

Since the weights were scaled, $w_i \geq 1$. So γ_i is the proportion of the total variance on the logit scale that is due to true differences among the states rather than measurement variance. For N_i sufficiently large,

$$\lim_{n_i \rightarrow N_i} \gamma_i – 1.$$

Focusing on $U\ 50^\wedge{}_i$ for an individual state, we note that the i-th row of $Z^t W(y-X\beta)$ has the form

$$\sum_{j=1}^{n_i} w_{ij}(y_{ij}\text{-}X_{ij}\beta\ 55^\wedge) = n_i w_i(y_{Di}\text{-}X_{Di}\beta\ 50^\wedge\})\quad ,$$

where $y_{Di} = \sum_{j=1}^{n_i} \frac{w_{ij}}{n_i w_i} y_{ij}$ is the design-consistent estimator of the mean of y in the i-th state, and

X_{Di} is defined similarly. Combining these results leads to the explicit solution for $U\ 50^\wedge{}_i$ of

$$U\ 50^\wedge\}_i = \gamma_i(y_{Di}\text{-}X_{Di}\beta\ 50^\wedge\}).$$

Given this solution for $U\ 50^\wedge{}_i$, the empirical Bayes estimator π_i^w for small area-i becomes

$$\pi\ 50^\wedge\}_i^w = \frac{1}{N_i}\sum_{j=1}^{N_i}(X_{ij}\beta\ 50^\wedge\} + U\ 50^\wedge\}_i)$$

$$= X_{\Omega i}\beta \; 50\hat{}_i + U \; 50\hat{}_i$$

where $X_{\Omega i}$ depicts the population mean of the X_{ij} vectors. Substituting in the explicit solution for U_i leads to

$$\pi_i^w = X_{\Omega i}\beta + \gamma_i(y_{Di} - X_{Di}\beta)$$

$$= (1-\gamma_i)X_{\Omega i}\beta + \gamma_i[y_{Di} - (X_{Di} - X_{\Omega i}\beta)]$$

Note that y_{Di} is the traditional Horvitz-Thompsen estimator and $y_{Di} - (X_{DI} - X_{\Omega i})\beta$ is design consistent for the mean of *y* in state *i* since X_{Di} is a design consistent estimator of $X_{\Omega i}$. Therefore with $\pi_i^D = y_{Di}$ for this linear mixed model case

$$\lim_{n_i \to N_i} (\pi_i^w \div \pi_i^D) = 1$$

since X_{Di} is a design consistent estimator of $X_{\Omega i}$. This establishes the desired result.

As mentioned earlier, the corresponding result for applying the method to binary variables has not yet been proven. The discreteness of the distribution for binary variables is not the primary difficulty in attempting a proof since with large enough *n*, the binary variables could be grouped into sets of binomial variables which would have approximately normal distributions by the Central Limit Theorem. The main difficulty in attempting a proof is that the matrix R is a function of β. The variance of a normal variable is usually not related to its mean. When there is a relationship, the method of proof given here fails.

Model Consistency

In addition to the question of design consistency, there is the issue of model consistency. Under fairly general conditions, maximum likelihood estimates are consistent for the parameters they are estimating. Less is known about the asymptotic properties of REML estimates, even for normally distributed errors. The model-based asymptotic characteristics of the approximations used here to estimate *G* are unknown, even when β is known. Given that in the procedure used here, there is a pattern of alternating iterations between an approximate maximum likelihood estimate for (β,*U*) and an approximate restricted maximum likelihood estimate for *G*, the asymptotic frequentist properties of the method will be difficult to ever determine.

In a sense, however, the asymptotic design-based and model-based frequentist properties of the estimator are of only limited interest since the sample sizes available for a small area are small. Mean square error of the point estimates of π is of greater interest, along with the coverage properties of the confidence intervals. Simulation studies will probably be the only fruitful approach to studying these issues. Breslow and Clayton (1993) had some small simulation studies, but more are needed.

C.5 Computational Tricks and the Relationship of the Survey-Weighted Empirical Bayes Estimate to Composite Estimation

In this section, we show versions of the equations from Section C.1 that are faster to calculate than the forms given in that section. First, a useful lemma is reproduced from Robinson (1991).

Lemma C.13: Let *A, U, B,* and *V* be matrices such that $A+UBV$ is well defined and invertible and such that A^{-1} exists. Then

$$(A+UBV)^{-1} = A^{-1} - A^{-1}U(I+BVA^{-1}U)^{-1}BVA^{-1}.$$

Proof: We demonstrate below that the right-hand form of the equation is the inverse of $A+UBV$ by multiplying it by $A+UBV$ showing that the product is the identity matrix. Since matrix multiplication is not commutative, a full-proof would require evaluating the results of left-hand and right-hand multiplication, but we leave the other side of the proof to the interested reader.

$$\begin{aligned}
&(A+UBV)[A^{-1}-A^{-1}U(I+BVA^{-1}U)^{-1}BVA^{-1}] \\
&= (A+UBV)A^{-1}[I-U(I+BVA^{-1}U)^{-1}BVA^{-1}] \\
&= (I+UBVA^{-1})[I-U(I+BVA^{-1}U)^{-1}BVA^{-1}] \\
&= I-U(I+BVA^{-1}U)^{-1}BVA^{-1}+UBVA^{-1}-UBVA^{-1}U(I+BVA^{-1}U)^{-1}BVA^{-1} \\
&= I-U[(I+BVA^{-1}U)^{-1}-I+BVA^{-1}U(I+BVA^{-1}U)^{-1}]BVA^{-1} \\
&= I-U[-I+(I+BVA^{-1}U)(I+BVA^{-1}U)^{-1}]BVA^{-1} \\
&= I-U[-I+I]BVA^{-1} \\
&= I.
\end{aligned}$$

Applying lemma C.13 to V, we have that since $V=R+ZGZ^t$, its inverse may be written as

$$V^{-1} = R^{-1} - R^{-1} Z(I + GZ^t R^{-1} Z)^{-1} GZ^t R^{-1} .$$

Since R is a diagonal matrix, it is quite easy to find its inverse. Also $I{+}GZ^tR^{-1}Z$ is of dimension qxq, so it is much easier to find its inverse than to find the inverse of V directly since V is nxn. Because the matrix $I{+}GZ^tR^{-1}Z$ needs to be inverted very often in the iterative procedure, we developed a closed-form expression for the inverse rather than relying on numerical methods to invert the matrix repeatedly.

Once this matrix has been inverted, simplifications in the formulas for the random effects can be achieved that also help the intuitive understanding of the nature of these effects. First, we use lemma C.13 to derive an expression for estimating the random effects that is fast to calculate. From Section C.1, we have that

Error!

We will make this general matrix equation more specific to our application and even faster to compute by actually finding a closed form expression for the inverse of $I{+}GZ^tR^{-1}Z$. To find such an expression is messy, but the computing rewards are strong. We now need to consider that for some characteristics and age ranges, we were unable to get the full model with state and PSU random effects to converge. Where we were unable to obtain convergence, we simply dropped the PSU random effects from the model. The matrix inversion is considerably simpler for this simpler model and so we go through the steps for it first. If there is only a state effect, then the columns of Z are mutually orthogonal so $GZ^tR^{-1}Z$ is diagonal with I-th diagonal element

$$\sigma_1^2 \sum_{j=1}^{r_i} \sum_{k=1}^{n_{ij}} w_{ijk}^2 \pi_{ijk}(1-\pi_{ijk})$$

We will simplify notation considerably by defining

$$\alpha_{ijk} = w_{ijk}^2 \pi_{ijk}(1-\pi_{ijk})$$

and

$$\alpha_{i++} = \sum_{j=1}^{r_i} \sum_{k=1}^{n_{ij}} \alpha_{ijk}$$

and then using these α factors as weights to get weighted averages of the ζ and X variables:

$$\zeta_i = \sum_{j=1}^{r_i}\sum_{k=1}^{n_{ij}} \frac{\alpha_{ijk}}{\alpha_{i++}}\zeta_{ijk} \qquad X_i = \sum_{j=1}^{r_i}\sum_{k=1}^{n_{ij}} \frac{\alpha_{ijk}}{\alpha_{i++}}X_{ijk}$$

Then $U_i \;=\; \dfrac{1}{1+\sigma_1^2\sum_j w_{ij}^2\pi_{ij}(1-\pi_{ij})}\;\sigma_1 sup2\sum_j^{n_i} w_{ij}^2\pi_{ij}(1-\pi_{ij})(\zeta_{ij}-X_{ij}\beta\;\;55\wedge)$

$$= \frac{\sigma_1^2\alpha_{i++}}{1+\sigma_1^2\alpha_{i++}}\sum_j^{r_i}\sum_k^{n_{ij}}\left(\frac{\alpha_{ijk}}{\alpha_{i++}}\right)(\zeta_{ij}-X_{ij}\beta\;\;55\wedge)$$

$$= \frac{\sigma_1^2}{\dfrac{1}{\alpha_{i++}}+\sigma_1^2}(\zeta_i-X_i\beta\;\;55\wedge)$$

$$= \gamma_i(\zeta_i - X_i\beta\;\;55\wedge)$$

where

$$\gamma_i = \frac{\sigma_1^2}{\dfrac{1}{\alpha_{i++}}+\sigma_1^2}$$

Note that since the weights were standardized to sum to the national sample size, α_{i++} is approximately equal to the sample size for the state multiplied by $\pi(1-\pi)$ where π is the true national average propensity. Thus,

$$\gamma_i \approx \frac{\sigma_1^2}{\dfrac{1}{n_i\pi(1-\pi)}+\sigma_1^2}$$

This form of the state random effect allows us to demonstrate a linkage between this estimation system and composite estimation. Recall that $\lambda_{ijk} = \ln\left(\frac{\pi_{ijk}}{1-\pi_{ijk}}\right)$ is the estimated logit

propensity to engage in the behavior of interest. If we think of averaging the estimated logit propensities across the entire population in the i-th state, we obtain

Error!

This equation looks just like a classical composite estimator with the design-weighted sample data in the second term and the model-based prediction in the first term. Because of this similarity, we refer to the factor γ_i as the shrinkage factor for the *i*-th state. If γ_i is close to 1, then the sample data from the state predominate the estimated logit propensity for the state. On the other hand, if γ_i is close to zero, then national data predominate the estimated logit propensity for the state through the model. However, it is probably not appropriate to push this analogy too hard since ζ_i is a complex function of the sample data in the state and the true propensities in the state:

$$\zeta_i = \frac{\sum_j \sum_k \alpha_{ijk} \zeta_{ijk}}{\alpha_{i++}}$$

$$= \frac{\sum_j \sum_k \alpha_{ijk} \left[X_{ijk} + U_i + \frac{y_{ijk} - \pi_{ijk}}{w_{ijk} \pi_{ijk} (1-\pi_{ijk})} \right]}{\alpha_{i++}}$$

$$= X_i \beta + U_i + \frac{\sum_j \sum_k w_{ijk} (y_{ijk} - \pi_{ijk})}{\alpha_{i++}}$$

$$= \frac{\sum_j \sum_k \alpha_{ijk} \operatorname{logit}(\pi_{ijk})}{\alpha_{i++}} + \frac{\sum_j \sum_k w_{ijk} (y_{ijk} - \pi_{ijk})}{\alpha_{i++}}$$

Also note that the composting takes place on the logit transform of the estimated propensity.

That completes the discussion of the simplified model with only state-level random effects. We now tackle the more difficult analogous equations for the full model with state-level and PSU-level random effects. With 2 levels of random effects,

$$I + GZ^t R^{-1} Z = \begin{bmatrix} I+\sigma_i^2 Z_1^t R^{-1} Z_1 & \sigma_1^2 Z_1^t R^{-1} Z_2 \\ \sigma_2^2 Z_2^t R^{-1} Z_1 & I+\sigma_2^2 Z_2^t R^{-1} Z_2 \end{bmatrix}$$

where Z1 and Z2 were defined in Section C.2. The columns of Z_1 are mutually orthogonal as are the columns of Z_2, making this matrix easier to invert, but Z_1 and Z_2, are not orthogonal to each other. We know that all columns involving different states are orthogonal. So we can re-order the random effects in the order U_1, U_{11}, ...,U_{1r_1}, ..., U_m, U_{m1}, ... U_{mrm} and reorder the columns of Z accordingly. Let each block of columns corresponding to a state be written B_i so that $Z = [B_1 DOTSAXIS B_m]$.

Also let

$$Q_i = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 I_{r_i} \end{bmatrix},$$

where r_i is number of sample PSUs in i-th state. Then

$$I+GZ^tR^{-1}Z = \begin{bmatrix} I+Q_iB_1^tR^{-1}B_1 & & 0 \\ & DOTSDIAG & \\ 0 & & I+Q_mB_m^tR^{-1}B_m \end{bmatrix},$$

where the I-th block on the main diagonal is $(1+r_i)$x$(1+r_i)$ in dimension. To invert this matrix, we need only invert each block. Taking the first block as an example, it has the form

$$I+Q_1B_1^tR^{-1}B_1 = \begin{bmatrix} 1+\sigma_1^2\alpha_{1++} & \sigma_1^2\alpha_{11+} & DOTSLOW & \sigma_1^2\alpha_{1r_1+} \\ \sigma_2^2\alpha_{11+} & 1+\sigma_2^2\alpha_{11+} & 0 & \\ & DOTSDIAG & & \\ \sigma_2^2\alpha_{1r_1+} & 0 & 1+\sigma_2^2\alpha_{1r_1+} & \end{bmatrix}$$

where $\alpha_{ijk} = w_{ijk}^2\pi_{ijk}(1-\pi_{ijk})$ and "+" denotes summation on a subscript.

Some grinding but routine linear algebra yields a closed-form expression for the inverse of this matrix. We don't give the derivation here, but merely the result.

Let $h = 1+\sigma_1^2\alpha_{1++} - \sum_{i=1}^{r_1} \frac{\sigma_1^2\sigma_2^2\alpha_{1i+}^2}{1+\sigma_2^2\alpha_{1i+}}$.

Then $(I+Q_1B_1^tR^{-1}B_1)^{-1}$ has the general form

$$\frac{1}{h}\begin{pmatrix} 1 & -\sigma_1^2A^t \\ -\sigma_2^2A & AA^t+hH \end{pmatrix},$$

where

$$A^t = \left[\frac{\alpha_{11+}}{1+\sigma_2^2\alpha_{11+}} \ \textit{DOTSLOW} \ \frac{\alpha_{1r_i+}}{1+\sigma_2^2\alpha_{1r_i+}} \right]$$

and

$$H = \begin{bmatrix} \frac{1}{1+\sigma_2^2\alpha_{11+}} & 0 \\ \textit{DOTSDIAG} & \\ 0 & \frac{1}{1+\sigma_2^2\alpha_{1r_i+}} \end{bmatrix},$$

This means that the random effect for the first state is given by the first row of

$$(I+Q_1B_1^tR^{-1}B_1)^{-1}Q_1B_1^tR^{-1}(\zeta\text{-}X\beta\ 55\text{^})$$

The random effects for the sample PSUs in that state are given by the remaining rows. Concentrating on the first row, we have that

$$U_1 = \frac{\sigma_1^2}{h}\sum_j\sum_k\alpha_{1jk}(\zeta_{1jk}\text{-}X_{1jk}\beta\ 55\text{^})\text{-}\ \frac{\sigma_1^2\sigma_2^2}{h}\sum_j\frac{\alpha_{1j+}}{1+\sigma_2^2\alpha_{1j+}}\sum_k\alpha_{1jk}(\zeta_{1jk}\text{-}X_{1jk}\beta\ 55\text{^}).$$

If we let $\zeta_{ij} = \sum_k\frac{\alpha_{ijk}}{\alpha_{ij+}}\zeta_{ijk}\ \ \zeta\ 60{=}_i = \frac{\sum_j\frac{\alpha_{ij+}}{\alpha_{ij+}\sigma_2^2+1}\zeta_{ij}}{\sum_j\frac{\alpha_{ij+}}{\alpha_{ij+}\sigma_2^2+1}}$

and define $X_{ij}\ \ X\ 60{=}_i$ similarly, then we have that

$$U_1 = \frac{\sigma_1^2}{h}\sum_j\sum_k\left[\left(1\text{-}\frac{\sigma_2^2\alpha_{1j+}}{1+\sigma_2^2\alpha_{1j+}}\right)\alpha_{1jk}(\zeta_{1jk}\text{-}X_{1jk}\beta\ 55\text{^})\right]$$

$$= \frac{\sigma_1^2}{h}\sum_j\sum_k\frac{1}{1+\sigma_2^2\alpha_{1j+}}\alpha_{1jk}(\zeta_{1jk}\text{-}X_{1jk}\beta\ 55\text{^})$$

$$= \frac{\sigma_1^2}{h} \sum_j \frac{\alpha_{1j+}}{1+\sigma_2^2\alpha_{1j+}} \sum_k \frac{\alpha_{1jk}}{\alpha_{1j+}} (\zeta_{1jk}-X_{1jk}\beta\ \ 55\hat{})$$

$$= \frac{\sigma_1^2}{h} \sum_j \frac{\alpha_{1j+}}{1+\sigma_2^2\alpha_{1j+}} (\zeta_{1j}-X_{1j}\beta\ \ 55\hat{})$$

$$= \frac{\sigma_1^2}{h} \left(\sum_j \frac{\alpha_{1j+}}{\sigma_2^2\alpha_{1j+}+1}\right)\left(\zeta\ \ 60{=}_1 - X\ \ 60{=}_1\beta\ \ 55\hat{}\right)$$

$$= \frac{\sigma_1^2\left(\sum_j \frac{\alpha_{1j+}}{\sigma_2^2\alpha_{1j+}+1}\right)}{1+\sigma_1^2\sum_j\left[\alpha_{1j+}-\left(\frac{\sigma_2^2\alpha_{1j+}^2}{\sigma_2^2\alpha_{1j+}+1}\right)\right]} \left(\zeta\ \ 60{=}_1 - X\ \ 60{=}_1\beta\ \ 55\hat{}\right)$$

$$= \frac{\sigma_1^2}{\sigma_1^2 + \left(\sum_j \frac{\alpha_{1j+}^2}{\sigma_2^2\alpha_{1j+}+1}\right)^{-1}} \left(\zeta\ \ 60{=}_1 - X\ \ 60{=}_1\beta\ \ 55\hat{}\}\right)$$

$$= \gamma_{11}\left(\zeta\ \ 60{=}_1 - X\ \ 60{=}_1\beta\ \ 55\hat{}\right)$$

where

$$\gamma_{11} = \frac{\sigma_1^2}{\sigma_1^2 + \left(\sum_j \frac{\alpha_{1j+}}{\sigma_2^2 \alpha_{1j+} + 1}\right)^{-1}}$$

is the state-level shrinkage factor in the two-level model.

Looking at a PSU effect, we have that

Error!

Error!

where

$$\gamma_{21j} = \frac{\sigma_2^2}{\sigma_2^2 + \frac{1}{\alpha_{1j+}}}$$

is the PSU-level shrinkage factor for the j-th sample PSU in the first state in the 2-level model.

Similarly for the other states,

$$U_i = \gamma_{1i}(\zeta\ 60{=}_i - X\ 60{=}_i \beta\ 55\text{^})$$

$$U_{ij} = \gamma_{2ij}\left[(\zeta_{ij} - X_{ij}\beta\ 55\text{^}) - U_i\right],$$

where

$$\gamma_{1i} = \frac{\sigma_1^2}{\sigma_1^2 + \left(\sum_{j=1}^{r_i} \frac{\alpha_{ij+}}{\sigma_2^2 \alpha_{ij+} + 1}\right)^{-1}}$$

and

$$\gamma_{2ij} = \frac{\sigma_2^2}{\sigma_2^2 + \frac{1}{\alpha_{ij+}}} \quad .$$

To get a better intuitive feeling for these shrinkage factors, it is useful to introduce some additional notation. Let

$$v_{ij} = Var\left[(\zeta_{ij} - X_{ij}\beta) \mid U_i\ U_{ij}\right] \quad .$$

This variance can be thought of as the measurement variance on the random perturbation for the j-th sample PSU in the i-th state given a particular manifestation from the superpopulation of state and PSU random effects. This variance may be derived as follows. Note that

$$\zeta_{ij} - X_{ij}\beta = U_i + U_{ij} + \sum_k \frac{\alpha_{ijk}}{\alpha_{ij+}} \frac{\varepsilon_{ijk}}{w_{ijk}}$$

Now, given U_i and U_{ij}, only ε_{ijk} is random. So

$$\begin{aligned} v_{ij} &= \sum_k \left(\frac{\alpha_{ijk}}{\alpha_{ij+}}\right)^2 \frac{Var(\varepsilon_{ijk} \mid U_i, U_{ij})}{w_{ijk}^2} \\ &= \sum_k \left(\frac{\alpha_{ijk}}{\alpha_{ij+}}\right)^2 \frac{1}{w_{ijk}^2 \, \pi_{ijk}(1-\pi_{ijk})} \\ &= \sum_k \frac{\alpha_{ijk}}{(\alpha_{ij+})^2} \\ &= \frac{1}{\alpha_{ij+}} \; . \end{aligned}$$

Thus,

$$\gamma_{2ij} = \frac{\sigma_2^2}{\sigma_2^2 + v_{ij}} \; .$$

Recall that σ_2^2 is the variance of the PSU perturbations on the logit scale. This is what was referred to in Chapters 1 and 3 as the process variation. So γ_{2ij} is the proportion of the total variance on the perturbation for the j-th PSU in the i-th state that is due to the process rather than to measurement error. This is closely analogous to the definition of the shrinkage factors for the linear mixed model discussed in Section C.4. If the measurement error in negligible, then there will be no shrinkage and so the estimated random effect for the PSU will be close to the design-based estimate for the PSU. If, on the other hand, the measurement error is very large, then there will be considerable shrinkage, meaning that the estimate for the PSU will be based largely on the fixed model and on the state perturbation.

Turning attention to the state-level shrinkage factors, note that similar algebra as used to derive v_{ij} can show that

$$Var\left[\, (\zeta_{ij} - X_{ij}\beta) \mid U_i \,\right] = \sigma_2^2 + v_{ij} \quad .$$

Forming the harmonic mean of these conditional variances across the r_i sample PSUs within the *i*-th state yields

$$v_{Hi} = \left[\frac{1}{r_i} \sum_{j=1}^{r_i} \frac{1}{\sigma_2^2 + v_{ij}} \right]^{-1} \quad .$$

Dividing this average conditional variance by the number of sample PSUs in the states gives us a sort of measurement error on the random perturbation for the i-th state.

Now note that the state-level shrinkage factor can be rewritten as

$$\gamma_{1i} = \frac{\sigma_1^2}{\sigma_1^2 + \frac{v_{Hi}}{r_i}}$$

So it can be thought of as the proportion of the total variance on the random perturbation for the i-th state that is due to the process rather than to measurement error. Again, this is clearly analogous to the shrinkage factors for the linear mixed model. If measurement error is negligible (due to a large sample size), then there will be no shrinkage and so the estimated random effect for the state will be close to that estimated based on the design. If, on the other hand, measurement error is large (due to a small sample size), then there will be considerable shrinkage, even to the point of saying that projections for the state should be based almost entirely on the fixed model.

Appendix D: Block Group and Tract Level Variables Considered in Small Area Estimation Models

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models

Variable	Label
DUMMY	O: Intercept Term
UCLASS9	T: Underclass Indicator
IFEM1	Female Interaction Of UCLASS9
IBLK1	Black Interaction Of UCLASS9
IHISP1	Hispanic Interaction Of UCLASS9
IOTH1	Other Interaction Of UCLASS9
PURBH	T: Percent Housing Units In Urban Areas
IFEM2	Female Interaction Of PURBH
IBLK2	Black Interaction Of PURBH
IHISP2	Hispanic Interaction Of PURBH
IOTH2	Other Interaction Of PURBH
FEMALE	O: Female Indicator
FEMBLCK	O: Black Interaction Of FEMALE
FEMHISP	O: Hispanic Interaction Of FEMALE
FEMOTHR	O: Other Interaction Of FEMALE
RACEBLCK	O: Race/Black Indicator
RACEHISP	O: Race/Hispanic Indicator
RACEOTHR	O: Race/Other Indicator
REGNOREA	O: Northeast Region Indicator
REGSOUTH	O: South Region Indicator
REGWEST	O: West Region Indicator
IFEM7	Female Interaction Of REGNOREA
IFEM8	Female Interaction Of REGSOUTH
IFEM9	Female Interaction Of REGWEST
IBLK7	Black Interaction Of REGNOREA
IBLK8	Black Interaction Of REGSOUTH
IBLK9	Black Interaction Of REGWEST
IHISP7	Hispanic Interaction Of REGNOREA
IHISP8	Hispanic Interaction Of REGSOUTH
IHISP9	Hispanic Interaction Of REGWEST
IOTH7	Other Interaction Of REGNOREA
IOTH8	Other Interaction Of REGSOUTH
IOTH9	Other Interaction Of REGWEST
PDENLEV1	O: Large MSA
PDENLEV2	O: Medium MSA
PDENLEV3	O: Small MSA
PDENLEV4	O: NonMSA, Urban
IFEM10	Female Interaction Of PDENLEV1
IFEM11	Female Interaction Of PDENLEV2
IFEM12	Female Interaction Of PDENLEV3
IFEM13	Female Interaction Of PDENLEV4
IOTH21	Other Interaction Of POPRMQU
IOTH22	Other Interaction Of POPRMCU

Variable	Label
IBLK10	Black Interaction Of PDENLEV1
IBLK11	Black Interaction Of PDENLEV2
IBLK12	Black Interaction Of PDENLEV3
IBLK13	Black Interaction Of PDENLEV4
IHISP10	Hispanic Interaction Of PDENLEV1
IHISP11	Hispanic Interaction Of PDENLEV2
IHISP12	Hispanic Interaction Of PDENLEV3
IHISP13	Hispanic Interaction Of PDENLEV4
IOTH10	Other Interaction Of PDENLEV1
IOTH11	Other Interaction Of PDENLEV2
IOTH12	Other Interaction Of PDENLEV3
IOTH13	Other Interaction Of PDENLEV4
PHH1PLN	T: Linear: Percent One Person Households
PHH1PQU	T: Quadratic: Percent One Person Households
PHH1PCU	T: Cubic: Percent One Person Households
PHH1PQR	T: Quartic: Percent One Person Households
IFEM15	Female Interaction Of PHH1PLN
IFEM16	Female Interaction Of PHH1PQU
IFEM17	Female Interaction Of PHH1PCU
IFEM18	Female Interaction Of PHH1PQR
IBLK15	Black Interaction Of PHH1PLN
IBLK16	Black Interaction Of PHH1PQU
IBLK17	Black Interaction Of PHH1PCU
IBLK18	Black Interaction Of PHH1PQR
IHISP15	Hispanic Interaction Of PHH1PLN
IHISP16	Hispanic Interaction Of PHH1PQU
IHISP17	Hispanic Interaction Of PHH1PCU
IHISP18	Hispanic Interaction Of PHH1PQR
IOTH15	Other Interaction Of PHH1PLN
IOTH16	Other Interaction Of PHH1PQU
IOTH17	Other Interaction Of PHH1PCU
IOTH18	Other Interaction Of PHH1PQR
POPRMLN	T: Linear: Average Persons Per Room
POPRMQU	T: Quadratic: Average Persons Per Room
POPRMCU	T: Cubic: Average Persons Per Room
POPRMQR	T: Quartic: Average Persons Per Room
IFEM20	Female Interaction Of POPRMLN
IFEM21	Female Interaction Of POPRMQU
IFEM22	Female Interaction Of POPRMCU
IFEM23	Female Interaction Of POPRMQR
IBLK20	Black Interaction Of POPRMLN
IBLK21	Black Interaction Of POPRMQU
IBLK22	Black Interaction Of POPRMCU
IBLK23	Black Interaction Of POPRMQR
IHISP20	Hispanic Interaction Of POPRMLN
IHISP21	Hispanic Interaction Of POPRMQU
IHISP22	Hispanic Interaction Of POPRMCU
IHISP23	Hispanic Interaction Of POPRMQR
IOTH20	Other Interaction Of POPRMLN
IOTH23	Other Interaction Of POPRMQR

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models
(continued)

Variable	Label
PAGE18LN	T: Linear: Percent Persons 0-18 Years
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years
PAGE18CU	T: Cubic: Percent Persons 0-18 Years
PAGE18QR	T: Quartic: Percent Persons 0-18 Years
IFEM25	Female Interaction Of PAGE18LN
IFEM26	Female Interaction Of PAGE18QU
IFEM27	Female Interaction Of PAGE18CU
IFEM28	Female Interaction Of PAGE18QR
IBLK25	Black Interaction Of PAGE18LN
IBLK26	Black Interaction Of PAGE18QU
IBLK27	Black Interaction Of PAGE18CU
IBLK28	Black Interaction Of PAGE18QR
IHISP25	Hispanic Interaction Of PAGE18LN
IHISP26	Hispanic Interaction Of PAGE18QU
IHISP27	Hispanic Interaction Of PAGE18CU
IHISP28	Hispanic Interaction Of PAGE18QR
IOTH25	Other Interaction Of PAGE18LN
IOTH26	Other Interaction Of PAGE18QU
IOTH27	Other Interaction Of PAGE18CU
IOTH28	Other Interaction Of PAGE18QR
PAGE24LN	T: Linear: Percent Persons 19-24 Years
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years
PAGE24CU	T: Cubic: Percent Persons 19-24 Years
PAGE24QR	T: Quartic: Percent Persons 19-24 Years
IFEM30	Female Interaction Of PAGE24LN
IFEM31	Female Interaction Of PAGE24QU
IFEM32	Female Interaction Of PAGE24CU
IFEM33	Female Interaction Of PAGE24QR
IBLK30	Black Interaction Of PAGE24LN
IBLK31	Black Interaction Of PAGE24QU
IBLK32	Black Interaction Of PAGE24CU
IBLK33	Black Interaction Of PAGE24QR
IHISP30	Hispanic Interaction Of PAGE24LN
IHISP31	Hispanic Interaction Of PAGE24QU
IHISP32	Hispanic Interaction Of PAGE24CU
IHISP33	Hispanic Interaction Of PAGE24QR
IOTH30	Other Interaction Of PAGE24LN
IOTH31	Other Interaction Of PAGE24QU
IOTH32	Other Interaction Of PAGE24CU
IOTH33	Other Interaction Of PAGE24QR
PAGE34LN	T: Linear: Percent Persons 25-34 Years
PAGE34QU	T: Quadratic: Percent Persons 25-34 Years
PAGE34CU	T: Cubic: Percent Persons 25-34 Years
PAGE34QR	T: Quartic: Percent Persons 25-34 Years
IFEM35	Female Interaction Of PAGE34LN
IFEM36	Female Interaction Of PAGE34QU
IFEM37	Female Interaction Of PAGE34CU
IFEM38	Female Interaction Of PAGE34QR
IHISP48	Hispanic Interaction Of PAGE54QR
IOTH45	Other Interaction Of PAGE54LN
IOTH46	Other Interaction Of PAGE54QU
IOTH47	Other Interaction Of PAGE54CU

Variable	Label
IBLK35	Black Interaction Of PAGE34LN
IBLK36	Black Interaction Of PAGE34QU
IBLK37	Black Interaction Of PAGE34CU
IBLK38	Black Interaction Of PAGE34QR
IHISP35	Hispanic Interaction Of PAGE34LN
IHISP36	Hispanic Interaction Of PAGE34QU
IHISP37	Hispanic Interaction Of PAGE34CU
IHISP38	Hispanic Interaction Of PAGE34QR
IOTH35	Other Interaction Of PAGE34LN
IOTH36	Other Interaction Of PAGE34QU
IOTH37	Other Interaction Of PAGE34CU
IOTH38	Other Interaction Of PAGE34QR
PAGE44LN	T: Linear: Percent Persons 35-44 Years
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years
PAGE44CU	T: Cubic: Percent Persons 35-44 Years
PAGE44QR	T: Quartic: Percent Persons 35-44 Years
IFEM40	Female Interaction Of PAGE44LN
IFEM41	Female Interaction Of PAGE44QU
IFEM42	Female Interaction Of PAGE44CU
IFEM43	Female Interaction Of PAGE44QR
IBLK40	Black Interaction Of PAGE44LN
IBLK41	Black Interaction Of PAGE44QU
IBLK42	Black Interaction Of PAGE44CU
IBLK43	Black Interaction Of PAGE44QR
IHISP40	Hispanic Interaction Of PAGE44LN
IHISP41	Hispanic Interaction Of PAGE44QU
IHISP42	Hispanic Interaction Of PAGE44CU
IHISP43	Hispanic Interaction Of PAGE44QR
IOTH40	Other Interaction Of PAGE44LN
IOTH41	Other Interaction Of PAGE44QU
IOTH42	Other Interaction Of PAGE44CU
IOTH43	Other Interaction Of PAGE44QR
PAGE54LN	T: Linear: Percent Persons 45-54 Years
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years
PAGE54CU	T: Cubic: Percent Persons 45-54 Years
PAGE54QR	T: Quartic: Percent Persons 45-54 Years
IFEM45	Female Interaction Of PAGE54LN
IFEM46	Female Interaction Of PAGE54QU
IFEM47	Female Interaction Of PAGE54CU
IFEM48	Female Interaction Of PAGE54QR
IBLK45	Black Interaction Of PAGE54LN
IBLK46	Black Interaction Of PAGE54QU
IBLK47	Black Interaction Of PAGE54CU
IBLK48	Black Interaction Of PAGE54QR
IHISP45	Hispanic Interaction Of PAGE54LN
IHISP46	Hispanic Interaction Of PAGE54QU
IHISP47	Hispanic Interaction Of PAGE54CU
IOTH48	Other Interaction Of PAGE54QR
PAGE64LN	T: Linear: Percent Persons 55-64 Years
PAGE64QU	T: Quadratic: Percent Persons 55-64 Years
PAGE64CU	T: Cubic: Percent Persons 55-64 Years

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models

(continued)

Variable	Label
PAGE64QR	T: Quartic: Percent Persons 55-64 Years
IFEM50	Female Interaction Of PAGE64LN
IFEM51	Female Interaction Of PAGE64QU
IFEM52	Female Interaction Of PAGE64CU
IFEM53	Female Interaction Of PAGE64QR
IBLK50	Black Interaction Of PAGE64LN
IBLK51	Black Interaction Of PAGE64QU
IBLK52	Black Interaction Of PAGE64CU
IBLK53	Black Interaction Of PAGE64QR
IHISP50	Hispanic Interaction Of PAGE64LN
IHISP51	Hispanic Interaction Of PAGE64QU
IHISP52	Hispanic Interaction Of PAGE64CU
IHISP53	Hispanic Interaction Of PAGE64QR
IOTH50	Other Interaction Of PAGE64LN
IOTH51	Other Interaction Of PAGE64QU
IOTH52	Other Interaction Of PAGE64CU
IOTH53	Other Interaction Of PAGE64QR
PSCH8LN	T: Linear: Percent 0-8 Years Of School
PSCH8QU	T: Quadratic: Percent 0-8 Years Of School
PSCH8CU	T: Cubic: Percent 0-8 Years Of School
PSCH8QR	T: Quartic: Percent 0-8 Years Of School
IFEM55	Female Interaction Of PSCH8LN
IFEM56	Female Interaction Of PSCH8QU
IFEM57	Female Interaction Of PSCH8CU
IFEM58	Female Interaction Of PSCH8QR
IBLK55	Black Interaction Of PSCH8LN
IBLK56	Black Interaction Of PSCH8QU
IBLK57	Black Interaction Of PSCH8CU
IBLK58	Black Interaction Of PSCH8QR
IHISP55	Hispanic Interaction Of PSCH8LN
IHISP56	Hispanic Interaction Of PSCH8QU
IHISP57	Hispanic Interaction Of PSCH8CU
IHISP58	Hispanic Interaction Of PSCH8QR
IOTH55	Other Interaction Of PSCH8LN
IOTH56	Other Interaction Of PSCH8QU
IOTH57	Other Interaction Of PSCH8CU
IOTH58	Other Interaction Of PSCH8QR
PSCH12LN	T: Linear: Percent 9-12 Years & No HS Diploma
PSCH12QU	T: Quadratic: Percent 9-12 Years & No HS Diploma
PSCH12CU	T: Cubic: Percent 9-12 Years & No HS Diploma
PSCH12QR	T: Quartic: Percent 9-12 Years & No HS Diploma
IFEM60	Female Interaction Of PSCH12LN
IFEM61	Female Interaction Of PSCH12QU
IFEM62	Female Interaction Of PSCH12CU
IBLK78	Black Interaction Of PSCHCOQR
IHISP75	Hispanic Interaction Of PSCHCOLN
IHISP76	Hispanic Interaction Of PSCHCOQU
IHISP77	Hispanic Interaction Of PSCHCOCU
IHISP78	Hispanic Interaction Of PSCHCOQR
IOTH75	Other Interaction Of PSCHCOLN

Variable	Label
IFEM63	Female Interaction Of PSCH12QR
IBLK60	Black Interaction Of PSCH12LN
IBLK61	Black Interaction Of PSCH12QU
IBLK62	Black Interaction Of PSCH12CU
IBLK63	Black Interaction Of PSCH12QR
IHISP60	Hispanic Interaction Of PSCH12LN
IHISP61	Hispanic Interaction Of PSCH12QU
IHISP62	Hispanic Interaction Of PSCH12CU
IHISP63	Hispanic Interaction Of PSCH12QR
IOTH60	Other Interaction Of PSCH12LN
IOTH61	Other Interaction Of PSCH12QU
IOTH62	Other Interaction Of PSCH12CU
IOTH63	Other Interaction Of PSCH12QR
PSCHASLN	T: Linear: Percent Associates Degree
PSCHASQU	T: Quadratic: Percent Associates Degree
PSCHASCU	T: Cubic: Percent Associates Degree
PSCHASQR	T: Quartic: Percent Associates Degree
IFEM65	Female Interaction Of PSCHASLN
IFEM66	Female Interaction Of PSCHASQU
IFEM67	Female Interaction Of PSCHASCU
IFEM68	Female Interaction Of PSCHASQR
IBLK65	Black Interaction Of PSCHASLN
IBLK66	Black Interaction Of PSCHASQU
IBLK67	Black Interaction Of PSCHASCU
IBLK68	Black Interaction Of PSCHASQR
IHISP65	Hispanic Interaction Of PSCHASLN
IHISP66	Hispanic Interaction Of PSCHASQU
IHISP67	Hispanic Interaction Of PSCHASCU
IHISP68	Hispanic Interaction Of PSCHASQR
IOTH65	Other Interaction Of PSCHASLN
IOTH66	Other Interaction Of PSCHASQU
IOTH67	Other Interaction Of PSCHASCU
IOTH68	Other Interaction Of PSCHASQR
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree
PSCHCOCU	T: Cubic: Bachelors, Graduate, Or Professional Degree
PSCHCOQR	T: Quartic: Bachelors, Graduate, Or Professional Degree
IFEM75	Female Interaction Of PSCHCOLN
IFEM76	Female Interaction Of PSCHCOQU
IFEM77	Female Interaction Of PSCHCOCU
IFEM78	Female Interaction Of PSCHCOQR
IBLK75	Black Interaction Of PSCHCOLN
IBLK76	Black Interaction Of PSCHCOQU
IBLK77	Black Interaction Of PSCHCOCU
IOTH76	Other Interaction Of PSCHCOQU
IOTH77	Other Interaction Of PSCHCOCU
IOTH78	Other Interaction Of PSCHCOQR
PSCHSCLN	T: Linear: Percent Some College And No Degree
PSCHSCQU	T: Quadratic: Percent Some College And No Degree
PSCHSCCU	T: Cubic: Percent Some College And No Degree
PSCHSCQR	T: Quartic: Percent Some College And No Degree

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models
(continued)

Variable	Label
IFEM80	Female Interaction Of PSCHSCLN
IFEM81	Female Interaction Of PSCHSCQU
IFEM82	Female Interaction Of PSCHSCCU
IFEM83	Female Interaction Of PSCHSCQR
IBLK80	Black Interaction Of PSCHSCLN
IBLK81	Black Interaction Of PSCHSCQU
IBLK82	Black Interaction Of PSCHSCCU
IBLK83	Black Interaction Of PSCHSCQR
IHISP80	Hispanic Interaction Of PSCHSCLN
IHISP81	Hispanic Interaction Of PSCHSCQU
IHISP82	Hispanic Interaction Of PSCHSCCU
IHISP83	Hispanic Interaction Of PSCHSCQR
IOTH80	Other Interaction Of PSCHSCLN
IOTH81	Other Interaction Of PSCHSCQU
IOTH82	Other Interaction Of PSCHSCCU
IOTH83	Other Interaction Of PSCHSCQR
PPOVERLN	T: Linear: Percent Families Below Poverty Level
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level
PPOVERCU	T: Cubic: Percent Families Below Poverty Level
PPOVERQR	T: Quartic: Percent Families Below Poverty Level
IFEM85	Female Interaction Of PPOVERLN
IFEM86	Female Interaction Of PPOVERQU
IFEM87	Female Interaction Of PPOVERCU
IFEM88	Female Interaction Of PPOVERQR
IBLK85	Black Interaction Of PPOVERLN
IBLK86	Black Interaction Of PPOVERQU
IBLK87	Black Interaction Of PPOVERCU
IBLK88	Black Interaction Of PPOVERQR
IHISP85	Hispanic Interaction Of PPOVERLN
IHISP86	Hispanic Interaction Of PPOVERQU
IHISP87	Hispanic Interaction Of PPOVERCU
IHISP88	Hispanic Interaction Of PPOVERQR
IOTH85	Other Interaction Of PPOVERLN
IOTH86	Other Interaction Of PPOVERQU
IOTH87	Other Interaction Of PPOVERCU
IOTH88	Other Interaction Of PPOVERQR
PPUBASLN	T: Linear: % HHS With Public Assist Income
PPUBASQU	T: Quadratic: % HHS With Public Assist Income
PPUBASCU	T: Cubic: % HHS With Public Assist Income
PPUBASQR	T: Quartic: % HHS With Public Assist Income
IFEM90	Female Interaction Of PPUBASLN
IFEM91	Female Interaction Of PPUBASQU
IFEM92	Female Interaction Of PPUBASCU
IBLK101	Black Interaction Of PBLACKQU
IBLK102	Black Interaction Of PBLACKCU
IBLK103	Black Interaction Of PBLACKQR
IHISP100	Hispanic Interaction Of PBLACKLN
IHISP101	Hispanic Interaction Of PBLACKQU
IHISP102	Hispanic Interaction Of PBLACKCU
IHISP103	Hispanic Interaction Of PBLACKQR
IOTH100	Other Interaction Of PBLACKLN
IOTH101	Other Interaction Of PBLACKQU

Variable	Label
IFEM93	Female Interaction Of PPUBASQR
IBLK90	Black Interaction Of PPUBASLN
IBLK91	Black Interaction Of PPUBASQU
IBLK92	Black Interaction Of PPUBASCU
IBLK93	Black Interaction Of PPUBASQR
IHISP90	Hispanic Interaction Of PPUBASLN
IHISP91	Hispanic Interaction Of PPUBASQU
IHISP92	Hispanic Interaction Of PPUBASCU
IHISP93	Hispanic Interaction Of PPUBASQR
IOTH90	Other Interaction Of PPUBASLN
IOTH91	Other Interaction Of PPUBASQU
IOTH92	Other Interaction Of PPUBASCU
IOTH93	Other Interaction Of PPUBASQR
P64DISLN	T: Linear: Percent 16-64 With A Work Disability
P64DISQU	T: Quadratic: Percent 16-64 With A Work Disability
P64DISCU	T: Cubic: Percent 16-64 With A Work Disability
P64DISQR	T: Quartic: Percent 16-64 With A Work Disability
IFEM95	Female Interaction Of P64DISLN
IFEM96	Female Interaction Of P64DISQU
IFEM97	Female Interaction Of P64DISCU
IFEM98	Female Interaction Of P64DISQR
IBLK95	Black Interaction Of P64DISLN
IBLK96	Black Interaction Of P64DISQU
IBLK97	Black Interaction Of P64DISCU
IBLK98	Black Interaction Of P64DISQR
IHISP95	Hispanic Interaction Of P64DISLN
IHISP96	Hispanic Interaction Of P64DISQU
IHISP97	Hispanic Interaction Of P64DISCU
IHISP98	Hispanic Interaction Of P64DISQR
IOTH95	Other Interaction Of P64DISLN
IOTH96	Other Interaction Of P64DISQU
IOTH97	Other Interaction Of P64DISCU
IOTH98	Other Interaction Of P64DISQR
PBLACKLN	T: Linear: Percent Black Nonhispanic
PBLACKQU	T: Quadratic: Percent Black Nonhispanic
PBLACKCU	T: Cubic: Percent Black Nonhispanic
PBLACKQR	T: Quartic: Percent Black Nonhispanic
IFEM100	Female Interaction Of PBLACKLN
IFEM101	Female Interaction Of PBLACKQU
IFEM102	Female Interaction Of PBLACKCU
IFEM103	Female Interaction Of PBLACKQR
IBLK100	Black Interaction Of PBLACKLN
IOTH102	Other Interaction Of PBLACKCU
IOTH103	Other Interaction Of PBLACKQR
PHISPLN	T: Linear: Percent Hispanic
PHISPQU	T: Quadratic: Percent Hispanic
PHISPCU	T: Cubic: Percent Hispanic
PHISPQR	T: Quartic: Percent Hispanic
IFEM105	Female Interaction Of PHISPLN
IFEM106	Female Interaction Of PHISPQU
IFEM107	Female Interaction Of PHISPCU

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models
(continued)

Variable	Label
IFEM108	Female Interaction Of PHISPQR
IBLK105	Black Interaction Of PHISPLN
IBLK106	Black Interaction Of PHISPQU
IBLK107	Black Interaction Of PHISPCU
IBLK108	Black Interaction Of PHISPQR
IHISP105	Hispanic Interaction Of PHISPLN
IHISP106	Hispanic Interaction Of PHISPQU
IHISP107	Hispanic Interaction Of PHISPCU
IHISP108	Hispanic Interaction Of PHISPQR
IOTH105	Other Interaction Of PHISPLN
IOTH106	Other Interaction Of PHISPQU
IOTH107	Other Interaction Of PHISPCU
IOTH108	Other Interaction Of PHISPQR
POTHLN	T: Linear: Percent Other Race/Hispanicity
POTHQU	T: Quadratic: Percent Other Race/Hispanicity
POTHCU	T: Cubic: Percent Other Race/Hispanicity
POTHQR	T: Quartic: Percent Other Race/Hispanicity
IFEM110	Female Interaction Of POTHLN
IFEM111	Female Interaction Of POTHQU
IFEM112	Female Interaction Of POTHCU
IFEM113	Female Interaction Of POTHQR
IBLK110	Black Interaction Of POTHLN
IBLK111	Black Interaction Of POTHQU
IBLK112	Black Interaction Of POTHCU
IBLK113	Black Interaction Of POTHQR
IHISP110	Hispanic Interaction Of POTHLN
IHISP111	Hispanic Interaction Of POTHQU
IHISP112	Hispanic Interaction Of POTHCU
IHISP113	Hispanic Interaction Of POTHQR
IOTH110	Other Interaction Of POTHLN
IOTH111	Other Interaction Of POTHQU
IOTH112	Other Interaction Of POTHCU
IOTH113	Other Interaction Of POTHQR
PHHF18LN	T: Linear: % F-Headed HH W/No Spouse & Chld <18
PHHF18QU	T: Quad.: % F-Headed HH W/No Spouse & Chld <18
PHHF18CU	T: Cubic: % F-Headed HH W/No Spouse & Chld <18
PHHF18QR	T: Quartic: % F-Headed HH W/No Spouse & Chld <18
IFEM120	Female Interaction Of PHHF18LN
IFEM121	Female Interaction Of PHHF18QU
IFEM132	Female Interaction Of PFNEVCU
IFEM133	Female Interaction Of PFNEVQR
IBLK130	Black Interaction Of PFNEVLN
IBLK131	Black Interaction Of PFNEVQU
IBLK132	Black Interaction Of PFNEVCU
IBLK133	Black Interaction Of PFNEVQR
IHISP130	Hispanic Interaction Of PFNEVLN
IHISP131	Hispanic Interaction Of PFNEVQU
IHISP132	Hispanic Interaction Of PFNEVCU
IHISP133	Hispanic Interaction Of PFNEVQR
IOTH130	Other Interaction Of PFNEVLN
IFEM122	Female Interaction Of PHHF18CU
IFEM123	Female Interaction Of PHHF18QR
IBLK120	Black Interaction Of PHHF18LN
IBLK121	Black Interaction Of PHHF18QU
IBLK122	Black Interaction Of PHHF18CU
IBLK123	Black Interaction Of PHHF18QR
IHISP120	Hispanic Interaction Of PHHF18LN
IHISP121	Hispanic Interaction Of PHHF18QU
IHISP122	Hispanic Interaction Of PHHF18CU
IHISP123	Hispanic Interaction Of PHHF18QR
IOTH120	Other Interaction Of PHHF18LN
IOTH121	Other Interaction Of PHHF18QU
IOTH122	Other Interaction Of PHHF18CU
IOTH123	Other Interaction Of PHHF18QR
PFLABLN	T: Linear: % Females 16+ Years Old In Labor Force
PFLABQU	T: Quadratic: % Females 16+ Years Old In Labor Force
PFLABCU	T: Cubic: % Females 16+ Years Old In Labor Force
PFLABQR	T: Quartic: % Females 16+ Years Old In Labor Force
IFEM125	Female Interaction Of PFLABLN
IFEM126	Female Interaction Of PFLABQU
IFEM127	Female Interaction Of PFLABCU
IFEM128	Female Interaction Of PFLABQR
IBLK125	Black Interaction Of PFLABLN
IBLK126	Black Interaction Of PFLABQU
IBLK127	Black Interaction Of PFLABCU
IBLK128	Black Interaction Of PFLABQR
IHISP125	Hispanic Interaction Of PFLABLN
IHISP126	Hispanic Interaction Of PFLABQU
IHISP127	Hispanic Interaction Of PFLABCU
IHISP128	Hispanic Interaction Of PFLABQR
IOTH125	Other Interaction Of PFLABLN
IOTH126	Other Interaction Of PFLABQU
IOTH127	Other Interaction Of PFLABCU
IOTH128	Other Interaction Of PFLABQR
PFNEVLN	T: Linear: Percent Females Never Married
PFNEVQU	T: Quadratic: Percent Females Never Married
PFNEVCU	T: Cubic: Percent Females Never Married
PFNEVQR	T: Quartic: Percent Females Never Married
IFEM130	Female Interaction Of PFNEVLN
IFEM131	Female Interaction Of PFNEVQU
IOTH131	Other Interaction Of PFNEVQU
IOTH132	Other Interaction Of PFNEVCU
IOTH133	Other Interaction Of PFNEVQR
PFNOTLN	T: Linear: % Females Separated, Divorced Or Widowed
PFNOTQU	T: Quadratic: % Fem Separated, Divorced Or Widowed
PFNOTCU	T: Cubic: % Fem Separated, Divorced Or Widowed
PFNOTQR	T: Quartic: % Fem Separated, Divorced Or Widowed
IFEM135	Female Interaction Of PFNOTLN
IFEM136	Female Interaction Of PFNOTQU
IFEM137	Female Interaction Of PFNOTCU
IFEM138	Female Interaction Of PFNOTQR

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models

(continued)

Variable	Label
IBLK135	Black Interaction Of PFNOTLN
IBLK136	Black Interaction Of PFNOTQU
IBLK137	Black Interaction Of PFNOTCU
IBLK138	Black Interaction Of PFNOTQR
IHISP135	Hispanic Interaction Of PFNOTLN
IHISP136	Hispanic Interaction Of PFNOTQU
IHISP137	Hispanic Interaction Of PFNOTCU
IHISP138	Hispanic Interaction Of PFNOTQR
IOTH135	Other Interaction Of PFNOTLN
IOTH136	Other Interaction Of PFNOTQU
IOTH137	Other Interaction Of PFNOTCU
IOTH138	Other Interaction Of PFNOTQR
PMNEVLN	T: Linear: Percent Males Never Married
PMNEVQU	T: Quadratic: Percent Males Never Married
PMNEVCU	T: Cubic: Percent Males Never Married
PMNEVQR	T: Quartic: Percent Males Never Married
IFEM140	Female Interaction Of PMNEVLN
IFEM141	Female Interaction Of PMNEVQU
IFEM142	Female Interaction Of PMNEVCU
IFEM143	Female Interaction Of PMNEVQR
IBLK140	Black Interaction Of PMNEVLN
IBLK141	Black Interaction Of PMNEVQU
IBLK142	Black Interaction Of PMNEVCU
IBLK143	Black Interaction Of PMNEVQR
IHISP140	Hispanic Interaction Of PMNEVLN
IHISP141	Hispanic Interaction Of PMNEVQU
IHISP142	Hispanic Interaction Of PMNEVCU
IHISP143	Hispanic Interaction Of PMNEVQR
IOTH140	Other Interaction Of PMNEVLN
IOTH141	Other Interaction Of PMNEVQU
IOTH142	Other Interaction Of PMNEVCU
IOTH143	Other Interaction Of PMNEVQR
PMLABLN	T: Linear: % Males 16+ Years Old In Labor Force
PMLABQU	T: Quadratic: % Males 16+ Years Old In Labor Force
PMLABCU	T: Cubic: % Males 16+ Years Old In Labor Force
PMLABQR	T: Quartic: % Males 16+ Years Old In Labor Force
IFEM145	Female Interaction Of PMLABLN
POLDHUQR	T: Quartic: % Housing Units Built 1939 Or Earlier
IFEM155	Female Interaction Of POLDHULN
IFEM156	Female Interaction Of POLDHUQU
IFEM157	Female Interaction Of POLDHUCU
IFEM158	Female Interaction Of POLDHUQR
IBLK155	Black Interaction Of POLDHULN
IBLK156	Black Interaction Of POLDHUQU
IBLK157	Black Interaction Of POLDHUCU
IBLK158	Black Interaction Of POLDHUQR
IHISP155	Hispanic Interaction Of POLDHULN
IHISP156	Hispanic Interaction Of POLDHUQU
IHISP157	Hispanic Interaction Of POLDHUCU
IHISP158	Hispanic Interaction Of POLDHUQR

Variable	Label
IFEM146	Female Interaction Of PMLABQU
IFEM147	Female Interaction Of PMLABCU
IFEM148	Female Interaction Of PMLABQR
IBLK145	Black Interaction Of PMLABLN
IBLK146	Black Interaction Of PMLABQU
IBLK147	Black Interaction Of PMLABCU
IBLK148	Black Interaction Of PMLABQR
IHISP145	Hispanic Interaction Of PMLABLN
IHISP146	Hispanic Interaction Of PMLABQU
IHISP147	Hispanic Interaction Of PMLABCU
IHISP148	Hispanic Interaction Of PMLABQR
IOTH145	Other Interaction Of PMLABLN
IOTH146	Other Interaction Of PMLABQU
IOTH147	Other Interaction Of PMLABCU
IOTH148	Other Interaction Of PMLABQR
PMNOTLN	T: Linear: % M Separated, Divorced Or Widowed
PMNOTQU	T: Quadratic: % M Separated, Divorced Or Widowed
PMNOTCU	T: Cubic: % M Separated, Divorced Or Widowed
PMNOTQR	T: Quartic: % M Separated, Divorced Or Widowed
IFEM150	Female Interaction Of PMNOTLN
IFEM151	Female Interaction Of PMNOTQU
IFEM152	Female Interaction Of PMNOTCU
IFEM153	Female Interaction Of PMNOTQR
IBLK150	Black Interaction Of PMNOTLN
IBLK151	Black Interaction Of PMNOTQU
IBLK152	Black Interaction Of PMNOTCU
IBLK153	Black Interaction Of PMNOTQR
IHISP150	Hispanic Interaction Of PMNOTLN
IHISP151	Hispanic Interaction Of PMNOTQU
IHISP152	Hispanic Interaction Of PMNOTCU
IHISP153	Hispanic Interaction Of PMNOTQR
IOTH150	Other Interaction Of PMNOTLN
IOTH151	Other Interaction Of PMNOTQU
IOTH152	Other Interaction Of PMNOTCU
IOTH153	Other Interaction Of PMNOTQR
POLDHULN	T: Linear: % Housing Units Built 1939 Or Earlier
POLDHUQU	T: Quadratic: % Housing Units Built 1939 Or Earlier
POLDHUCU	T: Cubic: % Housing Units Built 1939 Or Earlier
IOTH155	Other Interaction Of POLDHULN
IOTH156	Other Interaction Of POLDHUQU
IOTH157	Other Interaction Of POLDHUCU
IOTH158	Other Interaction Of POLDHUQR
P40HULN	T: Linear: Percent Housing Units Built 1940-1949
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949
P40HUCU	T: Cubic: Percent Housing Units Built 1940-1949
P40HUQR	T: Quartic: Percent Housing Units Built 1940-1949
IFEM160	Female Interaction Of P40HULN
IFEM161	Female Interaction Of P40HUQU
IFEM162	Female Interaction Of P40HUCU
IFEM163	Female Interaction Of P40HUQR
IBLK160	Black Interaction Of P40HULN
IBLK161	Black Interaction Of P40HUQU

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models

(continued)

Variable	Label
IBLK162	Black Interaction Of P40HUCU
IBLK163	Black Interaction Of P40HUQR
IHISP160	Hispanic Interaction Of P40HULN
IHISP161	Hispanic Interaction Of P40HUQU
IHISP162	Hispanic Interaction Of P40HUCU
IHISP163	Hispanic Interaction Of P40HUQR
IOTH160	Other Interaction Of P40HULN
IOTH161	Other Interaction Of P40HUQU
IOTH162	Other Interaction Of P40HUCU
IOTH163	Other Interaction Of P40HUQR
PRENTLN	T: Linear: Percent Housing Rented
PRENTQU	T: Quadratic: Percent Housing Rented
PRENTCU	T: Cubic: Percent Housing Rented
PRENTQR	T: Quartic: Percent Housing Rented
IFEM165	Female Interaction Of PRENTLN
IFEM166	Female Interaction Of PRENTQU
IFEM167	Female Interaction Of PRENTCU
IFEM168	Female Interaction Of PRENTQR
IBLK165	Black Interaction Of PRENTLN
IBLK166	Black Interaction Of PRENTQU
IBLK167	Black Interaction Of PRENTCU
IBLK168	Black Interaction Of PRENTQR
IHISP165	Hispanic Interaction Of PRENTLN
IHISP166	Hispanic Interaction Of PRENTQU
IHISP167	Hispanic Interaction Of PRENTCU
IHISP168	Hispanic Interaction Of PRENTQR
IOTH165	Other Interaction Of PRENTLN
IOTH166	Other Interaction Of PRENTQU
IOTH167	Other Interaction Of PRENTCU
IOTH168	Other Interaction Of PRENTQR
PURBPLN	T: Linear: Percent Population In Urban Areas
PURBPQU	T: Quadratic: Percent Population In Urban Areas
PURBPCU	T: Cubic: Percent Population In Urban Areas
PURBPQR	T: Quartic: Percent Population In Urban Areas
IFEM170	Female Interaction Of PURBPLN
ADRAT1LN	C: Linr: Death Rate With Explicit Mention Of Alcohol
ADRAT1QU	C: Quad: Death Rate With Explicit Mention Of Alcohol
ADRAT1CU	C: Cubic: Death Rate With Explicit Mention Of Alcohol
ADRAT1QR	C: Quart: Death Rate With Explicit Mention Of Alcohol
IFEM180	Female Interaction Of ADRAT1LN
IFEM181	Female Interaction Of ADRAT1QU
IFEM182	Female Interaction Of ADRAT1CU
IFEM183	Female Interaction Of ADRAT1QR
IBLK180	Black Interaction Of ADRAT1LN
IBLK181	Black Interaction Of ADRAT1QU
IBLK182	Black Interaction Of ADRAT1CU
IBLK183	Black Interaction Of ADRAT1QR
IHISP180	Hispanic Interaction Of ADRAT1LN
IHISP181	Hispanic Interaction Of ADRAT1QU
IHISP182	Hispanic Interaction Of ADRAT1CU
IHISP183	Hispanic Interaction Of ADRAT1QR
IFEM171	Female Interaction Of PURBPQU
IFEM172	Female Interaction Of PURBPCU
IFEM173	Female Interaction Of PURBPQR
IBLK170	Black Interaction Of PURBPLN
IBLK171	Black Interaction Of PURBPQU
IBLK172	Black Interaction Of PURBPCU
IBLK173	Black Interaction Of PURBPQR
IHISP170	Hispanic Interaction Of PURBPLN
IHISP171	Hispanic Interaction Of PURBPQU
IHISP172	Hispanic Interaction Of PURBPCU
IHISP173	Hispanic Interaction Of PURBPQR
IOTH170	Other Interaction Of PURBPLN
IOTH171	Other Interaction Of PURBPQU
IOTH172	Other Interaction Of PURBPCU
IOTH173	Other Interaction Of PURBPQR
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases
ADRATECU	C: Cubic: Death Rate For All Alcohol-Related Cases
ADRATEQR	C: Quartic: Death Rate For All Alcohol-Related Cases
IFEM175	Female Interaction Of ADRATELN
IFEM176	Female Interaction Of ADRATEQU
IFEM177	Female Interaction Of ADRATECU
IFEM178	Female Interaction Of ADRATEQR
IBLK175	Black Interaction Of ADRATELN
IBLK176	Black Interaction Of ADRATEQU
IBLK177	Black Interaction Of ADRATECU
IBLK178	Black Interaction Of ADRATEQR
IHISP175	Hispanic Interaction Of ADRATELN
IHISP176	Hispanic Interaction Of ADRATEQU
IHISP177	Hispanic Interaction Of ADRATECU
IHISP178	Hispanic Interaction Of ADRATEQR
IOTH175	Other Interaction Of ADRATELN
IOTH176	Other Interaction Of ADRATEQU
IOTH177	Other Interaction Of ADRATECU
IOTH178	Other Interaction Of ADRATEQR
IOTH180	Other Interaction Of ADRAT1LN
IOTH181	Other Interaction Of ADRAT1QU
IOTH182	Other Interaction Of ADRAT1CU
IOTH183	Other Interaction Of ADRAT1QR
V18FLN	C: Linear: Marijuana Posession Arrest Rate
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate
V18FCU	C: Cubic: Marijuana Posession Arrest Rate
V18FQR	C: Quartic: Marijuana Posession Arrest Rate
IFEM185	Female Interaction Of V18FLN
IFEM186	Female Interaction Of V18FQU
IFEM187	Female Interaction Of V18FCU
IFEM188	Female Interaction Of V18FQR
IBLK185	Black Interaction Of V18FLN
IBLK186	Black Interaction Of V18FQU
IBLK187	Black Interaction Of V18FCU
IBLK188	Black Interaction Of V18FQR

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models

(continued)

Variable	Label
IHISP185	Hispanic Interaction Of V18FLN
IHISP186	Hispanic Interaction Of V18FQU
IHISP187	Hispanic Interaction Of V18FCU
IHISP188	Hispanic Interaction Of V18FQR
IOTH185	Other Interaction Of V18FLN
IOTH186	Other Interaction Of V18FQU
IOTH187	Other Interaction Of V18FCU
IOTH188	Other Interaction Of V18FQR
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate
V18BCU	C: Cubic: Marijuana Sale/Manufacture Arrest Rate
V18BQR	C: Quartic: Marijuana Sale/Manufacture Arrest Rate
IFEM190	Female Interaction Of V18BLN
IFEM191	Female Interaction Of V18BQU
IFEM192	Female Interaction Of V18BCU
IFEM193	Female Interaction Of V18BQR
IBLK190	Black Interaction Of V18BLN
IBLK191	Black Interaction Of V18BQU
IBLK192	Black Interaction Of V18BCU
IBLK193	Black Interaction Of V18BQR
IHISP190	Hispanic Interaction Of V18BLN
IHISP191	Hispanic Interaction Of V18BQU
IHISP192	Hispanic Interaction Of V18BCU
IHISP193	Hispanic Interaction Of V18BQR
IOTH190	Other Interaction Of V18BLN
IOTH191	Other Interaction Of V18BQU
IOTH192	Other Interaction Of V18BCU
IOTH193	Other Interaction Of V18BQR
V18ALN	C: Linr: Opium/Cocaine & Deriv Sale/Manuf Arst Rate
V18AQU	C: Quad: Opium/Cocaine & Deriv Sale/Manuf Arst Rate
V18ACU	C: Cub: Opium/Cocaine & Deriv Sale/Manuf Arst Rate
V18AQR	C: Qurt: Opium/Cocaine & Deriv Sale/Manuf Arst Rate
IOTH205	Other Interaction Of V18HLN
IOTH206	Other Interaction Of V18HQU
IOTH207	Other Interaction Of V18HCU
IOTH208	Other Interaction Of V18HQR
V18DLN	C: Linear: Other (Non-Narcotics) Sale/Manuf Arst Rate
V18DQU	C: Quad: Other (Non-Narcotics) Sale/Manuf Arst Rate
V18DCU	C: Cubic: Other (Non-Narcotics) Sale/Manuf Arst Rate
V18DQR	C: Quartic: Other (Non-Narcotics) Sale/Manuf Arst Rate
IFEM210	Female Interaction Of V18DLN
IFEM211	Female Interaction Of V18DQU
IFEM212	Female Interaction Of V18DCU
IFEM213	Female Interaction Of V18DQR
IBLK210	Black Interaction Of V18DLN
IBLK211	Black Interaction Of V18DQU
IBLK212	Black Interaction Of V18DCU
IBLK213	Black Interaction Of V18DQR
IHISP210	Hispanic Interaction Of V18DLN
IHISP211	Hispanic Interaction Of V18DQU
IHISP212	Hispanic Interaction Of V18DCU

Variable	Label
IFEM200	Female Interaction Of V18ALN
IFEM201	Female Interaction Of V18AQU
IFEM202	Female Interaction Of V18ACU
IFEM203	Female Interaction Of V18AQR
IBLK200	Black Interaction Of V18ALN
IBLK201	Black Interaction Of V18AQU
IBLK202	Black Interaction Of V18ACU
IBLK203	Black Interaction Of V18AQR
IHISP200	Hispanic Interaction Of V18ALN
IHISP201	Hispanic Interaction Of V18AQU
IHISP202	Hispanic Interaction Of V18ACU
IHISP203	Hispanic Interaction Of V18AQR
IOTH200	Other Interaction Of V18ALN
IOTH201	Other Interaction Of V18AQU
IOTH202	Other Interaction Of V18ACU
IOTH203	Other Interaction Of V18AQR
V18HLN	C: Linear: Other (Non-Narcotics) Posession Arrest Rate
V18HQU	C: Quad: Other (Non-Narcotics) Posession Arst Rate
V18HCU	C: Cubic: Other (Non-Narcotics) Posession Arrest Rate
V18HQR	C: Quartic: Other (Non-Narcotics) Posession Arrest Rate
IFEM205	Female Interaction Of V18HLN
IFEM206	Female Interaction Of V18HQU
IFEM207	Female Interaction Of V18HCU
IFEM208	Female Interaction Of V18HQR
IBLK205	Black Interaction Of V18HLN
IBLK206	Black Interaction Of V18HQU
IBLK207	Black Interaction Of V18HCU
IBLK208	Black Interaction Of V18HQR
IHISP205	Hispanic Interaction Of V18HLN
IHISP206	Hispanic Interaction Of V18HQU
IHISP207	Hispanic Interaction Of V18HCU
IHISP208	Hispanic Interaction Of V18HQR
IHISP213	Hispanic Interaction Of V18DQR
IOTH210	Other Interaction Of V18DLN
IOTH211	Other Interaction Of V18DQU
IOTH212	Other Interaction Of V18DCU
IOTH213	Other Interaction Of V18DQR
V18GLN	C: Linear: Synthetic Narcotics Posession Arrest Rate
V18GQU	C: Quadratic: Synthetic Narcotics Posession Arrest Rate
V18GCU	C: Cubic: Synthetic Narcotics Posession Arrest Rate
V18GQR	C: Quartic: Synthetic Narcotics Posession Arrest Rate
IFEM215	Female Interaction Of V18GLN
IFEM216	Female Interaction Of V18GQU
IFEM217	Female Interaction Of V18GCU
IFEM218	Female Interaction Of V18GQR
IBLK215	Black Interaction Of V18GLN
IBLK216	Black Interaction Of V18GQU
IBLK217	Black Interaction Of V18GCU
IBLK218	Black Interaction Of V18GQR
IHISP215	Hispanic Interaction Of V18GLN

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models
(continued)

Variable	Label
IHISP216	Hispanic Interaction Of V18GQU
IHISP217	Hispanic Interaction Of V18GCU
IHISP218	Hispanic Interaction Of V18GQR
IOTH215	Other Interaction Of V18GLN
IOTH216	Other Interaction Of V18GQU
IOTH217	Other Interaction Of V18GCU
IOTH218	Other Interaction Of V18GQR
V18CLN	C: Linear: Synthetic Narcotics Sale/Manuf Arrest Rate
V18CQU	C: Quadratic: Synthetic Narcotics Sale/Manuf Arrest Rate
V18CCU	C: Cubic: Synthetic Narcotics Sale/Manuf Arrest Rate
V18CQR	C: Quartic: Synthetic Narcotics Sale/Manuf Arrest Rate
IFEM220	Female Interaction Of V18CLN
IFEM221	Female Interaction Of V18CQU
IFEM222	Female Interaction Of V18CCU
IFEM223	Female Interaction Of V18CQR
IBLK220	Black Interaction Of V18CLN
IBLK221	Black Interaction Of V18CQU
IBLK222	Black Interaction Of V18CCU
IBLK223	Black Interaction Of V18CQR
IHISP220	Hispanic Interaction Of V18CLN
IHISP221	Hispanic Interaction Of V18CQU
IHISP222	Hispanic Interaction Of V18CCU
IHISP223	Hispanic Interaction Of V18CQR
IOTH220	Other Interaction Of V18CLN
IOTH221	Other Interaction Of V18CQU
IOTH222	Other Interaction Of V18CCU
IOTH223	Other Interaction Of V18CQR
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate
IHISP231	Hispanic Interaction Of VIOLQU
IHISP232	Hispanic Interaction Of VIOLCU
IHISP233	Hispanic Interaction Of VIOLQR
IOTH230	Other Interaction Of VIOLLN
IOTH231	Other Interaction Of VIOLQU
IOTH232	Other Interaction Of VIOLCU
IOTH233	Other Interaction Of VIOLQR
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 92
DRATEQU	C: Quad: Mean Drug Client Treatment Rate 1991 & 92
DRATECU	C: Cubic: Mean Drug Client Treatment Rate 1991 & 92
DRATEQR	C: Quart: Mean Drug Client Treatment Rate 1991 & 92
IFEM235	Female Interaction Of DRATELN
IFEM236	Female Interaction Of DRATEQU
IFEM237	Female Interaction Of DRATECU
IFEM238	Female Interaction Of DRATEQR
IBLK235	Black Interaction Of DRATELN
IBLK236	Black Interaction Of DRATEQU
IBLK237	Black Interaction Of DRATECU
IBLK238	Black Interaction Of DRATEQR
IHISP235	Hispanic Interaction Of DRATELN
IHISP236	Hispanic Interaction Of DRATEQU

Variable	Label
IFEM225	Female Interaction Of V18LN
IFEM226	Female Interaction Of V18QU
IFEM227	Female Interaction Of V18CU
IFEM228	Female Interaction Of V18QR
IBLK225	Black Interaction Of V18LN
IBLK226	Black Interaction Of V18QU
IBLK227	Black Interaction Of V18CU
IBLK228	Black Interaction Of V18QR
IHISP225	Hispanic Interaction Of V18LN
IHISP226	Hispanic Interaction Of V18QU
IHISP227	Hispanic Interaction Of V18CU
IHISP228	Hispanic Interaction Of V18QR
IOTH225	Other Interaction Of V18LN
IOTH226	Other Interaction Of V18QU
IOTH227	Other Interaction Of V18CU
IOTH228	Other Interaction Of V18QR
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate
VIOLCU	C: Cubic: Total Violent Offenses Arrest Rate
VIOLQR	C: Quartic: Total Violent Offenses Arrest Rate
IFEM230	Female Interaction Of VIOLLN
IFEM231	Female Interaction Of VIOLQU
IFEM232	Female Interaction Of VIOLCU
IFEM233	Female Interaction Of VIOLQR
IBLK230	Black Interaction Of VIOLLN
IBLK231	Black Interaction Of VIOLQU
IBLK232	Black Interaction Of VIOLCU
IBLK233	Black Interaction Of VIOLQR
IHISP230	Hispanic Interaction Of VIOLLN
IHISP237	Hispanic Interaction Of DRATECU
IHISP238	Hispanic Interaction Of DRATEQR
IOTH235	Other Interaction Of DRATELN
IOTH236	Other Interaction Of DRATEQU
IOTH237	Other Interaction Of DRATECU
IOTH238	Other Interaction Of DRATEQR
RH43ALN	T: Linear: Recoded Median Rents For Rental Units
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units
IFEM240	Female Interaction Of RH43ALN
IFEM241	Female Interaction Of RH43AQU
IFEM242	Female Interaction Of RH43ACU
IFEM243	Female Interaction Of RH43AQR
IBLK240	Black Interaction Of RH43ALN
IBLK241	Black Interaction Of RH43AQU
IBLK242	Black Interaction Of RH43ACU
IBLK243	Black Interaction Of RH43AQR
IHISP240	Hispanic Interaction Of RH43ALN
IHISP241	Hispanic Interaction Of RH43AQU
IHISP242	Hispanic Interaction Of RH43ACU

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models
(continued)

Variable	Label
IHISP243	Hispanic Interaction Of RH43AQR
IOTH240	Other Interaction Of RH43ALN
IOTH241	Other Interaction Of RH43AQU
IOTH242	Other Interaction Of RH43ACU
IOTH243	Other Interaction Of RH43AQR
RH61ALN	T: Linear: Recod Median Value Of Owner Occup HUs
RH61AQU	T: Quad: Recod Median Value Of Owner Occup HUs
RH61ACU	T: Cubic: Recod Median Value Of Owner Occup HUs
RH61AQR	T: Quartic: Recod Median Value Of Owner Occup HUs
IFEM245	Female Interaction Of RH61ALN
IFEM246	Female Interaction Of RH61AQU
IFEM247	Female Interaction Of RH61ACU
IFEM248	Female Interaction Of RH61AQR
IBLK245	Black Interaction Of RH61ALN
IBLK246	Black Interaction Of RH61AQU
IBLK247	Black Interaction Of RH61ACU
IBLK248	Black Interaction Of RH61AQR
IHISP245	Hispanic Interaction Of RH61ALN
IHISP246	Hispanic Interaction Of RH61AQU
IHISP247	Hispanic Interaction Of RH61ACU
IHISP248	Hispanic Interaction Of RH61AQR
IOTH245	Other Interaction Of RH61ALN
IOTH246	Other Interaction Of RH61AQU
IOTH247	Other Interaction Of RH61ACU
IOTH248	Other Interaction Of RH61AQR
RP80ALN	T: Linear: Recoded Median Household Income
RP80AQU	T: Quadratic: Recoded Median Household Income
RP80ACU	T: Cubic: Recoded Median Household Income
IBLK258	Black Interaction Of ARATEQR
IHISP255	Hispanic Interaction Of ARATELN
IHISP256	Hispanic Interaction Of ARATEQU
IHISP257	Hispanic Interaction Of ARATECU
IHISP258	Hispanic Interaction Of ARATEQR
IOTH255	Other Interaction Of ARATELN
IOTH256	Other Interaction Of ARATEQU
IOTH257	Other Interaction Of ARATECU
IOTH258	Other Interaction Of ARATEQR
B64DISLN	B: Linear: % Persons 16-64 With A Work Disability
B64DISQU	B: Quadratic: % Persons 16-64 With A Work Disability
B64DISCU	B: Cubic: % Persons 16-64 With A Work Disability
B64DISQR	B: Quartic: % Persons 16-64 With A Work Disability
IFEM260	Female Interaction Of B64DISLN
IFEM261	Female Interaction Of B64DISQU
IFEM262	Female Interaction Of B64DISCU
IFEM263	Female Interaction Of B64DISQR
IBLK260	Black Interaction Of B64DISLN
IBLK261	Black Interaction Of B64DISQU
IBLK262	Black Interaction Of B64DISCU
IBLK263	Black Interaction Of B64DISQR
IHISP260	Hispanic Interaction Of B64DISLN
IHISP261	Hispanic Interaction Of B64DISQU

Variable	Label
RP80AQR	T: Quartic: Recoded Median Household Income
IFEM250	Female Interaction Of RP80ALN
IFEM251	Female Interaction Of RP80AQU
IFEM252	Female Interaction Of RP80ACU
IFEM253	Female Interaction Of RP80AQR
IBLK250	Black Interaction Of RP80ALN
IBLK251	Black Interaction Of RP80AQU
IBLK252	Black Interaction Of RP80ACU
IBLK253	Black Interaction Of RP80AQR
IHISP250	Hispanic Interaction Of RP80ALN
IHISP251	Hispanic Interaction Of RP80AQU
IHISP252	Hispanic Interaction Of RP80ACU
IHISP253	Hispanic Interaction Of RP80AQR
IOTH250	Other Interaction Of RP80ALN
IOTH251	Other Interaction Of RP80AQU
IOTH252	Other Interaction Of RP80ACU
IOTH253	Other Interaction Of RP80AQR
ARATELN	C: Linear: Mean A-Only Client Trmt Rate 1991 & 1992
ARATEQU	C: Quad: Mean A-Only Client Trmt Rate 1991 & 1992
ARATECU	C: Cubic: Mean A-Only Client Trmt Rate 1991 & 1992
ARATEQR	C: Quart: Mean A-Only Client Trmt Rate 1991 & 1992
IFEM255	Female Interaction Of ARATELN
IFEM256	Female Interaction Of ARATEQU
IFEM257	Female Interaction Of ARATECU
IFEM258	Female Interaction Of ARATEQR
IBLK255	Black Interaction Of ARATELN
IBLK256	Black Interaction Of ARATEQU
IBLK257	Black Interaction Of ARATECU
IHISP262	Hispanic Interaction Of B64DISCU
IHISP263	Hispanic Interaction Of B64DISQR
IOTH260	Other Interaction Of B64DISLN
IOTH261	Other Interaction Of B64DISQU
IOTH262	Other Interaction Of B64DISCU
IOTH263	Other Interaction Of B64DISQR
BAGE34LN	B: Linear: Percent Persons 25-34 Years
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years
BAGE34CU	B: Cubic: Percent Persons 25-34 Years
BAGE34QR	B: Quartic: Percent Persons 25-34 Years
IFEM265	Female Interaction Of BAGE34LN
IFEM266	Female Interaction Of BAGE34QU
IFEM267	Female Interaction Of BAGE34CU
IFEM268	Female Interaction Of BAGE34QR
IBLK265	Black Interaction Of BAGE34LN
IBLK266	Black Interaction Of BAGE34QU
IBLK267	Black Interaction Of BAGE34CU
IBLK268	Black Interaction Of BAGE34QR
IHISP265	Hispanic Interaction Of BAGE34LN
IHISP266	Hispanic Interaction Of BAGE34QU
IHISP267	Hispanic Interaction Of BAGE34CU
IHISP268	Hispanic Interaction Of BAGE34QR
IOTH265	Other Interaction Of BAGE34LN

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models
(continued)

Variable	Label
IOTH266	Other Interaction Of BAGE34QU
IOTH267	Other Interaction Of BAGE34CU
IOTH268	Other Interaction Of BAGE34QR
BAGE44LN	B: Linear: Percent Persons 35-44 Years
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years
BAGE44CU	B: Cubic: Percent Persons 35-44 Years
BAGE44QR	B: Quartic: Percent Persons 35-44 Years
IFEM270	Female Interaction Of BAGE44LN
IFEM271	Female Interaction Of BAGE44QU
IFEM272	Female Interaction Of BAGE44CU
IFEM273	Female Interaction Of BAGE44QR
IBLK270	Black Interaction Of BAGE44LN
IBLK271	Black Interaction Of BAGE44QU
IBLK272	Black Interaction Of BAGE44CU
IBLK273	Black Interaction Of BAGE44QR
IHISP270	Hispanic Interaction Of BAGE44LN
IHISP271	Hispanic Interaction Of BAGE44QU
IHISP272	Hispanic Interaction Of BAGE44CU
IHISP273	Hispanic Interaction Of BAGE44QR
IOTH270	Other Interaction Of BAGE44LN
IOTH271	Other Interaction Of BAGE44QU
IOTH272	Other Interaction Of BAGE44CU
IOTH273	Other Interaction Of BAGE44QR
BAGE54LN	B: Linear: Percent Persons 45-54 Years
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years
BAGE54CU	B: Cubic: Percent Persons 45-54 Years
BAGE54QR	B: Quartic: Percent Persons 45-54 Years
IBLK281	Black Interaction Of BAGE64QU
IBLK282	Black Interaction Of BAGE64CU
IBLK283	Black Interaction Of BAGE64QR
IHISP280	Hispanic Interaction Of BAGE64LN
IHISP281	Hispanic Interaction Of BAGE64QU
IHISP282	Hispanic Interaction Of BAGE64CU
IHISP283	Hispanic Interaction Of BAGE64QR
IOTH280	Other Interaction Of BAGE64LN
IOTH281	Other Interaction Of BAGE64QU
IOTH282	Other Interaction Of BAGE64CU
IOTH283	Other Interaction Of BAGE64QR
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander
IFEM285	Female Interaction Of BASIANLN
IFEM286	Female Interaction Of BASIANQU
IFEM287	Female Interaction Of BASIANCU
IFEM288	Female Interaction Of BASIANQR
IBLK285	Black Interaction Of BASIANLN
IBLK286	Black Interaction Of BASIANQU
IBLK287	Black Interaction Of BASIANCU
IBLK288	Black Interaction Of BASIANQR
IHISP285	Hispanic Interaction Of BASIANLN
IHISP286	Hispanic Interaction Of BASIANQU
IHISP287	Hispanic Interaction Of BASIANCU

Variable	Label
IFEM275	Female Interaction Of BAGE54LN
IFEM276	Female Interaction Of BAGE54QU
IFEM277	Female Interaction Of BAGE54CU
IFEM278	Female Interaction Of BAGE54QR
IBLK275	Black Interaction Of BAGE54LN
IBLK276	Black Interaction Of BAGE54QU
IBLK277	Black Interaction Of BAGE54CU
IBLK278	Black Interaction Of BAGE54QR
IHISP275	Hispanic Interaction Of BAGE54LN
IHISP276	Hispanic Interaction Of BAGE54QU
IHISP277	Hispanic Interaction Of BAGE54CU
IHISP278	Hispanic Interaction Of BAGE54QR
IOTH275	Other Interaction Of BAGE54LN
IOTH276	Other Interaction Of BAGE54QU
IOTH277	Other Interaction Of BAGE54CU
IOTH278	Other Interaction Of BAGE54QR
BAGE64LN	B: Linear: Percent Persons 55-64 Years
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years
BAGE64CU	B: Cubic: Percent Persons 55-64 Years
BAGE64QR	B: Quartic: Percent Persons 55-64 Years
IFEM280	Female Interaction Of BAGE64LN
IFEM281	Female Interaction Of BAGE64QU
IFEM282	Female Interaction Of BAGE64CU
IFEM283	Female Interaction Of BAGE64QR
IBLK280	Black Interaction Of BAGE64LN
IHISP288	Hispanic Interaction Of BASIANQR
IOTH285	Other Interaction Of BASIANLN
IOTH286	Other Interaction Of BASIANQU
IOTH287	Other Interaction Of BASIANCU
IOTH288	Other Interaction Of BASIANQR
BCUBANLN	B: Linear: Percent Hispanics: Cuban
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban
BCUBANCU	B: Cubic: Percent Hispanics: Cuban
BCUBANQR	B: Quartic: Percent Hispanics: Cuban
IFEM290	Female Interaction Of BCUBANLN
IFEM291	Female Interaction Of BCUBANQU
IFEM292	Female Interaction Of BCUBANCU
IFEM293	Female Interaction Of BCUBANQR
IBLK290	Black Interaction Of BCUBANLN
IBLK291	Black Interaction Of BCUBANQU
IBLK292	Black Interaction Of BCUBANCU
IBLK293	Black Interaction Of BCUBANQR
IHISP290	Hispanic Interaction Of BCUBANLN
IHISP291	Hispanic Interaction Of BCUBANQU
IHISP292	Hispanic Interaction Of BCUBANCU
IHISP293	Hispanic Interaction Of BCUBANQR
IOTH290	Other Interaction Of BCUBANLN
IOTH291	Other Interaction Of BCUBANQU
IOTH292	Other Interaction Of BCUBANCU
IOTH293	Other Interaction Of BCUBANQR

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models

(continued)

Variable	Label
BFNOTLN	B: Linear: % Fem Separated, Divorced or Widowed
BFNOTQU	B: Quadratic: % Fem Separated, Divorced or Widowed
BFNOTCU	B: Cubic: % Fem Separated, Divorced or Widowed
BFNOTQR	B: Quartic: % Fem Separated, Divorced or Widowed
IFEM295	Female Interaction Of BFNOTLN
IFEM296	Female Interaction Of BFNOTQU
IFEM297	Female Interaction Of BFNOTCU
IFEM298	Female Interaction Of BFNOTQR
IBLK295	Black Interaction Of BFNOTLN
IBLK296	Black Interaction Of BFNOTQU
IBLK297	Black Interaction Of BFNOTCU
IBLK298	Black Interaction Of BFNOTQR
IHISP295	Hispanic Interaction Of BFNOTLN
IHISP296	Hispanic Interaction Of BFNOTQU
IHISP297	Hispanic Interaction Of BFNOTCU
IHISP298	Hispanic Interaction Of BFNOTQR
IOTH295	Other Interaction Of BFNOTLN
IOTH296	Other Interaction Of BFNOTQU
IOTH297	Other Interaction Of BFNOTCU
IOTH298	Other Interaction Of BFNOTQR
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut
BINDIAQU	B: Quad: Percent Pop: American Indian, Eskimo, Aleut
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut
BINDIAQR	B: Quart: Percent Pop: American Indian, Eskimo, Aleut
IFEM308	Female Interaction Of BMNOTQR
IBLK305	Black Interaction Of BMNOTLN
IBLK306	Black Interaction Of BMNOTQU
IBLK307	Black Interaction Of BMNOTCU
IBLK308	Black Interaction Of BMNOTQR
IHISP305	Hispanic Interaction Of BMNOTLN
IHISP306	Hispanic Interaction Of BMNOTQU
IHISP307	Hispanic Interaction Of BMNOTCU
IHISP308	Hispanic Interaction Of BMNOTQR
IOTH305	Other Interaction Of BMNOTLN
IOTH306	Other Interaction Of BMNOTQU
IOTH307	Other Interaction Of BMNOTCU
IOTH308	Other Interaction Of BMNOTQR
BPOVERLN	B: Linear: Percent Families Below Poverty Level
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level
BPOVERCU	B: Cubic: Percent Families Below Poverty Level
BPOVERQR	B: Quartic: Percent Families Below Poverty Level
IFEM310	Female Interaction Of BPOVERLN
IFEM311	Female Interaction Of BPOVERQU
IFEM312	Female Interaction Of BPOVERCU
IFEM313	Female Interaction Of BPOVERQR
IBLK310	Black Interaction Of BPOVERLN
IBLK311	Black Interaction Of BPOVERQU
IBLK312	Black Interaction Of BPOVERCU
IBLK313	Black Interaction Of BPOVERQR
IHISP310	Hispanic Interaction Of BPOVERLN
IHISP311	Hispanic Interaction Of BPOVERQU
IHISP312	Hispanic Interaction Of BPOVERCU

Variable	Label
IFEM300	Female Interaction Of BINDIALN
IFEM301	Female Interaction Of BINDIAQU
IFEM302	Female Interaction Of BINDIACU
IFEM303	Female Interaction Of BINDIAQR
IBLK300	Black Interaction Of BINDIALN
IBLK301	Black Interaction Of BINDIAQU
IBLK302	Black Interaction Of BINDIACU
IBLK303	Black Interaction Of BINDIAQR
IHISP300	Hispanic Interaction Of BINDIALN
IHISP301	Hispanic Interaction Of BINDIAQU
IHISP302	Hispanic Interaction Of BINDIACU
IHISP303	Hispanic Interaction Of BINDIAQR
IOTH300	Other Interaction Of BINDIALN
IOTH301	Other Interaction Of BINDIAQU
IOTH302	Other Interaction Of BINDIACU
IOTH303	Other Interaction Of BINDIAQR
BMNOTLN	B: Linear: % Males Separated, Divorced or Widowed
BMNOTQU	B: Quad: % Males Separated, Divorced or Widowed
BMNOTCU	B: Cubic: % Males Separated, Divorced or Widowed
BMNOTQR	B: Quartic: % Males Separated, Divorced or Widowed
IFEM305	Female Interaction Of BMNOTLN
IFEM306	Female Interaction Of BMNOTQU
IFEM307	Female Interaction Of BMNOTCU
IHISP313	Hispanic Interaction Of BPOVERQR
IOTH310	Other Interaction Of BPOVERLN
IOTH311	Other Interaction Of BPOVERQU
IOTH312	Other Interaction Of BPOVERCU
IOTH313	Other Interaction Of BPOVERQR
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican
IFEM315	Female Interaction Of BPRICALN
IFEM316	Female Interaction Of BPRICAQU
IFEM317	Female Interaction Of BPRICACU
IFEM318	Female Interaction Of BPRICAQR
IBLK315	Black Interaction Of BPRICALN
IBLK316	Black Interaction Of BPRICAQU
IBLK317	Black Interaction Of BPRICACU
IBLK318	Black Interaction Of BPRICAQR
IHISP315	Hispanic Interaction Of BPRICALN
IHISP316	Hispanic Interaction Of BPRICAQU
IHISP317	Hispanic Interaction Of BPRICACU
IHISP318	Hispanic Interaction Of BPRICAQR
IOTH315	Other Interaction Of BPRICALN
IOTH316	Other Interaction Of BPRICAQU
IOTH317	Other Interaction Of BPRICACU
IOTH318	Other Interaction Of BPRICAQR
BSCHASLN	B: Linear: Percent Associates Degree
BSCHASQU	B: Quadratic: Percent Associates Degree
BSCHASCU	B: Cubic: Percent Associates Degree
BSCHASQR	B: Quartic: Percent Associates Degree

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models

(continued)

Variable	Label
IFEM320	Female Interaction Of BSCHASLN
IFEM321	Female Interaction Of BSCHASQU
IFEM322	Female Interaction Of BSCHASCU
IFEM323	Female Interaction Of BSCHASQR
IBLK320	Black Interaction Of BSCHASLN
IBLK321	Black Interaction Of BSCHASQU
IBLK322	Black Interaction Of BSCHASCU
IBLK323	Black Interaction Of BSCHASQR
IHISP320	Hispanic Interaction Of BSCHASLN
IHISP321	Hispanic Interaction Of BSCHASQU
IHISP322	Hispanic Interaction Of BSCHASCU
IHISP323	Hispanic Interaction Of BSCHASQR
IOTH320	Other Interaction Of BSCHASLN
IOTH321	Other Interaction Of BSCHASQU
IOTH322	Other Interaction Of BSCHASCU
IOTH323	Other Interaction Of BSCHASQR
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander
IFEM331	Female Interaction Of PCUBANQU
IFEM332	Female Interaction Of PCUBANCU
IFEM333	Female Interaction Of PCUBANQR
IBLK330	Black Interaction Of PCUBANLN
IBLK331	Black Interaction Of PCUBANQU
IBLK332	Black Interaction Of PCUBANCU
IBLK333	Black Interaction Of PCUBANQR
IHISP330	Hispanic Interaction Of PCUBANLN
IHISP331	Hispanic Interaction Of PCUBANQU
IHISP332	Hispanic Interaction Of PCUBANCU
IHISP333	Hispanic Interaction Of PCUBANQR
IOTH330	Other Interaction Of PCUBANLN
IOTH331	Other Interaction Of PCUBANQU
IOTH332	Other Interaction Of PCUBANCU
IOTH333	Other Interaction Of PCUBANQR
PINDIALN	T: Linear: % Pop: American Indian, Eskimo, Aleut
PINDIAQU	T: Quadratic: % Pop: American Indian, Eskimo, Aleut
PINDIACU	T: Cubic: % Pop: American Indian, Eskimo, Aleut
PINDIAQR	T: Quartic: % Pop: American Indian, Eskimo, Aleut
IFEM335	Female Interaction Of PINDIALN
IFEM336	Female Interaction Of PINDIAQU
IFEM337	Female Interaction Of PINDIACU
IFEM338	Female Interaction Of PINDIAQR
IBLK335	Black Interaction Of PINDIALN
IBLK336	Black Interaction Of PINDIAQU
IBLK337	Black Interaction Of PINDIACU
IBLK338	Black Interaction Of PINDIAQR
IHISP335	Hispanic Interaction Of PINDIALN
IHISP336	Hispanic Interaction Of PINDIAQU
IHISP337	Hispanic Interaction Of PINDIACU
IHISP338	Hispanic Interaction Of PINDIAQR

Variable	Label
IFEM325	Female Interaction Of PASIANLN
IFEM326	Female Interaction Of PASIANQU
IFEM327	Female Interaction Of PASIANCU
IFEM328	Female Interaction Of PASIANQR
IBLK325	Black Interaction Of PASIANLN
IBLK326	Black Interaction Of PASIANQU
IBLK327	Black Interaction Of PASIANCU
IBLK328	Black Interaction Of PASIANQR
IHISP325	Hispanic Interaction Of PASIANLN
IHISP326	Hispanic Interaction Of PASIANQU
IHISP327	Hispanic Interaction Of PASIANCU
IHISP328	Hispanic Interaction Of PASIANQR
IOTH325	Other Interaction Of PASIANLN
IOTH326	Other Interaction Of PASIANQU
IOTH327	Other Interaction Of PASIANCU
IOTH328	Other Interaction Of PASIANQR
PCUBANLN	T: Linear: Percent Hispanics: Cuban
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban
PCUBANCU	T: Cubic: Percent Hispanics: Cuban
PCUBANQR	T: Quartic: Percent Hispanics: Cuban
IFEM330	Female Interaction Of PCUBANLN
IOTH335	Other Interaction Of PINDIALN
IOTH336	Other Interaction Of PINDIAQU
IOTH337	Other Interaction Of PINDIACU
IOTH338	Other Interaction Of PINDIAQR
HSDROPLN	T: Linear: Underclass Indicator: High School Dropouts
HSDROPQU	T: Quadratic: Underclass Indicator: HS Dropouts
HSDROPCU	T: Cubic: Underclass Indicator: HS Dropouts
HSDROPQR	T: Quartic: Underclass Indicator: HS Dropouts
IFEM340	Female Interaction Of HSDROPLN
IFEM341	Female Interaction Of HSDROPQU
IFEM342	Female Interaction Of HSDROPCU
IFEM343	Female Interaction Of HSDROPQR
IBLK340	Black Interaction Of HSDROPLN
IBLK341	Black Interaction Of HSDROPQU
IBLK342	Black Interaction Of HSDROPCU
IBLK343	Black Interaction Of HSDROPQR
IHISP340	Hispanic Interaction Of HSDROPLN
IHISP341	Hispanic Interaction Of HSDROPQU
IHISP342	Hispanic Interaction Of HSDROPCU
IHISP343	Hispanic Interaction Of HSDROPQR
IOTH340	Other Interaction Of HSDROPLN
IOTH341	Other Interaction Of HSDROPQU
IOTH342	Other Interaction Of HSDROPCU
IOTH343	Other Interaction Of HSDROPQR
HSDROP9	T: Underclass Indicator: High School Dropouts
IFEM339	Female Interaction Of HSDROP9
IBLK339	Black Interaction Of HSDROP9

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Table D.1 Block Group and Tract Level Variables Considered in Small Area Estimation Models

(continued)

Variable	Label	Variable	Label
IHISP339	Hispanic Interaction Of HSDROP9		
IOTH339	Other Interaction Of HSDROP9		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Appendix E: Fixed Effect Coefficients for Logistic Models

E1. Coefficients of Model Parameters for Past Month Alcohol Use

		BIG CITY				REMAINDER			
Variable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	-1.371	0.862	1.003	0.392	-1.258	0.970	0.991	0.913
FEMALE	O: Female Indicator	-0.070	-0.277	-0.448	-0.570	-0.203	-0.481	-0.704	-0.797
FEMBLCK	O: Black Interaction Of FEMALE	-0.172	-0.105	-0.442	-0.340	-0.156	-0.138	0.220	-0.200
FEMHISP	O: Hispanic Interaction Of FEMALE	-0.309	-0.564	-0.823	-0.622	-0.005	-0.585	-0.511	-0.461
FEMOTHR	O: Other Interaction Of FEMALE	-0.167	-0.080	-0.883	-0.609	0.742	0.010	0.128	0.576
RACEBLCK	O: Race/Black Indicator	-0.731	-0.589	-0.455	-0.064	-0.303	-0.407	-0.471	-0.343
RACEHISP	O: Race/Hispanic Indicator	0.005	-0.208	-0.013	-0.005	-0.000	-0.105	-0.077	0.143
RACEOTHR	O: Race/Other Indicator	-0.631	-1.065	-0.912	-0.939	-0.967	-2.323	-0.896	-3.210
REGNOREA	O: Northeast Region Indicator						0.143		-0.173
REGSOUTH	O: South Region Indicator						-0.262		-0.451
REGWEST	O: West Region Indicator						0.097		-0.138
IOTH7	Other Interaction Of REGNOREA						1.864		2.478
IOTH8	Other Interaction Of REGSOUTH						1.558		3.241
IOTH9	Other Interaction Of REGWEST						1.154		1.051
UCLASS9	T: Underclass Indicator								0.864
PAGE18LN	T: Linear: Percent Persons 0-18 Years				-0.074				-0.014
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years								-0.038
IOTH25	Other Interaction Of PAGE18LN								0.375
IOTH26	Other Interaction Of PAGE18QU								-0.503
PAGE24LN	T: Linear: Percent Persons 19-24 Years						-0.037		
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years						-0.087		
IFEM30	Female Interaction Of PAGE24LN						-0.009		
IFEM31	Female Interaction Of PAGE24QU						0.141		
PAGE34LN	T: Linear: Percent Persons 25-34 Years	-0.076							
PAGE44LN	T: Linear: Percent Persons 35-44 Years				0.108				0.044
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years								0.013
PAGE44CU	T: Cubic: Percent Persons 35-44 Years								0.004
PAGE44QR	T: Quartic: Percent Persons 35-44 Years								0.023
PAGE64LN	T: Linear: Percent Persons 55-64 Years	0.043							
PSCH8LN	T: Linear: Percent 0-8 Years Of School		-0.097						
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree		0.155				0.165	0.153	
IBLK75	Black Interaction Of PSCHCOLN						-0.226	-0.170	
PSCHSCLN	T: Linear: Percent Some College And No Degree		-0.060						
PPOVERLN	T: Linear: Percent Families Below Poverty Level					-0.076			
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level					-0.030			
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					-0.116			
PPOVERQR	T: Quartic: Percent Families Below Poverty Level					-0.023			
IBLK85	Black Interaction Of PPOVERLN					0.235			
IBLK86	Black Interaction Of PPOVERQU					0.048			
IBLK87	Black Interaction Of PPOVERCU					0.078			
IBLK88	Black Interaction Of PPOVERQR					0.063			
PBLACKLN	T: Linear: Percent Black Nonhispanic		-0.132						
POTHLN	T: Linear: Percent Other Race/Hispanicity						0.034		
POTHQU	T: Quadratic: Percent Other Race/Hispanicity						-0.052		
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force	0.039							
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force	0.049							
PFNEVLN	T: Linear: Percent Females Never Married						0.091		
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed				-0.037				
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed				-0.002				
PFNOTCU	T: Cubic: Percent Females Separated, Divorced Or Widowed				-0.060				
PFNOTQR	T: Quartic: Percent Females Separated, Divorced Or Widowed				0.028				
PMNEVLN	T: Linear: Percent Males Never Married			0.067		0.103		0.046	0.069
PMNEVQU	T: Quadratic: Percent Males Never Married							-0.021	-0.044
PMNEVCU	T: Cubic: Percent Males Never Married							0.031	0.003
PMNEVQR	T: Quartic: Percent Males Never Married								0.027
IFEM140	Female Interaction Of PMNEVLN								0.038
IFEM141	Female Interaction Of PMNEVQU								0.162
IBLK140	Black Interaction Of PMNEVLN					-0.262			
IHISP140	Hispanic Interaction Of PMNEVLN			-0.152					
IOTH140	Other Interaction Of PMNEVLN			0.289				0.079	
IOTH141	Other Interaction Of PMNEVQU							0.039	
IOTH142	Other Interaction Of PMNEVCU							-0.425	
P40HULN	T: Linear: Percent Housing Units Built 1940-1949					0.124			
IFEM160	Female Interaction Of P40HULN					-0.161			
PRENTLN	T: Linear: Percent Housing Rented			-0.026		-0.008			
IBLK165	Black Interaction Of PRENTLN			0.059					
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases		0.002						
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases		-0.077						
V18FLN	C: Linear: Marijuana Posession Arrest Rate								0.068

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IOTH185	Other Interaction Of V18FLN								-0.514
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate		-0.003		-0.027				-0.082
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate		-0.044		-0.041				-0.026
V18ACU	C: Cubic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate				-0.034				-0.014
V18AQR	C: Quartic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate								-0.038
IHISP200	Hispanic Interaction Of V18ALN		-0.164						
IHISP201	Hispanic Interaction Of V18AQU		0.117						
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate								0.174
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992					0.051			
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992					-0.013			
DRATECU	C: Cubic: Mean Drug Client Treatment Rate 1991 & 1992					-0.024			
IOTH235	Other Interaction Of DRATELN					-0.020			
IOTH236	Other Interaction Of DRATEQU					0.070			
IOTH237	Other Interaction Of DRATECU					-0.436			
RH61ALN	T: Linear: Recoded Median Value Of Owner Occupied HUs		-0.003						0.227
RH61AQU	T: Quadratic: Recoded Median Value Of Owner Occupied HUs		-0.006						
IHISP245	Hispanic Interaction Of RH61ALN		-0.099						
IHISP246	Hispanic Interaction Of RH61AQU		0.103						
RP80ALN	T: Linear: Recoded Median Household Income				0.035				
RP80AQU	T: Quadratic: Recoded Median Household Income				0.051				
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992		0.015						
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992		0.013						
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992		-0.013						
IFEM255	Female Interaction Of ARATELN		-0.073						
IFEM256	Female Interaction Of ARATEQU		0.035						
IFEM257	Female Interaction Of ARATECU		-0.097						
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability				-0.009		-0.017		-0.150
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability				0.040		-0.007		
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability						0.054		
IHISP260	Hispanic Interaction Of B64DISLN						-0.046		0.202
IHISP261	Hispanic Interaction Of B64DISQU						0.116		
IHISP262	Hispanic Interaction Of B64DISCU						-0.160		
BAGE34LN	B: Linear: Percent Persons 25-34 Years							0.027	
IOTH265	Other Interaction Of BAGE34LN							-0.437	
BAGE44LN	B: Linear: Percent Persons 35-44 Years								-0.022
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years								-0.005
IHISP270	Hispanic Interaction Of BAGE44LN								0.130
IHISP271	Hispanic Interaction Of BAGE44QU								0.182
BAGE54LN	B: Linear: Percent Persons 45-54 Years				0.057				0.038
IBLK275	Black Interaction Of BAGE54LN								-0.164
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander							0.002	-0.040
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander							-0.007	-0.004
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander							-0.055	0.005
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander							-0.003	
IOTH285	Other Interaction Of BASIANLN							-0.221	0.642
IOTH286	Other Interaction Of BASIANQU							0.002	-0.761
IOTH287	Other Interaction Of BASIANCU							0.242	0.757
IOTH288	Other Interaction Of BASIANQR							0.199	
BCUBANLN	B: Linear: Percent Hispanics: Cuban		-0.032			0.091			
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban					0.042			
BCUBANCU	B: Cubic: Percent Hispanics: Cuban					0.057			
IHISP290	Hispanic Interaction Of BCUBANLN		0.185						
IOTH290	Other Interaction Of BCUBANLN					0.294			
IOTH291	Other Interaction Of BCUBANQU					0.055			
IOTH292	Other Interaction Of BCUBANCU					-1.073			
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed		0.034				-0.030		
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed		0.001				-0.034		
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed		0.024						
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed		-0.010						
IOTH295	Other Interaction Of BFNOTLN		0.103				0.093		
IOTH296	Other Interaction Of BFNOTQU		-0.001				0.263		
IOTH297	Other Interaction Of BFNOTCU		-0.145						
IOTH298	Other Interaction Of BFNOTQR		0.062						
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut					-0.109	-0.030		
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut						-0.017		
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut						0.171		
IFEM300	Female Interaction Of BINDIALN					0.225	-0.012		
IFEM301	Female Interaction Of BINDIAQU						-0.008		
IFEM302	Female Interaction Of BINDIACU						-0.185		
IBLK300	Black Interaction Of BINDIALN					0.291			
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed	0.045							-0.030
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed								0.002
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed								-0.042
IHISP305	Hispanic Interaction Of BMNOTLN	-0.127							0.086
IHISP306	Hispanic Interaction Of BMNOTQU								0.012
IHISP307	Hispanic Interaction Of BMNOTCU								0.171
BPOVERLN	B: Linear: Percent Families Below Poverty Level			-0.084					
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	-0.052				-0.016	-0.027		
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican	0.010				0.000	-0.083		
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican	0.075				-0.031	0.017		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican						0.072		
IBLK315	Black Interaction Of BPRICALN						-0.124		
IBLK316	Black Interaction Of BPRICAQU						0.260		
BSCHASLN	B: Linear: Percent Associates Degree				0.053				0.059
BSCHASQU	B: Quadratic: Percent Associates Degree				-0.026				
BSCHASCU	B: Cubic: Percent Associates Degree				0.013				
IBLK320	Black Interaction Of BSCHASLN				-0.072				
IBLK321	Black Interaction Of BSCHASQU				0.083				
IBLK322	Black Interaction Of BSCHASCU				-0.113				
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander			0.035					
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander			-0.062					
IBLK325	Black Interaction Of PASIANLN			0.055					
IBLK326	Black Interaction Of PASIANQU			0.090					
PCUBANLN	T: Linear: Percent Hispanics: Cuban				-0.025				
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban				-0.007				
PCUBANCU	T: Cubic: Percent Hispanics: Cuban				-0.109				
IFEM330	Female Interaction Of PCUBANLN				0.013				
IFEM331	Female Interaction Of PCUBANQU				0.019				
IFEM332	Female Interaction Of PCUBANCU				0.129				
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut			-0.040					-0.005
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			0.028					-0.086
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut			-0.011					0.051
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut			0.015					-0.039
IFEM335	Female Interaction Of PINDIALN								-0.026
IFEM336	Female Interaction Of PINDIAQU								0.035
IFEM337	Female Interaction Of PINDIACU								-0.099
IFEM338	Female Interaction Of PINDIAQR								0.059

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E2. Coefficients of Model Parameters for Past Month Any Illicit Drug Use

		BIG CITY				REMAINDER			
Variable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	-2.903	-1.573	-1.738	-3.271	-2.807	-1.406	-2.537	-3.825
FEMALE	O: Female Indicator	0.049	-0.420	-0.769	-0.581	0.088	-0.718	-0.353	-0.424
FEMBLCK	O: Black Interaction Of FEMALE	-0.534	-0.327	0.257	0.062	-0.050	0.203	-0.148	-0.277
FEMHISP	O: Hispanic Interaction Of FEMALE	-0.274	-0.085	0.330	-0.180		0.282	-0.258	0.116
FEMOTHR	O: Other Interaction Of FEMALE	0.459	-0.055	-0.549			-0.165	1.096	0.600
RACEBLCK	O: Race/Black Indicator	0.074	-0.103	-0.354	0.192	-0.357	-1.613	0.016	0.437
RACEHISP	O: Race/Hispanic Indicator	0.255	-0.518	-0.878	0.101	0.144	-0.581	0.614	-0.511
RACEOTHR	O: Race/Other Indicator	-1.180	-1.001	-1.000	-1.055		-0.541	-0.951	-1.295
REGNOREA	O: Northeast Region Indicator						0.336	0.411	
REGSOUTH	O: South Region Indicator						0.217	0.173	
REGWEST	O: West Region Indicator						0.662	0.813	
PDENLEV1	O: Large MSA					0.073	-0.476	-0.077	
PDENLEV2	O: Medium MSA					0.201	-0.321	-0.004	
PDENLEV3	O: Small MSA					0.297	-0.428	0.024	
PDENLEV4	O: NonMSA, Urban					0.121	-0.347	-0.054	
IBLK10	Black Interaction Of PDENLEV1						1.150		
IBLK11	Black Interaction Of PDENLEV2						0.770		
IBLK12	Black Interaction Of PDENLEV3						1.469		
IBLK13	Black Interaction Of PDENLEV4						1.484		
IHISP10	Hispanic Interaction Of PDENLEV1							-1.284	
IHISP11	Hispanic Interaction Of PDENLEV2							-1.705	
IHISP12	Hispanic Interaction Of PDENLEV3							-1.047	
IHISP13	Hispanic Interaction Of PDENLEV4							-0.962	
UCLASS9	T: Underclass Indicator						0.774		
PHH1PLN	T: Linear: Percent One Person Households						0.062		
PHH1PQU	T: Quadratic: Percent One Person Households						0.061		
POPRMLN	T: Linear: Average Persons Per Room			-0.012				-0.196	
POPRMQU	T: Quadratic: Average Persons Per Room							-0.138	
IFEM20	Female Interaction Of POPRMLN							0.157	
IFEM21	Female Interaction Of POPRMQU							0.137	
PAGE18LN	T: Linear: Percent Persons 0-18 Years						0.077		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years						0.020		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years						0.108		
IFEM25	Female Interaction Of PAGE18LN						-0.147		
IFEM26	Female Interaction Of PAGE18QU						-0.001		
IFEM27	Female Interaction Of PAGE18CU						-0.146		
PAGE24LN	T: Linear: Percent Persons 19-24 Years							0.136	
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years							-0.032	
PAGE24CU	T: Cubic: Percent Persons 19-24 Years							-0.089	
PAGE24QR	T: Quartic: Percent Persons 19-24 Years							0.005	
IHISP30	Hispanic Interaction Of PAGE24LN							0.047	
IHISP31	Hispanic Interaction Of PAGE24QU							0.039	
IHISP32	Hispanic Interaction Of PAGE24CU							0.023	
IHISP33	Hispanic Interaction Of PAGE24QR							0.020	
PAGE34LN	T: Linear: Percent Persons 25-34 Years					-0.002	-0.022		
PAGE34QU	T: Quadratic: Percent Persons 25-34 Years					-0.119	-0.038		
IHISP35	Hispanic Interaction Of PAGE34LN						-0.076		
IHISP36	Hispanic Interaction Of PAGE34QU						0.186		
PAGE44LN	T: Linear: Percent Persons 35-44 Years							0.060	
PAGE54LN	T: Linear: Percent Persons 45-54 Years				0.068		-0.136		-0.066
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years				0.012		-0.032		-0.057
PAGE54CU	T: Cubic: Percent Persons 45-54 Years						-0.057		
IBLK45	Black Interaction Of PAGE54LN				-0.133				
IBLK46	Black Interaction Of PAGE54QU				0.061				
IHISP45	Hispanic Interaction Of PAGE54LN						0.267		
PSCH8LN	T: Linear: Percent 0-8 Years Of School							-0.081	
PSCH8QU	T: Quadratic: Percent 0-8 Years Of School							0.026	
IBLK55	Black Interaction Of PSCH8LN							0.296	
IBLK56	Black Interaction Of PSCH8QU							-0.228	
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma	-0.002			-0.182	0.113			0.043
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma	0.067				0.133			0.137
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma	0.036							
PSCH12QR	T: Quartic: Percent 9-12 Years & No High School Diploma	0.028							
IFEM60	Female Interaction Of PSCH12LN					-0.115			
IFEM61	Female Interaction Of PSCH12QU					-0.181			
PSCHASLN	T: Linear: Percent Associates Degree							0.023	
PSCHASQU	T: Quadratic: Percent Associates Degree							-0.088	
IBLK65	Black Interaction Of PSCHASLN							0.100	
IBLK66	Black Interaction Of PSCHASQU							0.158	
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree					0.134			-0.003
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree					-0.007			-0.042
PSCHCOCU	T: Cubic: Bachelors, Graduate, Or Professional Degree					0.082			
IFEM75	Female Interaction Of PSCHCOLN					0.024			
IFEM76	Female Interaction Of PSCHCOQU					0.024			
PPOVERLN	T: Linear: Percent Families Below Poverty Level	-0.134				-0.122	-0.174	0.126	
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level	0.009				-0.006		0.021	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
PPOVERCU	T: Cubic: Percent Families Below Poverty Level	0.115				-0.052		0.037	
PPOVERQR	T: Quartic: Percent Families Below Poverty Level					-0.025		-0.023	
IFEM85	Female Interaction Of PPOVERLN						0.059		
PBLACKLN	T: Linear: Percent Black Nonhispanic			0.076			0.080	0.063	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic			-0.003			0.020	-0.053	
PBLACKCU	T: Cubic: Percent Black Nonhispanic			0.053			0.016	0.016	
PBLACKQR	T: Quartic: Percent Black Nonhispanic						-0.025		
PHISPLN	T: Linear: Percent Hispanic								0.082
PHISPQU	T: Quadratic: Percent Hispanic								-0.057
PHISPCU	T: Cubic: Percent Hispanic								-0.067
PHISPQR	T: Quartic: Percent Hispanic								-0.014
POTHLN	T: Linear: Percent Other Race/Hispanicity					-0.063			
POTHQU	T: Quadratic: Percent Other Race/Hispanicity					0.001			
POTHCU	T: Cubic: Percent Other Race/Hispanicity					-0.020			
PHHF18LN	T: Linear: % Female-Headed HH W/No Spouse & Chld Under 18					0.045			0.018
PHHF18QU	T: Quadratic: % Female-Headed HH W/No Spouse & Chld Under 18					0.020			
PHHF18CU	T: Cubic: % Female-Headed HH W/No Spouse & Chld Under 18					-0.014			
PHHF18QR	T: Quartic: % Female-Headed HH W/No Spouse & Chld Under 18					-0.014			
IFEM120	Female Interaction Of PHHF18LN								0.073
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed			0.086				0.003	
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed							0.033	
IHISP135	Hispanic Interaction Of PFNOTLN							-0.243	
IHISP136	Hispanic Interaction Of PFNOTQU							-0.077	
PMNEVLN	T: Linear: Percent Males Never Married				0.157				
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed							0.095	
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed							-0.037	
IHISP150	Hispanic Interaction Of PMNOTLN							0.041	
IHISP151	Hispanic Interaction Of PMNOTQU							0.157	
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier			0.056			0.100	0.007	
POLDHUQU	T: Quadratic: Percent Housing Units Built 1939 Or Earlier			-0.016			0.015	0.062	
POLDHUCU	T: Cubic: Percent Housing Units Built 1939 Or Earlier			-0.083					
P40HULN	T: Linear: Percent Housing Units Built 1940-1949					-0.025		0.145	
PRENTLN	T: Linear: Percent Housing Rented					-0.120	-0.037	-0.160	0.261
PRENTQU	T: Quadratic: Percent Housing Rented							0.059	0.137
PRENTCU	T: Cubic: Percent Housing Rented							0.038	0.007
PRENTQR	T: Quartic: Percent Housing Rented							0.032	-0.018
IFEM165	Female Interaction Of PRENTLN					0.280			
IBLK165	Black Interaction Of PRENTLN							0.232	
IBLK166	Black Interaction Of PRENTQU							-0.053	
IBLK167	Black Interaction Of PRENTCU							0.147	
IHISP165	Hispanic Interaction Of PRENTLN						0.086		
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases		0.004			-0.028	-0.100		
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases		0.004			-0.025	0.021		
ADRATECU	C: Cubic: Death Rate For All Alcohol-Related Cases		0.104			-0.059	-0.004		
ADRATEQR	C: Quartic: Death Rate For All Alcohol-Related Cases						0.054		
IBLK175	Black Interaction Of ADRATELN						-0.061		
IBLK176	Black Interaction Of ADRATEQU						0.098		
IBLK177	Black Interaction Of ADRATECU						0.059		
IBLK178	Black Interaction Of ADRATEQR						-0.089		
V18FLN	C: Linear: Marijuana Posession Arrest Rate					0.016			
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate						0.058	0.175	0.302
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate						-0.040		0.287
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate						0.030		0.086
V18EQR	C: Quartic: Opium/Cocaine & Deriv Posession Arrest Rate						-0.049		-0.002
IFEM195	Female Interaction Of V18ELN						0.128		
IFEM196	Female Interaction Of V18EQU						0.056		
IFEM197	Female Interaction Of V18ECU						-0.030		
IFEM198	Female Interaction Of V18EQR						0.058		
IBLK195	Black Interaction Of V18ELN						-0.282		
IHISP195	Hispanic Interaction Of V18ELN						-0.352		
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate	0.052	0.141			0.052		-0.007	
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate	0.091				0.048			
V18ACU	C: Cubic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate	-0.066				-0.119			
IFEM200	Female Interaction Of V18ALN	-0.000							
IFEM201	Female Interaction Of V18AQU	-0.047							
IFEM202	Female Interaction Of V18ACU	0.136							
IBLK200	Black Interaction Of V18ALN					-0.172			
IBLK201	Black Interaction Of V18AQU					-0.065			
IBLK202	Black Interaction Of V18ACU					0.211			
IHISP200	Hispanic Interaction Of V18ALN		-0.254						
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate		-0.049	0.003	0.066			-0.150	-0.076
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate			0.025				-0.070	-0.348
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate			0.009				-0.014	
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate			0.019				0.030	
IFEM225	Female Interaction Of V18LN							0.042	
IFEM226	Female Interaction Of V18QU							0.061	
IFEM227	Female Interaction Of V18CU							0.017	
IFEM228	Female Interaction Of V18QR							-0.027	
IHISP225	Hispanic Interaction Of V18LN		0.007						
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate				-0.022				-0.297
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate				-0.016				-0.147

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
IBLK230	Black Interaction Of VIOLLN				-0.146				
IBLK231	Black Interaction Of VIOLQU				0.029				
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992					0.099	0.020	-0.021	
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						-0.065	0.080	
IFEM235	Female Interaction Of DRATELN							0.069	
IFEM236	Female Interaction Of DRATEQU							-0.156	
RH43ALN	T: Linear: Recoded Median Rents For Rental Units	-0.024						-0.000	-0.255
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units	0.016							0.110
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units	0.020							
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units	-0.034							
IBLK240	Black Interaction Of RH43ALN								-0.012
IBLK241	Black Interaction Of RH43AQU								-0.233
RH61ALN	T: Linear: Recoded Median Value Of Owner Occupied HUs						-0.019		
RH61AQU	T: Quadratic: Recoded Median Value Of Owner Occupied HUs						-0.003		
RH61ACU	T: Cubic: Recoded Median Value Of Owner Occupied HUs						-0.097		
RP80ALN	T: Linear: Recoded Median Household Income					-0.144		-0.057	
RP80AQU	T: Quadratic: Recoded Median Household Income					0.014			
RP80ACU	T: Cubic: Recoded Median Household Income					0.056			
RP80AQR	T: Quartic: Recoded Median Household Income					-0.011			
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992	0.130				0.005			
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992					0.018			
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992					0.003			
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability					-0.003	0.004		
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability					-0.037	0.023		
IBLK260	Black Interaction Of B64DISLN						0.172		
BAGE34LN	B: Linear: Percent Persons 25-34 Years						0.097		
BAGE44LN	B: Linear: Percent Persons 35-44 Years					-0.118			
BAGE54LN	B: Linear: Percent Persons 45-54 Years						0.015	-0.015	0.051
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years						0.019	-0.068	-0.046
BAGE54CU	B: Cubic: Percent Persons 45-54 Years						-0.065	0.039	0.119
BAGE54QR	B: Quartic: Percent Persons 45-54 Years							0.034	
IFEM275	Female Interaction Of BAGE54LN						-0.017	-0.003	-0.023
IFEM276	Female Interaction Of BAGE54QU						-0.000	-0.003	0.120
IFEM277	Female Interaction Of BAGE54CU						0.144	-0.041	-0.175
IFEM278	Female Interaction Of BAGE54QR							-0.056	
IHISP275	Hispanic Interaction Of BAGE54LN							-0.159	
IHISP276	Hispanic Interaction Of BAGE54QU							-0.028	
IHISP277	Hispanic Interaction Of BAGE54CU							-0.204	
BAGE64LN	B: Linear: Percent Persons 55-64 Years	-0.032	-0.054	0.009	-0.101	-0.055	0.071		-0.239
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		-0.053	-0.016	0.048	-0.010	-0.034		-0.042
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		0.008	0.054		-0.075	0.048		0.065
IFEM280	Female Interaction Of BAGE64LN	0.140			0.112	0.045			0.082
IFEM281	Female Interaction Of BAGE64QU				-0.107	0.021			0.087
IFEM282	Female Interaction Of BAGE64CU					0.157			-0.020
IBLK280	Black Interaction Of BAGE64LN								0.130
IBLK281	Black Interaction Of BAGE64QU								-0.030
IHISP280	Hispanic Interaction Of BAGE64LN		0.095	-0.009			-0.017		
IHISP281	Hispanic Interaction Of BAGE64QU		0.013	-0.032			-0.028		
IHISP282	Hispanic Interaction Of BAGE64CU		-0.124	-0.128			-0.092		
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander		0.005	-0.009			0.061	0.027	0.051
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander		-0.024	0.001				0.165	0.026
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander			-0.048				-0.004	0.048
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander			0.012					0.025
IBLK285	Black Interaction Of BASIANLN								-0.082
IBLK286	Black Interaction Of BASIANQU								-0.239
IBLK287	Black Interaction Of BASIANCU								0.179
IBLK288	Black Interaction Of BASIANQR								-0.132
IHISP285	Hispanic Interaction Of BASIANLN			0.045				-0.032	
IHISP286	Hispanic Interaction Of BASIANQU			0.079				-0.076	
IHISP287	Hispanic Interaction Of BASIANCU			0.082				0.210	
IHISP288	Hispanic Interaction Of BASIANQR			-0.062					
BCUBANLN	B: Linear: Percent Hispanics: Cuban						0.270	-0.082	0.030
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban								-0.113
BCUBANCU	B: Cubic: Percent Hispanics: Cuban								0.330
IFEM290	Female Interaction Of BCUBANLN							0.214	
IBLK290	Black Interaction Of BCUBANLN								-0.017
IBLK291	Black Interaction Of BCUBANQU								0.347
IBLK292	Black Interaction Of BCUBANCU								-1.115
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed	0.074		-0.002			-0.031	0.128	0.142
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed	0.025		-0.030			-0.007	-0.033	
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed	0.033					0.070		
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed	0.025					0.018		
IFEM295	Female Interaction Of BFNOTLN						0.006		
IFEM296	Female Interaction Of BFNOTQU						-0.050		
IFEM297	Female Interaction Of BFNOTCU						-0.149		
IBLK295	Black Interaction Of BFNOTLN			-0.056				-0.237	
IBLK296	Black Interaction Of BFNOTQU			0.115				0.208	
IHISP295	Hispanic Interaction Of BFNOTLN						0.012		
IHISP296	Hispanic Interaction Of BFNOTQU						-0.068		
IHISP297	Hispanic Interaction Of BFNOTCU						-0.023		
IHISP298	Hispanic Interaction Of BFNOTQR						-0.039		
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut					0.108	0.074	0.158	-0.043
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut					-0.082			-0.029
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut					0.148			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IBLK300	Black Interaction Of BINDIALN					0.563	0.099		
IHISP300	Hispanic Interaction Of BINDIALN								0.351
IHISP301	Hispanic Interaction Of BINDIAQU								-0.480
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed				0.259	0.064	0.019	-0.008	
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed				-0.007	0.019	0.015	-0.007	
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed				0.082			0.049	
IFEM305	Female Interaction Of BMNOTLN				-0.276			-0.081	
IFEM306	Female Interaction Of BMNOTQU				0.000			-0.007	
IFEM307	Female Interaction Of BMNOTCU				-0.038			-0.108	
IHISP305	Hispanic Interaction Of BMNOTLN						-0.049		
IHISP306	Hispanic Interaction Of BMNOTQU						-0.112		
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.059				-0.030		-0.271
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.007						-0.035
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		-0.012						0.043
BPOVERQR	B: Quartic: Percent Families Below Poverty Level		-0.010						-0.101
IFEM310	Female Interaction Of BPOVERLN						0.151		0.063
IFEM311	Female Interaction Of BPOVERQU								-0.029
IFEM312	Female Interaction Of BPOVERCU								-0.135
IFEM313	Female Interaction Of BPOVERQR								0.138
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	-0.071	-0.098			-0.168	0.010	0.046	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican		-0.094			-0.159	-0.004	-0.051	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					-0.016	-0.078	0.022	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican					0.010			
IFEM315	Female Interaction Of BPRICALN					0.064			
IFEM316	Female Interaction Of BPRICAQU					0.150			
IFEM317	Female Interaction Of BPRICACU					-0.064			
IFEM318	Female Interaction Of BPRICAQR					-0.102			
IBLK315	Black Interaction Of BPRICALN		-0.035				-0.114	-0.168	
IBLK316	Black Interaction Of BPRICAQU		0.200				0.210	0.080	
IBLK317	Black Interaction Of BPRICACU							0.204	
IHISP315	Hispanic Interaction Of BPRICALN						-0.147		
IHISP316	Hispanic Interaction Of BPRICAQU						-0.019		
IHISP317	Hispanic Interaction Of BPRICACU						0.303		
BSCHASLN	B: Linear: Percent Associates Degree		-0.037			0.057	-0.016	0.031	0.110
BSCHASQU	B: Quadratic: Percent Associates Degree		0.013			-0.029	0.035	0.031	0.026
BSCHASCU	B: Cubic: Percent Associates Degree					-0.013	-0.028	0.066	-0.012
BSCHASQR	B: Quartic: Percent Associates Degree					-0.026			
IFEM320	Female Interaction Of BSCHASLN					-0.071			
IFEM321	Female Interaction Of BSCHASQU					0.080			
IFEM322	Female Interaction Of BSCHASCU					0.053			
IFEM323	Female Interaction Of BSCHASQR					0.064			
IBLK320	Black Interaction Of BSCHASLN					0.038		0.041	
IBLK321	Black Interaction Of BSCHASQU					0.126		0.063	
IBLK322	Black Interaction Of BSCHASCU					0.281		-0.087	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander				-0.143		-0.073	0.002	
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander						-0.000	-0.051	
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander						-0.067	-0.111	
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander						0.023		
IFEM325	Female Interaction Of PASIANLN						-0.035		
IFEM326	Female Interaction Of PASIANQU						-0.053		
IFEM327	Female Interaction Of PASIANCU						0.015		
IFEM328	Female Interaction Of PASIANQR						-0.045		
IHISP325	Hispanic Interaction Of PASIANLN						0.035		
IHISP326	Hispanic Interaction Of PASIANQU						0.099		
IHISP327	Hispanic Interaction Of PASIANCU						0.169		
PCUBANLN	T: Linear: Percent Hispanics: Cuban	-0.060		-0.055		-0.141		-0.019	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban	-0.114		0.015		-0.167		0.052	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban	-0.080				-0.123		0.012	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban							0.007	
IFEM330	Female Interaction Of PCUBANLN					0.137			
IFEM331	Female Interaction Of PCUBANQU					0.155			
IFEM332	Female Interaction Of PCUBANCU					0.356			
IBLK330	Black Interaction Of PCUBANLN	0.016							
IBLK331	Black Interaction Of PCUBANQU	0.060							
IBLK332	Black Interaction Of PCUBANCU	0.044							
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut	-0.038		0.008			-0.019		-0.071
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut	0.069		-0.035					0.004
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut	-0.108							-0.017
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut								0.025
IFEM335	Female Interaction Of PINDIALN								0.055
IFEM336	Female Interaction Of PINDIAQU								0.055
IFEM337	Female Interaction Of PINDIACU								0.054
IFEM338	Female Interaction Of PINDIAQR								-0.096
IHISP335	Hispanic Interaction Of PINDIALN			-0.083					
IHISP336	Hispanic Interaction Of PINDIAQU			0.152					

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E3. Coefficients of Model Parameters for Past Month Cigarette Use

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
DUMMY	O: Intercept Term	-2.096	-0.561	-0.606	-1.067	-2.495	-0.560	-0.930	-0.919
FEMALE	O: Female Indicator	0.115	-0.224	-0.098	-0.133	0.040	0.083	0.155	-0.328
FEMBLCK	O: Black Interaction Of FEMALE	-0.681	0.081	-0.271	-0.385	-0.022	-0.273	-0.376	-0.080
FEMHISP	O: Hispanic Interaction Of FEMALE	-0.186	-0.555	-0.640	-0.511			-0.947	-0.240
FEMOTHR	O: Other Interaction Of FEMALE	-0.212	-0.942	-1.580	-1.273	0.319	-1.145	-0.980	-0.064
RACEBLCK	O: Race/Black Indicator	-1.242	-1.112	0.049	0.355	-1.334	-0.903	0.075	0.145
RACEHISP	O: Race/Hispanic Indicator	-0.441	-0.251	-0.194	0.282			0.124	-0.094
RACEOTHR	O: Race/Other Indicator	-1.992	-0.251	0.047	-0.217	-2.685	-0.008	0.200	0.350
REGNOREA	O: Northeast Region Indicator						0.332		
REGSOUTH	O: South Region Indicator						0.073		
REGWEST	O: West Region Indicator						-0.060		
PDENLEV1	O: Large MSA					0.191	-0.121	0.105	
PDENLEV2	O: Medium MSA					0.476	-0.085	0.122	
PDENLEV3	O: Small MSA					0.353	-0.108	0.010	
PDENLEV4	O: NonMSA, Urban					0.436	0.004	0.014	
IHISP10	Hispanic Interaction Of PDENLEV1							-0.255	
IHISP11	Hispanic Interaction Of PDENLEV2							-0.193	
POPRMLN	T: Linear: Average Persons Per Room			0.081		-0.023		-0.014	
POPRMQU	T: Quadratic: Average Persons Per Room					-0.059			
POPRMCU	T: Cubic: Average Persons Per Room					-0.083			
IFEM20	Female Interaction Of POPRMLN							0.095	
IBLK20	Black Interaction Of POPRMLN					0.126			
IBLK21	Black Interaction Of POPRMQU					0.070			
IBLK22	Black Interaction Of POPRMCU					0.351			
IOTH20	Other Interaction Of POPRMLN					-0.949			
IOTH21	Other Interaction Of POPRMQU					-0.118			
IOTH22	Other Interaction Of POPRMCU					0.894			
PAGE18LN	T: Linear: Percent Persons 0-18 Years		0.123				0.174		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years						-0.020		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years						0.094		
IFEM25	Female Interaction Of PAGE18LN		-0.150				-0.085		
IFEM26	Female Interaction Of PAGE18QU						0.056		
IFEM27	Female Interaction Of PAGE18CU						-0.110		
PAGE24LN	T: Linear: Percent Persons 19-24 Years					0.069		-0.019	0.006
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years					-0.074		-0.030	-0.032
PAGE24CU	T: Cubic: Percent Persons 19-24 Years							-0.027	-0.011
PAGE24QR	T: Quartic: Percent Persons 19-24 Years							-0.007	0.015
IHISP30	Hispanic Interaction Of PAGE24LN							0.157	
IHISP31	Hispanic Interaction Of PAGE24QU							0.085	
IHISP32	Hispanic Interaction Of PAGE24CU							-0.010	
IHISP33	Hispanic Interaction Of PAGE24QR							0.021	
IOTH30	Other Interaction Of PAGE24LN								0.153
IOTH31	Other Interaction Of PAGE24QU								0.067
IOTH32	Other Interaction Of PAGE24CU								0.062
IOTH33	Other Interaction Of PAGE24QR								0.261
PAGE44LN	T: Linear: Percent Persons 35-44 Years							-0.023	
PAGE54LN	T: Linear: Percent Persons 45-54 Years				0.029	0.028			0.056
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years				0.001	0.064			-0.034
IBLK45	Black Interaction Of PAGE54LN				-0.057	0.018			
IBLK46	Black Interaction Of PAGE54QU				0.078	-0.238			
PAGE64LN	T: Linear: Percent Persons 55-64 Years					0.069			
PAGE64QU	T: Quadratic: Percent Persons 55-64 Years					0.022			
PAGE64CU	T: Cubic: Percent Persons 55-64 Years					0.010			
PAGE64QR	T: Quartic: Percent Persons 55-64 Years					0.031			
PSCH8LN	T: Linear: Percent 0-8 Years Of School					-0.043			
IOTH55	Other Interaction Of PSCH8LN					1.716			
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma					0.036			
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma					0.033			
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma					-0.045			
PSCHASLN	T: Linear: Percent Associates Degree					0.034		0.010	
PSCHASQU	T: Quadratic: Percent Associates Degree					0.043		0.003	
PSCHASCU	T: Cubic: Percent Associates Degree					-0.029			
PSCHASQR	T: Quartic: Percent Associates Degree					-0.063			
IFEM65	Female Interaction Of PSCHASLN					0.075			
IFEM66	Female Interaction Of PSCHASQU					-0.087			
IFEM67	Female Interaction Of PSCHASCU					0.005			
IFEM68	Female Interaction Of PSCHASQR					0.075			
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree					-0.103			-0.111
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree					-0.051			-0.048
IFEM75	Female Interaction Of PSCHCOLN					0.148			
IFEM76	Female Interaction Of PSCHCOQU					0.130			
PPOVERLN	T: Linear: Percent Families Below Poverty Level					0.031			0.084
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level					0.047			0.092
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					-0.072			
PPOVERQR	T: Quartic: Percent Families Below Poverty Level					0.002			
IFEM85	Female Interaction Of PPOVERLN					-0.170			-0.203
IFEM86	Female Interaction Of PPOVERQU					-0.075			-0.103
IFEM87	Female Interaction Of PPOVERCU					0.014			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
IFEM88	Female Interaction Of PPOVERQR					-0.069			
IOTH85	Other Interaction Of PPOVERLN					-0.409			
IOTH86	Other Interaction Of PPOVERQU					-0.275			
IOTH87	Other Interaction Of PPOVERCU					0.640			
PPUBASLN	T: Linear: Percent Households With Public Assist Income					-0.131			
IFEM90	Female Interaction Of PPUBASLN					0.264			
P64DISLN	T: Linear: Percent 16-64 With A Work Disability	0.086				-0.051			
IOTH95	Other Interaction Of P64DISLN	-0.516				0.704			
PBLACKLN	T: Linear: Percent Black Nonhispanic			-0.025				-0.033	-0.023
PBLACKQU	T: Quadratic: Percent Black Nonhispanic			-0.018				-0.048	-0.086
PBLACKCU	T: Cubic: Percent Black Nonhispanic			0.023				-0.031	
PHISPLN	T: Linear: Percent Hispanic					-0.067			-0.104
PHISPQU	T: Quadratic: Percent Hispanic					-0.042			-0.079
PHISPCU	T: Cubic: Percent Hispanic					0.042			-0.024
PHISPQR	T: Quartic: Percent Hispanic								0.008
POTHLN	T: Linear: Percent Other Race/Hispanicity					-0.198			0.046
POTHQU	T: Quadratic: Percent Other Race/Hispanicity					0.008			
POTHCU	T: Cubic: Percent Other Race/Hispanicity					-0.005			
IOTH110	Other Interaction Of POTHLN								-0.460
PHHF18LN	T: Linear: % Female-Headed HH W/No Spouse & Chld Under 18					-0.039			0.055
PHHF18QU	T: Quadratic: % Female-Headed HH W/No Spouse & Chld Under 18					-0.000			
PHHF18CU	T: Cubic: % Female-Headed HH W/No Spouse & Chld Under 18					0.036			
PHHF18QR	T: Quartic: % Female-Headed HH W/No Spouse & Chld Under 18					0.002			
IFEM120	Female Interaction Of PHHF18LN								0.051
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force						0.047		
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force						0.003		
PFLABCU	T: Cubic: Percent Females 16+ Years Old In Labor Force						0.050		
PFLABQR	T: Quartic: Percent Females 16+ Years Old In Labor Force						0.018		
IBLK125	Black Interaction Of PFLABLN						0.007		
IBLK126	Black Interaction Of PFLABQU						0.033		
IBLK127	Black Interaction Of PFLABCU						-0.046		
IBLK128	Black Interaction Of PFLABQR						-0.056		
PFNEVLN	T: Linear: Percent Females Never Married								0.044
PFNEVQU	T: Quadratic: Percent Females Never Married								0.059
PFNEVCU	T: Cubic: Percent Females Never Married								0.016
PFNEVQR	T: Quartic: Percent Females Never Married								-0.014
IHISP130	Hispanic Interaction Of PFNEVLN								0.002
IHISP131	Hispanic Interaction Of PFNEVQU								-0.058
IHISP132	Hispanic Interaction Of PFNEVCU								-0.005
IHISP133	Hispanic Interaction Of PFNEVQR								0.048
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed		0.106	0.001			0.023	0.039	
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed						-0.004	0.014	
PFNOTCU	T: Cubic: Percent Females Separated, Divorced Or Widowed						0.015	0.035	
PFNOTQR	T: Quartic: Percent Females Separated, Divorced Or Widowed						-0.009	0.014	
IHISP135	Hispanic Interaction Of PFNOTLN							0.021	
IHISP136	Hispanic Interaction Of PFNOTQU							-0.014	
IHISP137	Hispanic Interaction Of PFNOTCU							-0.031	
IHISP138	Hispanic Interaction Of PFNOTQR							-0.038	
IOTH135	Other Interaction Of PFNOTLN						0.089		
IOTH136	Other Interaction Of PFNOTQU						0.103		
IOTH137	Other Interaction Of PFNOTCU						-0.259		
IOTH138	Other Interaction Of PFNOTQR						-0.145		
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force					-0.027			
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force					-0.065			
IOTH145	Other Interaction Of PMLABLN					-0.283			
IOTH146	Other Interaction Of PMLABQU					0.465			
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed						0.015	0.090	
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed						0.012	-0.005	
IFEM150	Female Interaction Of PMNOTLN						0.084		
IFEM151	Female Interaction Of PMNOTQU						-0.069		
IHISP150	Hispanic Interaction Of PMNOTLN							-0.118	
IHISP151	Hispanic Interaction Of PMNOTQU							-0.020	
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier				-0.044				0.096
P40HULN	T: Linear: Percent Housing Units Built 1940-1949			-0.013		0.017		-0.014	
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949			-0.008		-0.014		-0.012	
P40HUCU	T: Cubic: Percent Housing Units Built 1940-1949			-0.000		-0.050		0.002	
P40HUQR	T: Quartic: Percent Housing Units Built 1940-1949							-0.009	
IBLK160	Black Interaction Of P40HULN			0.031				0.084	
IBLK161	Black Interaction Of P40HUQU			-0.041				0.001	
IBLK162	Black Interaction Of P40HUCU			0.092				-0.029	
IBLK163	Black Interaction Of P40HUQR							-0.044	
IOTH160	Other Interaction Of P40HULN					-0.589			
IOTH161	Other Interaction Of P40HUQU					0.797			
IOTH162	Other Interaction Of P40HUCU					0.496			
PRENTLN	T: Linear: Percent Housing Rented						0.060	-0.091	-0.012
PRENTQU	T: Quadratic: Percent Housing Rented						0.016		0.003
PRENTCU	T: Cubic: Percent Housing Rented						0.046		-0.007
PRENTQR	T: Quartic: Percent Housing Rented								0.001
IOTH165	Other Interaction Of PRENTLN						0.039		
IOTH166	Other Interaction Of PRENTQU						0.329		
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases					0.032			
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases					-0.036			
ADRATECU	C: Cubic: Death Rate For All Alcohol-Related Cases					-0.064			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol		-0.034				0.011		0.030
ADRAT1QU	C: Quadratic: Death Rate With Explicit Mention Of Alcohol		0.004				0.019		0.035
ADRAT1CU	C: Cubic: Death Rate With Explicit Mention Of Alcohol		-0.003				-0.006		0.072
ADRAT1QR	C: Quartic: Death Rate With Explicit Mention Of Alcohol		0.000				-0.021		0.027
IBLK180	Black Interaction Of ADRAT1LN		0.050				-0.097		
IBLK181	Black Interaction Of ADRAT1QU		-0.039				0.108		
IBLK182	Black Interaction Of ADRAT1CU		0.054				-0.050		
IBLK183	Black Interaction Of ADRAT1QR		-0.061				0.070		
V18FLN	C: Linear: Marijuana Posession Arrest Rate			-0.158		-0.017		-0.048	
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate							-0.008	
V18FCU	C: Cubic: Marijuana Posession Arrest Rate							-0.010	
V18FQR	C: Quartic: Marijuana Posession Arrest Rate							-0.026	
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate				0.042			-0.017	
IBLK190	Black Interaction Of V18BLN				0.138			0.189	
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate						-0.010		-0.111
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate						0.001		0.002
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate						-0.056		-0.006
V18EQR	C: Quartic: Opium/Cocaine & Deriv Posession Arrest Rate								-0.009
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate		0.026				0.074	0.010	0.132
IBLK200	Black Interaction Of V18ALN								-0.168
IHISP200	Hispanic Interaction Of V18ALN		-0.195						
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate		-0.016	-0.041			-0.065	-0.047	
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate			0.029			-0.020	0.047	
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate			0.097				0.067	
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate			0.010				0.011	
IFEM225	Female Interaction Of V18LN							0.095	
IFEM226	Female Interaction Of V18QU							-0.044	
IFEM227	Female Interaction Of V18CU							-0.044	
IFEM228	Female Interaction Of V18QR							-0.024	
IHISP225	Hispanic Interaction Of V18LN		0.061						
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate				-0.055				-0.060
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate				-0.001				-0.046
IBLK230	Black Interaction Of VIOLLN				-0.141				0.173
IBLK231	Black Interaction Of VIOLQU				-0.050				
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992					0.042	0.061		0.009
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						0.002		-0.005
IFEM235	Female Interaction Of DRATELN						-0.037		
IFEM236	Female Interaction Of DRATEQU						-0.029		
IOTH235	Other Interaction Of DRATELN								-0.134
IOTH236	Other Interaction Of DRATEQU								-0.615
RH43ALN	T: Linear: Recoded Median Rents For Rental Units	-0.015				-0.077		-0.054	
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units	0.055				-0.057			
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units	-0.001				-0.126			
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units	-0.001							
RP80ALN	T: Linear: Recoded Median Household Income		0.056	-0.086		-0.056	0.003	-0.116	-0.016
RP80AQU	T: Quadratic: Recoded Median Household Income					-0.008	-0.035		0.026
RP80ACU	T: Cubic: Recoded Median Household Income					0.025	-0.025		
RP80AQR	T: Quartic: Recoded Median Household Income					-0.027			
IBLK250	Black Interaction Of RP80ALN						-0.060		
IBLK251	Black Interaction Of RP80AQU						-0.014		
IBLK252	Black Interaction Of RP80ACU						0.005		
IHISP250	Hispanic Interaction Of RP80ALN		0.168						-0.034
IHISP251	Hispanic Interaction Of RP80AQU								-0.157
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992	0.056				0.038			
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992					0.031			
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992					-0.040			
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability			0.075	0.069	0.008	0.089	0.088	
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability					-0.065			
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability					-0.026			
B64DISQR	B: Quartic: Percent Persons 16-64 With A Work Disability					-0.025			
IFEM260	Female Interaction Of B64DISLN					0.022			
IFEM261	Female Interaction Of B64DISQU					0.078			
IFEM262	Female Interaction Of B64DISCU					0.071			
IFEM263	Female Interaction Of B64DISQR					0.071			
BAGE34LN	B: Linear: Percent Persons 25-34 Years				0.086	-0.005		0.038	0.116
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years					0.058		-0.033	0.028
BAGE34CU	B: Cubic: Percent Persons 25-34 Years							-0.018	0.066
IFEM265	Female Interaction Of BAGE34LN								-0.039
IFEM266	Female Interaction Of BAGE34QU								-0.035
IFEM267	Female Interaction Of BAGE34CU								-0.112
IBLK265	Black Interaction Of BAGE34LN								-0.130
IOTH265	Other Interaction Of BAGE34LN					0.953		-0.322	
IOTH266	Other Interaction Of BAGE34QU					-0.677		0.352	
IOTH267	Other Interaction Of BAGE34CU							0.445	
BAGE44LN	B: Linear: Percent Persons 35-44 Years							-0.027	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years							0.047	
IHISP270	Hispanic Interaction Of BAGE44LN							0.073	
IHISP271	Hispanic Interaction Of BAGE44QU							-0.121	
BAGE54LN	B: Linear: Percent Persons 45-54 Years					-0.017	0.001	-0.061	0.012
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years					0.003	-0.049	-0.023	0.017
BAGE54CU	B: Cubic: Percent Persons 45-54 Years					0.039	-0.099	-0.019	-0.017
BAGE54QR	B: Quartic: Percent Persons 45-54 Years							0.007	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

		BIG CITY				REMAINDER			
Variable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IFEM275	Female Interaction Of BAGE54LN						0.024		-0.001
IFEM276	Female Interaction Of BAGE54QU						0.051		0.021
IFEM277	Female Interaction Of BAGE54CU						0.135		0.008
IBLK275	Black Interaction Of BAGE54LN						-0.005		
IBLK276	Black Interaction Of BAGE54QU						0.013		
IOTH275	Other Interaction Of BAGE54LN					0.404		0.057	
IOTH276	Other Interaction Of BAGE54QU					-0.054		-0.049	
IOTH277	Other Interaction Of BAGE54CU					-0.754		0.187	
IOTH278	Other Interaction Of BAGE54QR							0.144	
BAGE64LN	B: Linear: Percent Persons 55-64 Years		-0.010	0.004	-0.084	-0.097	0.018		-0.036
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		-0.043	0.008	0.014		0.006		0.001
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		0.007	-0.004			0.025		-0.021
BAGE64QR	B: Quartic: Percent Persons 55-64 Years						-0.008		
IFEM280	Female Interaction Of BAGE64LN				0.028				0.056
IFEM281	Female Interaction Of BAGE64QU				-0.004				-0.026
IFEM282	Female Interaction Of BAGE64CU								-0.001
IBLK280	Black Interaction Of BAGE64LN								-0.040
IBLK281	Black Interaction Of BAGE64QU								0.033
IHISP280	Hispanic Interaction Of BAGE64LN		0.003	-0.063					
IHISP281	Hispanic Interaction Of BAGE64QU		0.007	-0.030					
IHISP282	Hispanic Interaction Of BAGE64CU		-0.042	0.008					
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander		-0.040	-0.005			-0.002	-0.079	-0.078
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander		-0.014	0.008			0.052	0.001	0.010
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander			-0.008			0.085	0.035	-0.140
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander			0.003			0.012	0.030	
IFEM285	Female Interaction Of BASIANLN						-0.033		0.031
IFEM286	Female Interaction Of BASIANQU						-0.028		0.006
IFEM287	Female Interaction Of BASIANCU						-0.055		0.183
IFEM288	Female Interaction Of BASIANQR						-0.047		
IHISP285	Hispanic Interaction Of BASIANLN			-0.003				0.197	
IHISP286	Hispanic Interaction Of BASIANQU			0.058					
IHISP287	Hispanic Interaction Of BASIANCU			0.003					
IHISP288	Hispanic Interaction Of BASIANQR			-0.041					
BCUBANLN	B: Linear: Percent Hispanics: Cuban						-0.079	0.052	
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban						0.029		
BCUBANCU	B: Cubic: Percent Hispanics: Cuban						0.064		
IFEM290	Female Interaction Of BCUBANLN							-0.114	
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed					0.025		0.004	
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed					0.008		-0.095	
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed					0.061			
IFEM295	Female Interaction Of BFNOTLN					0.008		0.026	
IFEM296	Female Interaction Of BFNOTQU					-0.045		0.085	
IFEM297	Female Interaction Of BFNOTCU					-0.176			
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut		0.061	-0.003			0.057		
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			0.024			-0.014		
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut						0.004		
BINDIAQR	B: Quartic: Percent Pop: American Indian, Eskimo, Aleut						0.006		
IBLK300	Black Interaction Of BINDIALN						0.222		
IHISP300	Hispanic Interaction Of BINDIALN			-0.120					
IHISP301	Hispanic Interaction Of BINDIAQU			0.168					
IOTH300	Other Interaction Of BINDIALN						0.440		
IOTH301	Other Interaction Of BINDIAQU						-0.426		
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed			0.041	0.132	0.024	0.058		0.067
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed			0.016	0.030	-0.111	0.000		-0.025
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed			-0.014	0.070		0.023		
BMNOTQR	B: Quartic: Percent Males Separated, Divorced or Widowed			-0.011					
IFEM305	Female Interaction Of BMNOTLN				-0.146	-0.030			
IFEM306	Female Interaction Of BMNOTQU				-0.060	0.157			
IFEM307	Female Interaction Of BMNOTCU				-0.011				
IHISP305	Hispanic Interaction Of BMNOTLN			-0.011					-0.048
IHISP306	Hispanic Interaction Of BMNOTQU			-0.011					0.171
IHISP307	Hispanic Interaction Of BMNOTCU			0.028					
IHISP308	Hispanic Interaction Of BMNOTQR			0.041					
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.105			0.006	-0.035		0.089
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.025			0.015	-0.017		-0.030
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		0.040			-0.039	-0.045		0.054
BPOVERQR	B: Quartic: Percent Families Below Poverty Level		0.008				0.004		
IFEM310	Female Interaction Of BPOVERLN								0.057
IFEM311	Female Interaction Of BPOVERQU								-0.012
IFEM312	Female Interaction Of BPOVERCU								-0.126
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	0.040				-0.082		0.097	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican					0.046		-0.009	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					-0.189		-0.050	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican							-0.039	
IFEM315	Female Interaction Of BPRICALN					0.128			
IBLK315	Black Interaction Of BPRICALN							-0.177	
IBLK316	Black Interaction Of BPRICAQU							0.094	
IBLK317	Black Interaction Of BPRICACU							0.026	
IBLK318	Black Interaction Of BPRICAQR							0.080	
BSCHASLN	B: Linear: Percent Associates Degree		0.031			0.019	0.048	-0.003	0.011
BSCHASQU	B: Quadratic: Percent Associates Degree		0.047			-0.051	0.004	-0.023	0.042
BSCHASCU	B: Cubic: Percent Associates Degree						0.016	0.007	0.023
BSCHASQR	B: Quartic: Percent Associates Degree						-0.014		
IFEM320	Female Interaction Of BSCHASLN					-0.157	-0.072		
IFEM321	Female Interaction Of BSCHASQU					0.094			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IBLK320	Black Interaction Of BSCHASLN							-0.023	
IBLK321	Black Interaction Of BSCHASQU							0.116	
IBLK322	Black Interaction Of BSCHASCU							-0.003	
IHISP320	Hispanic Interaction Of BSCHASLN		-0.075						
IHISP321	Hispanic Interaction Of BSCHASQU		-0.088						
IOTH320	Other Interaction Of BSCHASLN								-0.130
IOTH321	Other Interaction Of BSCHASQU								0.569
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander			-0.004		0.102	-0.018		
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander			-0.016		0.006	-0.014		
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander					0.005	-0.035		
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander						-0.019		
IBLK325	Black Interaction Of PASIANLN			0.073					
IBLK326	Black Interaction Of PASIANQU			0.101					
IOTH325	Other Interaction Of PASIANLN					0.286			
IOTH326	Other Interaction Of PASIANQU					-0.117			
IOTH327	Other Interaction Of PASIANCU					0.575			
PCUBANLN	T: Linear: Percent Hispanics: Cuban	0.011		0.016	-0.179	0.009	0.224	0.051	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban	0.073		0.035		0.093		0.010	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban	-0.065		0.013		-0.099		0.024	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban			0.027				-0.005	
IFEM330	Female Interaction Of PCUBANLN	0.067				-0.096			
IFEM331	Female Interaction Of PCUBANQU	-0.219				-0.052			
IFEM332	Female Interaction Of PCUBANCU					0.300			
IBLK330	Black Interaction Of PCUBANLN	-0.319					-0.278	0.043	
IBLK331	Black Interaction Of PCUBANQU	-0.058						-0.133	
IBLK332	Black Interaction Of PCUBANCU	0.128						0.073	
IBLK333	Black Interaction Of PCUBANQR							0.068	
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut			0.049		0.018		0.025	-0.015
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut							-0.004	-0.013
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut							-0.039	0.005
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut							0.008	0.026
IFEM335	Female Interaction Of PINDIALN							-0.007	
IFEM336	Female Interaction Of PINDIAQU							0.001	
IFEM337	Female Interaction Of PINDIACU							0.016	
IFEM338	Female Interaction Of PINDIAQR							-0.042	
IBLK335	Black Interaction Of PINDIALN					0.457		0.082	
IBLK336	Black Interaction Of PINDIAQU							0.125	
IHISP335	Hispanic Interaction Of PINDIALN							-0.041	
IHISP336	Hispanic Interaction Of PINDIAQU							0.050	
IHISP337	Hispanic Interaction Of PINDIACU							-0.016	
IHISP338	Hispanic Interaction Of PINDIAQR							0.068	
IOTH335	Other Interaction Of PINDIALN							0.318	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E4. Coefficients of Model Parameters for Past Month Cocaine Use

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	-6.184	-3.510	-3.721	-5.496	-6.529	-4.421	-5.037	-7.819
FEMALE	O: Female Indicator	0.941	-0.810	-0.624	-1.825	-0.262	-0.811	-0.335	-0.100
FEMBLCK	O: Black Interaction Of FEMALE	-1.977	0.086	0.319	0.875	0.246	-0.386	-0.538	-2.219
FEMHISP	O: Hispanic Interaction Of FEMALE	-0.554	0.197	-0.208	0.432	0.330	-0.419	-0.621	-1.322
RACEBLCK	O: Race/Black Indicator	0.539	-0.182	-0.175	1.287	-0.530	-0.286	-0.215	2.995
RACEHISP	O: Race/Hispanic Indicator	0.992	-0.703	-0.488	0.850	1.358	0.247	0.797	1.442
RACEOTHR	O: Race/Other Indicator	-0.937	-2.396				-0.246	-0.051	
PDENLEV1	O: Large MSA						0.186	0.133	
PDENLEV2	O: Medium MSA						0.285	0.277	
PDENLEV3	O: Small MSA						0.538	-0.319	
PDENLEV4	O: NonMSA, Urban						0.856	-0.344	
POPRMLN	T: Linear: Average Persons Per Room			-0.141				-0.439	
IFEM20	Female Interaction Of POPRMLN							0.529	
PAGE24LN	T: Linear: Percent Persons 19-24 Years							0.091	
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years							-0.002	
PAGE24CU	T: Cubic: Percent Persons 19-24 Years							-0.223	
PAGE24QR	T: Quartic: Percent Persons 19-24 Years							0.077	
IHISP30	Hispanic Interaction Of PAGE24LN							0.114	
IHISP31	Hispanic Interaction Of PAGE24QU							-0.010	
IHISP32	Hispanic Interaction Of PAGE24CU							0.344	
IHISP33	Hispanic Interaction Of PAGE24QR							-0.154	
PAGE44LN	T: Linear: Percent Persons 35-44 Years							0.250	
PAGE54LN	T: Linear: Percent Persons 45-54 Years				-0.128				-0.065
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years				-0.160				-0.315
IBLK45	Black Interaction Of PAGE54LN				0.103				
IBLK46	Black Interaction Of PAGE54QU				0.325				
PSCHASLN	T: Linear: Percent Associates Degree							-0.042	
PSCHASQU	T: Quadratic: Percent Associates Degree							-0.185	
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree					-0.169			-0.117
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree					0.356			-0.290
IFEM75	Female Interaction Of PSCHCOLN					0.606			
IFEM76	Female Interaction Of PSCHCOQU					-0.710			
PBLACKLN	T: Linear: Percent Black Nonhispanic			0.155				0.116	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic			0.035				0.052	
PBLACKCU	T: Cubic: Percent Black Nonhispanic			0.107				0.189	
PHISPLN	T: Linear: Percent Hispanic								0.128
PHISPQU	T: Quadratic: Percent Hispanic								-0.112
PHISPCU	T: Cubic: Percent Hispanic								0.063
PHISPQR	T: Quartic: Percent Hispanic								0.140
POTHLN	T: Linear: Percent Other Race/Hispanicity					-0.046			
POTHQU	T: Quadratic: Percent Other Race/Hispanicity					-0.165			
POTHCU	T: Cubic: Percent Other Race/Hispanicity					0.239			
PHHF18LN	T: Linear: % Female-Headed HH W/No Spouse & Chld Under 18					0.252			-0.498
PHHF18QU	T: Quadratic: % Female-Headed HH W/No Spouse & Chld Under 18					-0.070			
PHHF18CU	T: Cubic: % Female-Headed HH W/No Spouse & Chld Under 18					0.140			
PHHF18QR	T: Quartic: % Female-Headed HH W/No Spouse & Chld Under 18					0.129			
IFEM120	Female Interaction Of PHHF18LN								0.979
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed			0.172				0.176	
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed							0.191	
IHISP135	Hispanic Interaction Of PFNOTLN							-0.322	
IHISP136	Hispanic Interaction Of PFNOTQU							-0.381	
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed						-0.178	-0.066	
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed						0.171	-0.121	
IFEM150	Female Interaction Of PMNOTLN						0.550		
IFEM151	Female Interaction Of PMNOTQU						-0.299		
IHISP150	Hispanic Interaction Of PMNOTLN							-0.053	
IHISP151	Hispanic Interaction Of PMNOTQU							0.285	
PRENTLN	T: Linear: Percent Housing Rented								0.586
PRENTQU	T: Quadratic: Percent Housing Rented								0.155
PRENTCU	T: Cubic: Percent Housing Rented								0.274
PRENTQR	T: Quartic: Percent Housing Rented								-0.136
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate							0.192	
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate							0.028	
V18BCU	C: Cubic: Marijuana Sale/Manufacture Arrest Rate							0.288	
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate								0.250
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate								0.446
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate								0.088
V18EQR	C: Quartic: Opium/Cocaine & Deriv Posession Arrest Rate								0.165
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate		0.240					0.272	
IHISP200	Hispanic Interaction Of V18ALN		-0.361						
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate		-0.279	0.072			0.028	-0.691	
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate			-0.004			-0.072	0.018	
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate			-0.106				-0.069	
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate			0.012				0.140	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IFEM225	Female Interaction Of V18LN							0.161	
IFEM226	Female Interaction Of V18QU							-0.039	
IFEM227	Female Interaction Of V18CU							0.174	
IFEM228	Female Interaction Of V18QR							-0.132	
IHISP225	Hispanic Interaction Of V18LN		0.432						
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate				0.164				-0.310
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate				-0.116				-0.551
IBLK230	Black Interaction Of VIOLLN				-0.331				
IBLK231	Black Interaction Of VIOLQU				0.208				
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992						0.046		
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						-0.171		
IFEM235	Female Interaction Of DRATELN						-0.048		
IFEM236	Female Interaction Of DRATEQU						0.334		
RH43ALN	T: Linear: Recoded Median Rents For Rental Units	-0.068						0.489	
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units	0.052							
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units	-0.178							
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units	-0.083							
RP80ALN	T: Linear: Recoded Median Household Income					-0.026	-0.027	-0.686	
RP80AQU	T: Quadratic: Recoded Median Household Income					0.190	-0.071		
RP80ACU	T: Cubic: Recoded Median Household Income					0.014	-0.183		
RP80AQR	T: Quartic: Recoded Median Household Income					-0.054			
IBLK250	Black Interaction Of RP80ALN						-0.280		
IBLK251	Black Interaction Of RP80AQU						0.175		
IBLK252	Black Interaction Of RP80ACU						0.438		
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992	0.339							
BAGE54LN	B: Linear: Percent Persons 45-54 Years						-0.016		-0.270
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years						-0.194		-0.154
BAGE54CU	B: Cubic: Percent Persons 45-54 Years								0.163
IFEM275	Female Interaction Of BAGE54LN								0.578
IFEM276	Female Interaction Of BAGE54QU								0.417
IFEM277	Female Interaction Of BAGE54CU								-0.840
IBLK275	Black Interaction Of BAGE54LN						0.013		
IBLK276	Black Interaction Of BAGE54QU						0.438		
BAGE64LN	B: Linear: Percent Persons 55-64 Years		-0.089	-0.020	-0.156		0.176		-0.004
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		-0.100	0.005	0.195		-0.032		-0.579
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		0.039	0.058			0.110		0.204
BAGE64QR	B: Quartic: Percent Persons 55-64 Years						-0.003		
IFEM280	Female Interaction Of BAGE64LN				0.027				-0.343
IFEM281	Female Interaction Of BAGE64QU				-0.272				0.491
IFEM282	Female Interaction Of BAGE64CU								-0.629
IBLK280	Black Interaction Of BAGE64LN								0.208
IBLK281	Black Interaction Of BAGE64QU								0.628
IHISP280	Hispanic Interaction Of BAGE64LN		-0.002	-0.110					
IHISP281	Hispanic Interaction Of BAGE64QU		0.057	0.047					
IHISP282	Hispanic Interaction Of BAGE64CU		-0.352	-0.415					
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander		-0.037				0.074		
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander		-0.115				0.088		
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander						-0.022		
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander						-0.083		
BCUBANLN	B: Linear: Percent Hispanics: Cuban						0.028	-0.605	
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban						0.071		
BCUBANCU	B: Cubic: Percent Hispanics: Cuban						-0.348		
IFEM290	Female Interaction Of BCUBANLN							0.686	
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed				0.373	0.250	0.124		
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed				-0.187	0.269	-0.062		
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed				0.303		0.120		
IFEM305	Female Interaction Of BMNOTLN				-0.394				
IFEM306	Female Interaction Of BMNOTQU				0.159				
IFEM307	Female Interaction Of BMNOTCU				-0.380				
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.172				0.077		
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		-0.050				0.022		
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		-0.050				-0.070		
BPOVERQR	B: Quartic: Percent Families Below Poverty Level		-0.049				-0.055		
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	0.165				0.072		-0.248	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican							0.174	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican							-0.296	
IFEM315	Female Interaction Of BPRICALN					-0.388			
IBLK315	Black Interaction Of BPRICALN							0.028	
IBLK316	Black Interaction Of BPRICAQU							-0.352	
IBLK317	Black Interaction Of BPRICACU							1.067	
BSCHASLN	B: Linear: Percent Associates Degree		-0.083			0.181	0.120	-0.167	0.052
BSCHASQU	B: Quadratic: Percent Associates Degree		0.106			-0.336	0.029	-0.038	0.033
BSCHASCU	B: Cubic: Percent Associates Degree							0.155	-0.370
IFEM320	Female Interaction Of BSCHASLN					-0.243	-0.369		
IFEM321	Female Interaction Of BSCHASQU					0.552			
IBLK320	Black Interaction Of BSCHASLN							-0.174	
IBLK321	Black Interaction Of BSCHASQU							0.164	
IBLK322	Black Interaction Of BSCHASCU							-0.379	
PCUBANLN	T: Linear: Percent Hispanics: Cuban	0.284		-0.147				0.143	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban	-0.227		0.102				-0.083	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban	-0.144						0.172	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban							0.126	
IBLK330	Black Interaction Of PCUBANLN	0.256							

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IBLK331	Black Interaction Of PCUBANQU	-0.567							
IBLK332	Black Interaction Of PCUBANCU	0.598							

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E5. Coefficients of Model Parameters for Past Month Any Illicit Drug But Marijuana

		BIG CITY				REMAINDER			
Variable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	-3.502	-2.451	-2.988	-4.111	-3.613	-2.902	-3.810	-5.295
FEMALE	O: Female Indicator	0.173	-0.545	-0.343	-0.442	-0.199	-0.457	-0.271	0.051
FEMBLCK	O: Black Interaction Of FEMALE	-0.547	-0.150	0.136	0.267	0.696	0.263	-0.345	-0.580
FEMHISP	O: Hispanic Interaction Of FEMALE	0.093	0.358	0.370	-0.239		-0.001	-0.035	0.087
FEMOTHR	O: Other Interaction Of FEMALE	0.380	1.261	-0.015		-0.670	-0.526		1.072
RACEBLCK	O: Race/Black Indicator	-0.409	-0.732	-0.371	0.225	-0.914	-0.497	-0.313	1.159
RACEHISP	O: Race/Hispanic Indicator	-0.022	-0.717	-0.549	0.615		-0.320	-0.227	-0.305
RACEOTHR	O: Race/Other Indicator	-1.119	-2.089	-1.395		-0.929	-3.417	-1.208	-0.812
REGNOREA	O: Northeast Region Indicator							0.540	
REGSOUTH	O: South Region Indicator							0.380	
REGWEST	O: West Region Indicator						0.606	1.056	
PDENLEV1	O: Large MSA								0.371
PDENLEV2	O: Medium MSA								-0.035
PDENLEV3	O: Small MSA								1.149
PDENLEV4	O: NonMSA, Urban								0.023
PHH1PLN	T: Linear: Percent One Person Households					0.222			
PAGE24LN	T: Linear: Percent Persons 19-24 Years							0.087	-0.124
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years							-0.021	0.127
PAGE24CU	T: Cubic: Percent Persons 19-24 Years							-0.121	
PAGE24QR	T: Quartic: Percent Persons 19-24 Years							0.045	
PAGE34LN	T: Linear: Percent Persons 25-34 Years					-0.080			
PAGE34QU	T: Quadratic: Percent Persons 25-34 Years					-0.168			
PAGE44LN	T: Linear: Percent Persons 35-44 Years					0.037			
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years					-0.002			
PAGE54LN	T: Linear: Percent Persons 45-54 Years								-0.115
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years								-0.347
IFEM45	Female Interaction Of PAGE54LN								0.010
IFEM46	Female Interaction Of PAGE54QU								0.328
PSCH8LN	T: Linear: Percent 0-8 Years Of School							-0.171	
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma	0.042				0.103			
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma	0.104				0.101			
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma	0.005				0.010			
PSCH12QR	T: Quartic: Percent 9-12 Years & No High School Diploma	0.052				0.040			
IFEM60	Female Interaction Of PSCH12LN	-0.122				-0.166			
IFEM61	Female Interaction Of PSCH12QU	-0.183				-0.187			
PSCHASLN	T: Linear: Percent Associates Degree					0.065			
PSCHASQU	T: Quadratic: Percent Associates Degree					-0.081			
IBLK65	Black Interaction Of PSCHASLN					0.228			
IBLK66	Black Interaction Of PSCHASQU					0.370			
PSCHSCLN	T: Linear: Percent Some College And No Degree					0.012			
PSCHSCQU	T: Quadratic: Percent Some College And No Degree					0.106			
PPOVERLN	T: Linear: Percent Families Below Poverty Level							0.074	
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level							-0.037	
PPOVERCU	T: Cubic: Percent Families Below Poverty Level							-0.003	
PPOVERQR	T: Quartic: Percent Families Below Poverty Level							-0.051	
P64DISLN	T: Linear: Percent 16-64 With A Work Disability	0.073				0.126			
P64DISQU	T: Quadratic: Percent 16-64 With A Work Disability	0.057				0.085			
P64DISCU	T: Cubic: Percent 16-64 With A Work Disability	0.084				0.075			
IBLK95	Black Interaction Of P64DISLN	-0.215				0.094			
IBLK96	Black Interaction Of P64DISQU	0.007				0.252			
IBLK97	Black Interaction Of P64DISCU	-0.296				-0.579			
PBLACKLN	T: Linear: Percent Black Nonhispanic			0.064				0.129	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic			0.004				-0.015	
PBLACKCU	T: Cubic: Percent Black Nonhispanic			0.078				0.125	
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed			0.168				0.228	
PMNEVLN	T: Linear: Percent Males Never Married					-0.177			
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed					0.015			
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed					0.004			
PMNOTCU	T: Cubic: Percent Males Separated, Divorced Or Widowed					0.035			
PMNOTQR	T: Quartic: Percent Males Separated, Divorced Or Widowed					-0.036			
P40HULN	T: Linear: Percent Housing Units Built 1940-1949					-0.145		0.123	
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949							-0.039	
P40HUCU	T: Cubic: Percent Housing Units Built 1940-1949							-0.077	
P40HUQR	T: Quartic: Percent Housing Units Built 1940-1949							-0.044	
PRENTLN	T: Linear: Percent Housing Rented							-0.084	0.354
PRENTQU	T: Quadratic: Percent Housing Rented							0.111	
ADRATEQR	C: Quartic: Death Rate For All Alcohol-Related Cases						0.048		
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol						-0.176		
V18FLN	C: Linear: Marijuana Posession Arrest Rate					0.092		-0.119	
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate					-0.081		-0.035	
V18FCU	C: Cubic: Marijuana Posession Arrest Rate							-0.083	
V18FQR	C: Quartic: Marijuana Posession Arrest Rate							-0.053	
IHISP185	Hispanic Interaction Of V18FLN							0.246	
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate								-0.016
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate								-0.155

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate					-0.026			
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate					0.035			
V18ACU	C: Cubic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate					-0.119			
V18AQR	C: Quartic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate					-0.051			
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate		-0.184						
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992						0.131		
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						-0.117		
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability					-0.062			
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability					-0.078			
BAGE34LN	B: Linear: Percent Persons 25-34 Years					0.008	0.158	0.191	
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years					-0.035		-0.113	
BAGE34CU	B: Cubic: Percent Persons 25-34 Years					-0.055		0.104	
IFEM265	Female Interaction Of BAGE34LN							-0.142	
IFEM266	Female Interaction Of BAGE34QU							0.112	
IFEM267	Female Interaction Of BAGE34CU							-0.186	
IOTH265	Other Interaction Of BAGE34LN					-0.576		-1.421	
IOTH266	Other Interaction Of BAGE34QU					0.798			
IOTH267	Other Interaction Of BAGE34CU					1.010			
BAGE44LN	B: Linear: Percent Persons 35-44 Years					-0.182	-0.028	-0.004	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years						0.066	-0.046	
IBLK270	Black Interaction Of BAGE44LN							0.028	
IBLK271	Black Interaction Of BAGE44QU							0.259	
IOTH270	Other Interaction Of BAGE44LN						0.272		
IOTH271	Other Interaction Of BAGE44QU						1.219		
BAGE54LN	B: Linear: Percent Persons 45-54 Years							-0.031	
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years							-0.102	
BAGE64LN	B: Linear: Percent Persons 55-64 Years		-0.098	-0.023		-0.143	0.138		
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		-0.062	0.002			-0.055		
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		0.045	0.083			0.107		
BAGE64QR	B: Quartic: Percent Persons 55-64 Years						-0.032		
IHISP280	Hispanic Interaction Of BAGE64LN		0.114	0.041			0.144		
IHISP281	Hispanic Interaction Of BAGE64QU		0.001	-0.100			0.077		
IHISP282	Hispanic Interaction Of BAGE64CU		-0.233	-0.250			0.095		
IHISP283	Hispanic Interaction Of BAGE64QR						0.100		
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander						0.012	-0.050	
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander						0.099	0.167	
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander						-0.034		
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander						-0.046		
IOTH285	Other Interaction Of BASIANLN						2.215		
BCUBANLN	B: Linear: Percent Hispanics: Cuban				-0.071	0.260			
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban				-0.117	0.264			
BCUBANCU	B: Cubic: Percent Hispanics: Cuban					-0.155			
IFEM290	Female Interaction Of BCUBANLN					-0.070			
IFEM291	Female Interaction Of BCUBANQU					0.008			
IFEM292	Female Interaction Of BCUBANCU					0.408			
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed	-0.038				0.004	-0.032		
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed	-0.014				0.017	-0.040		
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed	-0.008				-0.056	0.035		
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed	0.035				0.024			
IFEM295	Female Interaction Of BFNOTLN						0.117		
IFEM296	Female Interaction Of BFNOTQU						0.028		
IFEM297	Female Interaction Of BFNOTCU						-0.190		
IBLK295	Black Interaction Of BFNOTLN					0.102	-0.024		
IBLK296	Black Interaction Of BFNOTQU					-0.032	0.298		
IBLK297	Black Interaction Of BFNOTCU					0.219			
IBLK298	Black Interaction Of BFNOTQR					-0.118			
IHISP295	Hispanic Interaction Of BFNOTLN						-0.244		
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut					0.095			
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut					-0.056			
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut					0.215			
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed				0.164				-0.018
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed								0.062
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed								0.231
IFEM305	Female Interaction Of BMNOTLN				-0.351				
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.092				-0.072		
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.043				-0.132		
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		-0.033				-0.151		
BPOVERQR	B: Quartic: Percent Families Below Poverty Level		-0.027				-0.032		
IFEM310	Female Interaction Of BPOVERLN		0.061				0.188		
IFEM311	Female Interaction Of BPOVERQU		-0.108				0.129		
IFEM312	Female Interaction Of BPOVERCU		0.191				0.204		
IHISP310	Hispanic Interaction Of BPOVERLN		-0.280				-0.015		
IHISP311	Hispanic Interaction Of BPOVERQU						0.246		
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	-0.161				-0.206		-0.087	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican					-0.246		0.162	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					-0.160		-0.224	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican					0.022			
IFEM315	Female Interaction Of BPRICALN					-0.027		0.225	
IFEM316	Female Interaction Of BPRICAQU					0.444		-0.100	
IFEM317	Female Interaction Of BPRICACU					0.118		0.379	
IFEM318	Female Interaction Of BPRICAQR					-0.269			
BSCHASLN	B: Linear: Percent Associates Degree	0.020	-0.092			-0.040	-0.004	-0.080	
BSCHASQU	B: Quadratic: Percent Associates Degree	0.026				-0.137	0.077	0.005	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
BSCHASCU	B: Cubic: Percent Associates Degree	0.072				-0.089		0.119	
BSCHASQR	B: Quartic: Percent Associates Degree	0.041				-0.038			
IFEM320	Female Interaction Of BSCHASLN					-0.033			
IFEM321	Female Interaction Of BSCHASQU					0.257			
IFEM322	Female Interaction Of BSCHASCU					0.190			
IFEM323	Female Interaction Of BSCHASQR					0.061			
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander						-0.004	0.039	0.215
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander						-0.036	-0.046	
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander						-0.110	-0.124	
IHISP325	Hispanic Interaction Of PASIANLN						-0.161		
IHISP326	Hispanic Interaction Of PASIANQU						0.093		
IHISP327	Hispanic Interaction Of PASIANCU						0.209		
IOTH325	Other Interaction Of PASIANLN						-1.705		
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut						-0.056	0.116	
IHISP335	Hispanic Interaction Of PINDIALN							-0.325	
IOTH335	Other Interaction Of PINDIALN						1.008		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E6. Coefficients of Model Parameters for Past Year Alcohol Treatment

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
DUMMY	O: Intercept Term	-6.554	-5.328	-4.383	-5.549	-6.694	-6.307	-4.376	-7.191
FEMALE	O: Female Indicator	0.006	-1.113	-0.878	-0.744	-0.111	-1.695	-1.074	-1.606
FEMBLCK	O: Black Interaction Of FEMALE	0.075		0.354	-1.245	0.372	0.461	-0.522	-1.905
FEMHISP	O: Hispanic Interaction Of FEMALE	0.658	-1.013	-0.610	-1.864	0.406	-1.665	-1.319	-1.947
RACEBLCK	O: Race/Black Indicator	-0.946	-1.951	-1.432	-1.170	-1.783	-1.335	-0.602	0.362
RACEHISP	O: Race/Hispanic Indicator	-0.130	-0.067	-2.533	-1.033	-0.244	-0.632	-0.649	-3.137
REGNOREA	O: Northeast Region Indicator						-0.058		
REGWEST	O: West Region Indicator						0.079		
PAGE18LN	T: Linear: Percent Persons 0-18 Years		0.513				-0.254		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years						-0.284		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years						0.085		
PAGE18QR	T: Quartic: Percent Persons 0-18 Years						-0.071		
IHISP25	Hispanic Interaction Of PAGE18LN						0.037		
IHISP26	Hispanic Interaction Of PAGE18QU						0.330		
IHISP27	Hispanic Interaction Of PAGE18CU						0.272		
IHISP28	Hispanic Interaction Of PAGE18QR						0.225		
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma		-0.486				0.128		
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma		-0.404				-0.801		
IFEM60	Female Interaction Of PSCH12LN						-0.170		
IFEM61	Female Interaction Of PSCH12QU						0.760		
IHISP60	Hispanic Interaction Of PSCH12LN						0.270		
IHISP61	Hispanic Interaction Of PSCH12QU						0.727		
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree							-0.219	0.172
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree							-0.192	-0.006
PSCHCOCU	T: Cubic: Bachelors, Graduate, Or Professional Degree								0.388
PSCHCOQR	T: Quartic: Bachelors, Graduate, Or Professional Degree								0.081
PSCHSCLN	T: Linear: Percent Some College And No Degree	0.548				-0.093			
PSCHSCQU	T: Quadratic: Percent Some College And No Degree	-0.093				0.180			
PSCHSCCU	T: Cubic: Percent Some College And No Degree					-0.177			
IHISP80	Hispanic Interaction Of PSCHSCLN	-0.393							
IHISP81	Hispanic Interaction Of PSCHSCQU	0.491							
PPOVERLN	T: Linear: Percent Families Below Poverty Level	-0.246				0.255			
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level	-0.410				-0.247			
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					-0.473			
PPUBASLN	T: Linear: Percent Households With Public Assist Income		-0.549				0.317		
PPUBASQU	T: Quadratic: Percent Households With Public Assist Income		0.118				0.185		
PPUBASCU	T: Cubic: Percent Households With Public Assist Income		0.288						
IHISP90	Hispanic Interaction Of PPUBASLN		0.603				-0.253		
IHISP91	Hispanic Interaction Of PPUBASQU						-0.712		
PBLACKLN	T: Linear: Percent Black Nonhispanic								-0.066
PBLACKQU	T: Quadratic: Percent Black Nonhispanic								-0.303
POTHLN	T: Linear: Percent Other Race/Hispanicity						0.062		
POTHQU	T: Quadratic: Percent Other Race/Hispanicity						-0.087		
POTHCU	T: Cubic: Percent Other Race/Hispanicity						-0.235		
IFEM110	Female Interaction Of POTHLN						0.249		
IFEM111	Female Interaction Of POTHQU						0.146		
IFEM112	Female Interaction Of POTHCU						0.689		
PFNEVLN	T: Linear: Percent Females Never Married						0.632		
PMNEVLN	T: Linear: Percent Males Never Married		0.407		-0.026		-0.566		0.288
PMNEVQU	T: Quadratic: Percent Males Never Married						-0.213		0.037
PMNEVCU	T: Cubic: Percent Males Never Married						0.162		-0.329
IBLK140	Black Interaction Of PMNEVLN						0.028		0.120
IBLK141	Black Interaction Of PMNEVQU						0.135		-0.126
IBLK142	Black Interaction Of PMNEVCU						-0.758		0.873
IHISP140	Hispanic Interaction Of PMNEVLN		-0.737		-0.815		-0.633		-1.569
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force				-0.011	-0.003			0.355
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force				0.209	0.053			0.060
PMLABCU	T: Cubic: Percent Males 16+ Years Old In Labor Force					0.038			0.033
PMLABQR	T: Quartic: Percent Males 16+ Years Old In Labor Force					-0.108			
IFEM145	Female Interaction Of PMLABLN				0.336				-0.023
IFEM146	Female Interaction Of PMLABQU				-0.559				0.248
IFEM147	Female Interaction Of PMLABCU								0.920
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier				0.230	0.133			
POLDHUQU	T: Quadratic: Percent Housing Units Built 1939 Or Earlier					-0.229			
POLDHUCU	T: Cubic: Percent Housing Units Built 1939 Or Earlier					-0.156			
IFEM155	Female Interaction Of POLDHULN				-0.578	-0.255			
IFEM156	Female Interaction Of POLDHUQU					0.317			
IFEM157	Female Interaction Of POLDHUCU					0.480			
P40HULN	T: Linear: Percent Housing Units Built 1940-1949								0.320
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol			0.026			-0.371		
ADRAT1QU	C: Quadratic: Death Rate With Explicit Mention Of Alcohol			-0.078			-0.360		
ADRAT1CU	C: Cubic: Death Rate With Explicit Mention Of Alcohol			0.134			-0.551		
ADRAT1QR	C: Quartic: Death Rate With Explicit Mention Of Alcohol			-0.095					
V18FLN	C: Linear: Marijuana Posession Arrest Rate						-0.416		
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate						-0.311		
IBLK185	Black Interaction Of V18FLN						0.381		
IBLK186	Black Interaction Of V18FQU						0.519		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate						-0.129		
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate						-0.033		
IFEM195	Female Interaction Of V18ELN						-0.959		
IFEM196	Female Interaction Of V18EQU						-1.036		
IBLK195	Black Interaction Of V18ELN					-0.345			
IBLK198	Black Interaction Of V18EQR					0.158			
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate					0.126			
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate					-0.006			
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate					0.379			
IFEM225	Female Interaction Of V18LN					-0.047			
IFEM226	Female Interaction Of V18QU					0.024			
IFEM227	Female Interaction Of V18CU					-0.584			
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate					0.006			
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate					-0.033			
VIOLCU	C: Cubic: Total Violent Offenses Arrest Rate					0.235			
VIOLQR	C: Quartic: Total Violent Offenses Arrest Rate					0.109			
RH43ALN	T: Linear: Recoded Median Rents For Rental Units				0.295				-0.162
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units				0.335				0.288
RH61ALN	T: Linear: Recoded Median Value Of Owner Occupied HUs				-0.451				
IBLK245	Black Interaction Of RH61ALN				1.166				
RP80ALN	T: Linear: Recoded Median Household Income					0.374	-0.221		
RP80AQU	T: Quadratic: Recoded Median Household Income					0.072	0.206		
RP80ACU	T: Cubic: Recoded Median Household Income					-0.295			
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992			0.095		-0.243		-0.037	
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992			0.168				-0.097	
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992			-0.250				0.184	
IHISP255	Hispanic Interaction Of ARATELN			2.239				0.539	
IHISP256	Hispanic Interaction Of ARATEQU			-1.707				-0.108	
IHISP257	Hispanic Interaction Of ARATECU			1.671				-0.715	
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability						-0.340		
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability						-0.101		
IBLK260	Black Interaction Of B64DISLN						-0.190		
IBLK261	Black Interaction Of B64DISQU						0.785		
BAGE34LN	B: Linear: Percent Persons 25-34 Years				0.207	-0.086			0.037
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years				0.161	-0.179			-0.236
BAGE34CU	B: Cubic: Percent Persons 25-34 Years				-0.332				0.450
BAGE34QR	B: Quartic: Percent Persons 25-34 Years								0.021
IFEM265	Female Interaction Of BAGE34LN				-0.263				-0.568
IFEM266	Female Interaction Of BAGE34QU				-0.233				-0.052
IFEM267	Female Interaction Of BAGE34CU				0.586				-0.659
IFEM268	Female Interaction Of BAGE34QR								-0.296
IBLK265	Black Interaction Of BAGE34LN					-0.087			
IBLK266	Black Interaction Of BAGE34QU					0.759			
IHISP265	Hispanic Interaction Of BAGE34LN				-0.087				0.142
IHISP266	Hispanic Interaction Of BAGE34QU				-0.729				-1.552
IHISP267	Hispanic Interaction Of BAGE34CU								-0.894
BAGE44LN	B: Linear: Percent Persons 35-44 Years						-0.175		
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years						-0.022		
BAGE44CU	B: Cubic: Percent Persons 35-44 Years						0.029		
BAGE44QR	B: Quartic: Percent Persons 35-44 Years						0.086		
BAGE54LN	B: Linear: Percent Persons 45-54 Years					-0.436		-0.076	
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years					-0.077		0.225	
BAGE54CU	B: Cubic: Percent Persons 45-54 Years					0.210			
BAGE54QR	B: Quartic: Percent Persons 45-54 Years					-0.066			
IFEM275	Female Interaction Of BAGE54LN					0.439			
IFEM276	Female Interaction Of BAGE54QU					0.334			
IFEM277	Female Interaction Of BAGE54CU					0.001			
IFEM278	Female Interaction Of BAGE54QR					0.219			
BAGE64LN	B: Linear: Percent Persons 55-64 Years				-0.223				0.295
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years								-0.046
BAGE64CU	B: Cubic: Percent Persons 55-64 Years								0.162
BAGE64QR	B: Quartic: Percent Persons 55-64 Years								0.053
IBLK280	Black Interaction Of BAGE64LN				0.817				-0.567
IBLK282	Black Interaction Of BAGE64CU								0.128
IBLK283	Black Interaction Of BAGE64QR								-0.292
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander	0.242				-0.029			0.151
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander	0.177				0.252			0.466
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander	-0.121				-0.324			
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander	0.096				-0.198			
IFEM285	Female Interaction Of BASIANLN	-0.474				-0.250			
IFEM286	Female Interaction Of BASIANQU	-0.248				-0.104			
IFEM287	Female Interaction Of BASIANCU	-0.010				0.071			
IFEM288	Female Interaction Of BASIANQR	-0.210				0.402			
IHISP285	Hispanic Interaction Of BASIANLN								-0.555
IHISP286	Hispanic Interaction Of BASIANQU								-0.735
BCUBANLN	B: Linear: Percent Hispanics: Cuban				0.251				-0.198
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban								-0.443
IHISP290	Hispanic Interaction Of BCUBANLN				-1.268				-0.680
IHISP291	Hispanic Interaction Of BCUBANQU								1.153
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed				0.238	-0.048			0.092
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed					-0.054			-0.006
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed					-0.396			0.135
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed								0.125

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut			0.045	0.001			-0.165	0.624
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			-0.052	-0.167			-0.155	-0.291
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut				-0.112				0.509
BINDIAQR	B: Quartic: Percent Pop: American Indian, Eskimo, Aleut				-0.123				
IHISP300	Hispanic Interaction Of BINDIALN			0.070				-0.014	
IHISP301	Hispanic Interaction Of BINDIAQU			0.395				0.614	
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed	0.082				-0.041			
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed	-0.209				-0.175			
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed					0.221			
BMNOTQR	B: Quartic: Percent Males Separated, Divorced or Widowed					-0.078			
IFEM305	Female Interaction Of BMNOTLN	-0.304				0.115			
IFEM306	Female Interaction Of BMNOTQU	0.447				0.231			
IFEM307	Female Interaction Of BMNOTCU					-0.008			
IFEM308	Female Interaction Of BMNOTQR					0.219			
BPOVERLN	B: Linear: Percent Families Below Poverty Level	0.209						0.049	
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level	0.070						-0.024	
BPOVERCU	B: Cubic: Percent Families Below Poverty Level	-0.022						-0.236	
BPOVERQR	B: Quartic: Percent Families Below Poverty Level	-0.088							
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican							-0.128	0.100
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican							0.098	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican							-0.171	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican							-0.082	
IHISP315	Hispanic Interaction Of BPRICALN							0.033	-0.866
IHISP316	Hispanic Interaction Of BPRICAQU							-0.564	
IHISP317	Hispanic Interaction Of BPRICACU							0.464	
IHISP318	Hispanic Interaction Of BPRICAQR							0.201	
BSCHASLN	B: Linear: Percent Associates Degree				0.214				0.046
BSCHASQU	B: Quadratic: Percent Associates Degree				-0.082				0.324
BSCHASCU	B: Cubic: Percent Associates Degree				-0.092				-0.289
BSCHASQR	B: Quartic: Percent Associates Degree				0.090				-0.143
IBLK320	Black Interaction Of BSCHASLN				-1.164				-0.180
IBLK321	Black Interaction Of BSCHASQU								-1.163
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander				0.284			-0.276	-0.328
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander							-0.235	-0.431
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander								-0.076
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander								0.140
IFEM325	Female Interaction Of PASIANLN				-0.664			0.372	0.856
IFEM326	Female Interaction Of PASIANQU							0.455	
IHISP325	Hispanic Interaction Of PASIANLN							0.625	
PCUBANLN	T: Linear: Percent Hispanics: Cuban						0.068	0.426	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban						0.395	0.195	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban						-0.119	-0.068	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban						-0.102	0.135	
IFEM330	Female Interaction Of PCUBANLN						-0.160		
IFEM331	Female Interaction Of PCUBANQU						-0.334		
IFEM332	Female Interaction Of PCUBANCU						-0.150		
IFEM333	Female Interaction Of PCUBANQR						0.236		
IBLK330	Black Interaction Of PCUBANLN							-0.435	
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut			0.173				0.213	
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			0.100					
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut			-0.228					
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut			0.121					

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E7. Coefficients of Model Parameters for Past Year Illicit Drug Use Treatment

		BIG CITY				REMAINDER			
Variable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	-5.043	-4.773	-4.581	-6.415	-5.359	-4.595	-4.979	-7.462
FEMALE	O: Female Indicator	0.860	-0.370	-0.419	-0.020	0.745	-0.472	-0.527	-1.312
FEMBLCK	O: Black Interaction Of FEMALE	0.387	0.089	-0.072	-1.335	-3.185	-0.594	-0.410	0.329
FEMHISP	O: Hispanic Interaction Of FEMALE	-1.317	-1.305	-0.844	-0.362	-0.910	-0.367	-0.679	-0.798
FEMOTHR	O: Other Interaction Of FEMALE							-1.065	
RACEBLCK	O: Race/Black Indicator	-3.492	-0.534	0.334	1.470	-0.932	-1.037	0.498	0.976
RACEHISP	O: Race/Hispanic Indicator	0.091	-0.029	-2.206	-0.778	-0.163	-1.958	-0.944	0.648
RACEOTHR	O: Race/Other Indicator		-1.761			-2.205		0.171	1.545
REGNOREA	O: Northeast Region Indicator					-1.227			
REGSOUTH	O: South Region Indicator					-0.207			
IHISP7	Hispanic Interaction Of REGNOREA			2.957					
IHISP8	Hispanic Interaction Of REGSOUTH			2.157					
IHISP9	Hispanic Interaction Of REGWEST			1.724					
PHH1PLN	T: Linear: Percent One Person Households							0.262	0.202
PHH1PQU	T: Quadratic: Percent One Person Households							0.149	-0.042
PHH1PCU	T: Cubic: Percent One Person Households							-0.118	-0.090
PHH1PQR	T: Quartic: Percent One Person Households							0.085	-0.098
POPRMLN	T: Linear: Average Persons Per Room							0.347	
PAGE34LN	T: Linear: Percent Persons 25-34 Years					-0.111			
IFEM35	Female Interaction Of PAGE34LN					0.661			
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma				-0.670				-0.233
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma								0.284
PSCHSCLN	T: Linear: Percent Some College And No Degree		0.294				-0.152		0.320
PSCHSCQU	T: Quadratic: Percent Some College And No Degree		-0.123				-0.147		-0.381
IFEM80	Female Interaction Of PSCHSCLN		-0.355						-0.277
IFEM81	Female Interaction Of PSCHSCQU		0.327						0.521
IHISP80	Hispanic Interaction Of PSCHSCLN						0.381		
IHISP81	Hispanic Interaction Of PSCHSCQU						0.658		
PPOVERLN	T: Linear: Percent Families Below Poverty Level				0.807				0.806
PBLACKLN	T: Linear: Percent Black Nonhispanic							-0.214	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic							-0.170	
PHISPLN	T: Linear: Percent Hispanic			0.003				0.037	
PHISPQU	T: Quadratic: Percent Hispanic			0.116				0.173	
PHISPCU	T: Cubic: Percent Hispanic			-0.172				0.241	
PHISPQR	T: Quartic: Percent Hispanic			0.062				0.058	
POTHLN	T: Linear: Percent Other Race/Hispanicity						-0.075		
POTHQU	T: Quadratic: Percent Other Race/Hispanicity						-0.042		
POTHCU	T: Cubic: Percent Other Race/Hispanicity						-0.165		
IHISP110	Hispanic Interaction Of POTHLN						-0.156		
IHISP111	Hispanic Interaction Of POTHQU						0.614		
IHISP112	Hispanic Interaction Of POTHCU						0.498		
PHHF18LN	T: Linear: % Female-Headed HH W/No Spouse & Chld Under 18					0.136			
PHHF18QU	T: Quadratic: % Female-Headed HH W/No Spouse & Chld Under 18					-0.035			
PHHF18CU	T: Cubic: % Female-Headed HH W/No Spouse & Chld Under 18					-0.150			
PHHF18QR	T: Quartic: % Female-Headed HH W/No Spouse & Chld Under 18					-0.105			
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force	0.029				-0.196			
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force	-0.204				0.050			
PFLABCU	T: Cubic: Percent Females 16+ Years Old In Labor Force					0.027			
PFLABQR	T: Quartic: Percent Females 16+ Years Old In Labor Force					0.140			
PFNEVLN	T: Linear: Percent Females Never Married			0.203				-0.032	
PFNEVQU	T: Quadratic: Percent Females Never Married							0.074	
PFNEVCU	T: Cubic: Percent Females Never Married							-0.052	
PFNEVQR	T: Quartic: Percent Females Never Married							-0.089	
PMNEVLN	T: Linear: Percent Males Never Married						0.263		0.115
PMNEVQU	T: Quadratic: Percent Males Never Married								0.354
IBLK140	Black Interaction Of PMNEVLN						-0.707		
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force								0.075
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force								0.042
PMLABCU	T: Cubic: Percent Males 16+ Years Old In Labor Force								-0.091
PMLABQR	T: Quartic: Percent Males 16+ Years Old In Labor Force								-0.136
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases						-0.032		
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases						0.245		
V18FLN	C: Linear: Marijuana Posession Arrest Rate				-0.856		0.072		-0.829
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate						0.159		-0.304
V18FCU	C: Cubic: Marijuana Posession Arrest Rate						-0.088		-0.300
V18FQR	C: Quartic: Marijuana Posession Arrest Rate						-0.081		
IHISP185	Hispanic Interaction Of V18FLN								0.695
IHISP186	Hispanic Interaction Of V18FQU								0.108
IHISP187	Hispanic Interaction Of V18FCU								0.932
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate		0.239				-0.484		
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate						-0.322		
IBLK190	Black Interaction Of V18BLN						-0.003		
IBLK191	Black Interaction Of V18BQU						0.491		
IHISP190	Hispanic Interaction Of V18BLN		-0.724				0.426		
IHISP191	Hispanic Interaction Of V18BQU						0.590		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate						0.380		
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate				0.774				0.484
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992	-0.158							
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992	0.207							
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992	-0.332							
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability			0.182	0.022			0.484	
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability			0.047	0.358			-0.135	
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability			0.015				0.009	
B64DISQR	B: Quartic: Percent Persons 16-64 With A Work Disability			0.065				-0.057	
IHISP260	Hispanic Interaction Of B64DISLN							-0.142	
IHISP261	Hispanic Interaction Of B64DISQU							0.099	
IHISP262	Hispanic Interaction Of B64DISCU							0.681	
IHISP263	Hispanic Interaction Of B64DISQR							0.183	
BAGE34LN	B: Linear: Percent Persons 25-34 Years					-0.080			
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years					-0.021			
BAGE34CU	B: Cubic: Percent Persons 25-34 Years					-0.007			
BAGE44LN	B: Linear: Percent Persons 35-44 Years			-0.119		-0.074	0.063	-0.014	0.135
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years			0.015		-0.037	0.030	-0.080	-0.189
BAGE44CU	B: Cubic: Percent Persons 35-44 Years			0.159		0.196	-0.085	-0.109	0.473
BAGE44QR	B: Quartic: Percent Persons 35-44 Years						-0.044	0.062	
IFEM270	Female Interaction Of BAGE44LN						-0.008		
IFEM271	Female Interaction Of BAGE44QU						-0.115		
IFEM272	Female Interaction Of BAGE44CU						0.381		
IFEM273	Female Interaction Of BAGE44QR						0.117		
BAGE54LN	B: Linear: Percent Persons 45-54 Years	-0.233				0.193	0.207		
IFEM275	Female Interaction Of BAGE54LN						-0.499		
IBLK275	Black Interaction Of BAGE54LN	1.518							
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander		-0.269	0.288			-0.235		
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander			-0.202			0.039		
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander						-0.092		
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander						-0.045		
IFEM285	Female Interaction Of BASIANLN						0.226		
IFEM286	Female Interaction Of BASIANQU						-0.276		
IFEM287	Female Interaction Of BASIANCU						0.077		
IFEM288	Female Interaction Of BASIANQR						0.201		
IBLK285	Black Interaction Of BASIANLN		0.470				1.131		
BCUBANLN	B: Linear: Percent Hispanics: Cuban	0.419					-0.352		
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban	-0.071					-0.140		
BCUBANCU	B: Cubic: Percent Hispanics: Cuban	0.479							
IFEM290	Female Interaction Of BCUBANLN						-0.175		
IFEM291	Female Interaction Of BCUBANQU						0.568		
IBLK290	Black Interaction Of BCUBANLN	-1.371							
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed					0.318	0.095	-0.187	-0.325
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed						0.009	-0.019	0.216
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed						0.211	0.057	
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed						0.072	0.062	
IFEM295	Female Interaction Of BFNOTLN					-0.419			0.193
IFEM296	Female Interaction Of BFNOTQU								-0.818
IBLK295	Black Interaction Of BFNOTLN						-0.107		
IBLK297	Black Interaction Of BFNOTCU						-0.225		
IBLK298	Black Interaction Of BFNOTQR						-0.189		
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut		0.010		-0.569		0.358		0.304
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut		0.097		-0.540		-0.064		0.642
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut		0.258		-0.557		0.294		-1.237
BINDIAQR	B: Quartic: Percent Pop: American Indian, Eskimo, Aleut		-0.092				-0.103		0.308
IBLK300	Black Interaction Of BINDIALN		0.487		0.668		-0.004		0.253
IBLK301	Black Interaction Of BINDIAQU		-0.122		0.686		0.419		-0.353
IBLK302	Black Interaction Of BINDIACU		0.226		1.054		0.085		0.879
IBLK303	Black Interaction Of BINDIAQR		0.139				0.246		-0.373
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed								0.460
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.053						-0.488
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.132						-0.336
IOTH311	Other Interaction Of BPOVERQU		-0.470						
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican		-0.354				0.251		0.530
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican						0.104		
IHISP315	Hispanic Interaction Of BPRICALN						0.588		
IHISP316	Hispanic Interaction Of BPRICAQU						-1.177		
BSCHASLN	B: Linear: Percent Associates Degree							0.148	
BSCHASQU	B: Quadratic: Percent Associates Degree							-0.063	
BSCHASCU	B: Cubic: Percent Associates Degree							0.150	
IBLK320	Black Interaction Of BSCHASLN							-0.633	
IBLK321	Black Interaction Of BSCHASQU							-0.272	
IBLK322	Black Interaction Of BSCHASCU							-0.648	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander					0.122		0.016	
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander							0.131	
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander							-0.037	
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander							0.082	
PCUBANLN	T: Linear: Percent Hispanics: Cuban						-0.011		
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban						0.169		
PCUBANCU	T: Cubic: Percent Hispanics: Cuban						0.024		
PCUBANQR	T: Quartic: Percent Hispanics: Cuban						-0.083		
IFEM330	Female Interaction Of PCUBANLN						0.682		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IFEM331	Female Interaction Of PCUBANQU						-0.233		
IFEM332	Female Interaction Of PCUBANCU						-0.096		
IFEM333	Female Interaction Of PCUBANQR						0.242		
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut							0.254	
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut							-0.115	
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut							0.301	
IBLK335	Black Interaction Of PINDIALN							-0.108	
IBLK336	Black Interaction Of PINDIAQU							0.093	
IBLK337	Black Interaction Of PINDIACU							-0.494	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E8. Coefficients of Model Parameters for Past Year Dependency on Alcohol Only

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
DUMMY	O: Intercept Term	-3.963	-2.644	-3.080	-3.729	-4.195	-2.365	-2.898	-4.946
FEMALE	O: Female Indicator	-0.383	-0.366	-0.418	-0.598	0.113	-0.845	-1.140	-1.297
FEMBLCK	O: Black Interaction Of FEMALE	-1.158	-0.018	-0.071	-0.648	-0.850	-0.065	0.021	-0.187
FEMHISP	O: Hispanic Interaction Of FEMALE	0.022	-1.116	-0.961	-1.060	0.122	-0.073	-0.427	-0.924
FEMOTHR	O: Other Interaction Of FEMALE	0.373	-0.253	-0.200	-1.735		2.041	-0.388	
RACEBLCK	O: Race/Black Indicator	-0.176	-0.939	-0.428	0.542	-2.878	-0.844	0.024	-2.764
RACEHISP	O: Race/Hispanic Indicator	-0.581	0.066	0.186	0.017	0.067	-0.229	0.449	0.312
RACEOTHR	O: Race/Other Indicator	-1.575	-0.675	-2.097	-3.071		-1.577	-0.957	-0.134
REGNOREA	O: Northeast Region Indicator					-0.118			
REGSOUTH	O: South Region Indicator					-0.211			
REGWEST	O: West Region Indicator					0.225			
IBLK7	Black Interaction Of REGNOREA					1.705			
IBLK8	Black Interaction Of REGSOUTH					1.512			
IBLK9	Black Interaction Of REGWEST					2.175			
PDENLEV1	O: Large MSA					-0.530			0.498
PDENLEV2	O: Medium MSA					-0.419			0.074
PDENLEV3	O: Small MSA					0.076			0.648
PDENLEV4	O: NonMSA, Urban					0.174			0.902
IBLK10	Black Interaction Of PDENLEV1					0.911			3.298
IBLK11	Black Interaction Of PDENLEV2					0.208			3.914
IBLK12	Black Interaction Of PDENLEV3					2.387			1.964
IBLK13	Black Interaction Of PDENLEV4					1.470			1.283
PHH1PLN	T: Linear: Percent One Person Households								0.112
PHH1PQU	T: Quadratic: Percent One Person Households								-0.153
IFEM15	Female Interaction Of PHH1PLN								0.033
IFEM16	Female Interaction Of PHH1PQU								0.288
POPRMLN	T: Linear: Average Persons Per Room		-0.194				-0.223		
POPRMQU	T: Quadratic: Average Persons Per Room						-0.141		
POPRMCU	T: Cubic: Average Persons Per Room						-0.141		
IBLK20	Black Interaction Of POPRMLN						-0.045		
IBLK21	Black Interaction Of POPRMQU						0.238		
PAGE44LN	T: Linear: Percent Persons 35-44 Years					-0.004			
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years					0.111			
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma				0.003				-0.365
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma				0.137				0.069
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma								-0.138
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree								-0.436
PSCHSCLN	T: Linear: Percent Some College And No Degree			0.048				0.046	
PSCHSCQU	T: Quadratic: Percent Some College And No Degree			-0.005				0.089	
PSCHSCCU	T: Cubic: Percent Some College And No Degree			0.118				-0.008	
IFEM80	Female Interaction Of PSCHSCLN							0.061	
IFEM81	Female Interaction Of PSCHSCQU							-0.023	
IFEM82	Female Interaction Of PSCHSCCU							0.260	
PPOVERLN	T: Linear: Percent Families Below Poverty Level					0.013		0.184	
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level					-0.058		-0.104	
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					-0.079		-0.027	
PPOVERQR	T: Quartic: Percent Families Below Poverty Level					0.001		-0.051	
IFEM85	Female Interaction Of PPOVERLN					0.040			
IFEM86	Female Interaction Of PPOVERQU					0.194			
IFEM87	Female Interaction Of PPOVERCU					-0.072			
IFEM88	Female Interaction Of PPOVERQR					-0.104			
IHISP85	Hispanic Interaction Of PPOVERLN					-0.398		0.042	
IHISP86	Hispanic Interaction Of PPOVERQU					-0.362		0.249	
P64DISLN	T: Linear: Percent 16-64 With A Work Disability		-0.092				-0.161		
P64DISQU	T: Quadratic: Percent 16-64 With A Work Disability		0.033				0.051		
P64DISCU	T: Cubic: Percent 16-64 With A Work Disability		-0.033				-0.036		
P64DISQR	T: Quartic: Percent 16-64 With A Work Disability		-0.010				0.046		
IBLK95	Black Interaction Of P64DISLN		-0.122				0.266		
IBLK96	Black Interaction Of P64DISQU		0.272				0.099		
IBLK97	Black Interaction Of P64DISCU		-0.006				-0.204		
IBLK98	Black Interaction Of P64DISQR		-0.133				-0.146		
PHISPLN	T: Linear: Percent Hispanic								-0.233
PHISPQU	T: Quadratic: Percent Hispanic								-0.315
IBLK105	Black Interaction Of PHISPLN								0.094
IBLK106	Black Interaction Of PHISPQU								0.356
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed						0.033		
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed						-0.020		
IOTH135	Other Interaction Of PFNOTLN						0.125		
IOTH136	Other Interaction Of PFNOTQU						0.445		
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed								-0.079
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed								-0.072
PMNOTCU	T: Cubic: Percent Males Separated, Divorced Or Widowed								-0.189
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier								0.046
POLDHUQU	T: Quadratic: Percent Housing Units Built 1939 Or Earlier								0.015
POLDHUCU	T: Cubic: Percent Housing Units Built 1939 Or Earlier								0.181
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol	-0.113				-0.101			
ADRAT1QU	C: Quadratic: Death Rate With Explicit Mention Of Alcohol	-0.029				0.006			
ADRAT1CU	C: Cubic: Death Rate With Explicit Mention Of Alcohol	-0.067				0.010			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
ADRATIQR	C: Quartic: Death Rate With Explicit Mention Of Alcohol	-0.098				0.066			
V18FLN	C: Linear: Marijuana Posession Arrest Rate					-0.028			
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate					0.097			
IFEM185	Female Interaction Of V18FLN					0.011			
IFEM186	Female Interaction Of V18FQU					-0.283			
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate			0.363	-0.013		0.117	0.108	0.034
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate			0.253	0.192			0.112	0.131
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate				0.218				0.327
IHISP195	Hispanic Interaction Of V18ELN								0.255
IHISP196	Hispanic Interaction Of V18EQU								-0.345
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate			0.121				0.019	0.051
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate			-0.060				-0.097	-0.155
IHISP200	Hispanic Interaction Of V18ALN								-0.349
IHISP201	Hispanic Interaction Of V18AQU								0.356
IOTH200	Other Interaction Of V18ALN			1.082					
IOTH201	Other Interaction Of V18AQU			-0.680					
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate	0.156		-0.455	0.027	-0.259		-0.218	0.058
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate			-0.125	0.006	-0.131		-0.112	-0.156
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate				-0.080	-0.114			-0.271
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate					-0.088			
IBLK225	Black Interaction Of V18LN								-0.137
IBLK226	Black Interaction Of V18QU								0.371
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992	0.193				0.162			
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992	0.137				-0.133			
DRATECU	C: Cubic: Mean Drug Client Treatment Rate 1991 & 1992	0.026							
DRATEQR	C: Quartic: Mean Drug Client Treatment Rate 1991 & 1992	0.097							
IBLK235	Black Interaction Of DRATELN	-0.113							
IBLK236	Black Interaction Of DRATEQU	-0.300							
IBLK237	Black Interaction Of DRATECU	0.460							
IBLK238	Black Interaction Of DRATEQR	-0.159							
RP80ALN	T: Linear: Recoded Median Household Income		0.057				-0.150		
RP80AQU	T: Quadratic: Recoded Median Household Income		0.049				-0.014		
RP80ACU	T: Cubic: Recoded Median Household Income		-0.028				-0.010		
RP80AQR	T: Quartic: Recoded Median Household Income		-0.007				-0.038		
IBLK250	Black Interaction Of RP80ALN		-0.057				0.056		
IBLK251	Black Interaction Of RP80AQU		-0.137				0.298		
IBLK252	Black Interaction Of RP80ACU		0.134				0.393		
IBLK253	Black Interaction Of RP80AQR		0.085				0.087		
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992					-0.225			
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability		0.094				-0.008		
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability		0.009				-0.018		
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability		0.019				0.050		
B64DISQR	B: Quartic: Percent Persons 16-64 With A Work Disability		-0.005						
IHISP260	Hispanic Interaction Of B64DISLN		-0.097						
IHISP261	Hispanic Interaction Of B64DISQU		-0.015						
IHISP262	Hispanic Interaction Of B64DISCU		-0.075						
IHISP263	Hispanic Interaction Of B64DISQR		0.059						
BAGE34LN	B: Linear: Percent Persons 25-34 Years		0.037				0.018		0.067
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years		-0.017				-0.066		-0.064
BAGE34CU	B: Cubic: Percent Persons 25-34 Years						0.109		0.050
BAGE34QR	B: Quartic: Percent Persons 25-34 Years								-0.080
IFEM265	Female Interaction Of BAGE34LN		-0.166						
IFEM266	Female Interaction Of BAGE34QU		0.166						
BAGE44LN	B: Linear: Percent Persons 35-44 Years			0.124		0.030		0.045	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years			-0.025		0.067		0.054	
BAGE44CU	B: Cubic: Percent Persons 35-44 Years			0.080		-0.140		-0.037	
BAGE44QR	B: Quartic: Percent Persons 35-44 Years							-0.037	
IOTH270	Other Interaction Of BAGE44LN			-0.168				0.511	
IOTH271	Other Interaction Of BAGE44QU			-0.286					
BAGE54LN	B: Linear: Percent Persons 45-54 Years						-0.023		
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years						0.053		
IOTH275	Other Interaction Of BAGE54LN						-0.121		
IOTH276	Other Interaction Of BAGE54QU						-0.374		
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander			-0.009	-0.064			-0.031	
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander				-0.123			-0.019	
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander				-0.001			-0.010	
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander				-0.035			-0.033	
IHISP285	Hispanic Interaction Of BASIANLN			-0.277				-0.086	
IHISP286	Hispanic Interaction Of BASIANQU							0.136	
IHISP287	Hispanic Interaction Of BASIANCU							0.122	
IHISP288	Hispanic Interaction Of BASIANQR							0.116	
IOTH285	Other Interaction Of BASIANLN				0.164				
IOTH286	Other Interaction Of BASIANQU				1.819				
IOTH287	Other Interaction Of BASIANCU				-3.294				
IOTH288	Other Interaction Of BASIANQR				0.795				
BCUBANLN	B: Linear: Percent Hispanics: Cuban		0.212	0.171			-0.096		
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban		-0.153	0.041			0.085		
BCUBANCU	B: Cubic: Percent Hispanics: Cuban						-0.088		
BCUBANQR	B: Quartic: Percent Hispanics: Cuban						0.135		
IFEM290	Female Interaction Of BCUBANLN		-0.320				0.073		
IFEM291	Female Interaction Of BCUBANQU		0.211				-0.082		
IFEM292	Female Interaction Of BCUBANCU						0.422		
IFEM293	Female Interaction Of BCUBANQR						-0.216		
IBLK290	Black Interaction Of BCUBANLN			-0.104					
IBLK291	Black Interaction Of BCUBANQU			0.252					

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed							0.074	-0.227
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed							-0.002	-0.015
IFEM295	Female Interaction Of BFNOTLN								0.097
IFEM296	Female Interaction Of BFNOTQU								0.236
IOTH295	Other Interaction Of BFNOTLN							0.417	
IOTH296	Other Interaction Of BFNOTQU							0.721	
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut			-0.214		-0.053		-0.093	0.069
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut							0.048	0.220
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut							0.087	
IBLK300	Black Interaction Of BINDIALN			0.418		0.916		0.122	
IBLK301	Black Interaction Of BINDIAQU							-0.070	
IBLK302	Black Interaction Of BINDIACU							-0.442	
IHISP300	Hispanic Interaction Of BINDIALN			0.309				0.160	
IHISP301	Hispanic Interaction Of BINDIAQU							-0.268	
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed			0.085	-0.075	-0.043	-0.052	-0.053	-0.059
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed				0.097	0.091	0.021	-0.094	-0.064
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed				-0.004	0.009			-0.175
BMNOTQR	B: Quartic: Percent Males Separated, Divorced or Widowed				-0.045	-0.050			0.006
IBLK305	Black Interaction Of BMNOTLN				0.200				0.644
IBLK306	Black Interaction Of BMNOTQU				-0.200				0.040
IBLK307	Black Interaction Of BMNOTCU				0.308				0.312
IBLK308	Black Interaction Of BMNOTQR								-0.183
IHISP305	Hispanic Interaction Of BMNOTLN					0.299			
IHISP306	Hispanic Interaction Of BMNOTQU					-0.128			
IHISP307	Hispanic Interaction Of BMNOTCU					-0.048			
IHISP308	Hispanic Interaction Of BMNOTQR					0.169			
BPOVERLN	B: Linear: Percent Families Below Poverty Level				-0.069				0.093
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level				-0.061				
BPOVERCU	B: Cubic: Percent Families Below Poverty Level				0.118				
IHISP310	Hispanic Interaction Of BPOVERLN								0.408
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican		-0.229	0.089			-0.173	-0.112	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican			-0.040				-0.060	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican			-0.140				-0.062	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican							-0.079	
IBLK315	Black Interaction Of BPRICALN							-0.265	
IBLK316	Black Interaction Of BPRICAQU							0.090	
IBLK317	Black Interaction Of BPRICACU							-0.070	
IBLK318	Black Interaction Of BPRICAQR							0.192	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander	-0.019				0.078	-0.039		-0.144
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander	0.157				0.147	-0.032		-0.048
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander	-0.145				0.075	-0.076		-0.204
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander					-0.041			
IFEM325	Female Interaction Of PASIANLN	-0.009				-0.287			
IFEM326	Female Interaction Of PASIANQU	-0.242				-0.119			
IFEM327	Female Interaction Of PASIANCU	0.249				0.033			
IFEM328	Female Interaction Of PASIANQR					0.137			
IBLK325	Black Interaction Of PASIANLN						0.151		
IBLK326	Black Interaction Of PASIANQU						0.072		
IBLK327	Black Interaction Of PASIANCU						0.415		
PCUBANLN	T: Linear: Percent Hispanics: Cuban	-0.057	-0.185		0.036	0.032	0.069		-0.142
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban				-0.065		0.006		-0.037
PCUBANCU	T: Cubic: Percent Hispanics: Cuban								-0.129
IBLK330	Black Interaction Of PCUBANLN	0.519			0.197	0.696			0.277
IBLK331	Black Interaction Of PCUBANQU				0.296				-0.301
IBLK332	Black Interaction Of PCUBANCU								0.674
IOTH330	Other Interaction Of PCUBANLN						0.535		
IOTH331	Other Interaction Of PCUBANQU						0.515		
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut	0.055			-0.179	-0.064			
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut	-0.062			0.114	0.159			
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut	0.015			-0.057	0.154			
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut				0.042				
IFEM335	Female Interaction Of PINDIALN	0.134				0.007			
IFEM336	Female Interaction Of PINDIAQU	0.168				-0.264			
IFEM337	Female Interaction Of PINDIACU	-0.275				-0.319			
IHISP335	Hispanic Interaction Of PINDIALN				0.282				
IHISP336	Hispanic Interaction Of PINDIAQU				-0.091				
IHISP337	Hispanic Interaction Of PINDIACU				0.151				
IHISP338	Hispanic Interaction Of PINDIAQR				-0.115				
IOTH335	Other Interaction Of PINDIALN				0.359				
IOTH336	Other Interaction Of PINDIAQU				-0.987				

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E9. Coefficients of Model Parameters for Past Year Illicit Dependency

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	-4.381	-3.268	-4.208	-5.485	-4.057	-3.268	-4.618	-6.476
FEMALE	O: Female Indicator	0.449	-0.684	-0.706	-0.607	-0.320	-0.878	-0.774	-0.071
FEMBLCK	O: Black Interaction Of FEMALE	-0.149	0.076	0.230	-0.177	0.859	0.533	-0.178	-0.466
FEMHISP	O: Hispanic Interaction Of FEMALE	-0.388	0.306	0.913	-0.262	0.637	0.240	0.094	-0.306
RACEBLCK	O: Race/Black Indicator	-1.289	-0.503	0.181	0.696	-1.851	-0.864	0.220	0.884
RACEHISP	O: Race/Hispanic Indicator	0.150	-1.149	-1.313	-0.980	-0.302	-0.585	-0.778	-0.234
RACEOTHR	O: Race/Other Indicator		-1.719				-1.661		
REGNOREA	O: Northeast Region Indicator							0.349	
REGSOUTH	O: South Region Indicator							0.286	
REGWEST	O: West Region Indicator							0.840	
PHH1PLN	T: Linear: Percent One Person Households						0.258		
POPRMLN	T: Linear: Average Persons Per Room			-0.156				-0.232	
POPRMQU	T: Quadratic: Average Persons Per Room			-0.130				0.040	
POPRMCU	T: Cubic: Average Persons Per Room							0.082	
POPRMQR	T: Quartic: Average Persons Per Room							0.074	
PAGE18LN	T: Linear: Percent Persons 0-18 Years		0.051				0.248		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years		-0.067				0.108		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years		-0.072				0.078		
PAGE18QR	T: Quartic: Percent Persons 0-18 Years		0.044				0.067		
IFEM25	Female Interaction Of PAGE18LN		-0.201				-0.128		
IFEM26	Female Interaction Of PAGE18QU		0.084				0.036		
IFEM27	Female Interaction Of PAGE18CU		0.158				-0.085		
IFEM28	Female Interaction Of PAGE18QR		-0.135				-0.096		
PAGE34LN	T: Linear: Percent Persons 25-34 Years		0.072			-0.183		0.127	
PAGE34QU	T: Quadratic: Percent Persons 25-34 Years					-0.202		-0.187	
IBLK35	Black Interaction Of PAGE34LN							-0.691	
IHISP35	Hispanic Interaction Of PAGE34LN		-0.416						
PAGE44LN	T: Linear: Percent Persons 35-44 Years								0.381
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years								0.267
PAGE54LN	T: Linear: Percent Persons 45-54 Years				-0.073	0.057			
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years				0.146	0.040			
PAGE54CU	T: Cubic: Percent Persons 45-54 Years				-0.299	-0.158			
IHISP45	Hispanic Interaction Of PAGE54LN				-0.761				
IHISP46	Hispanic Interaction Of PAGE54QU				-0.915				
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma	-0.109							
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma	0.111							
IFEM60	Female Interaction Of PSCH12LN	0.011							
IFEM61	Female Interaction Of PSCH12QU	-0.248							
PPOVERLN	T: Linear: Percent Families Below Poverty Level				-0.112				
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level				-0.011				
PPOVERCU	T: Cubic: Percent Families Below Poverty Level				-0.220				
PMNEVLN	T: Linear: Percent Males Never Married			0.159					
PMNEVQU	T: Quadratic: Percent Males Never Married			-0.057					
PMNEVCU	T: Cubic: Percent Males Never Married			0.141					
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force		0.140						
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force		-0.073						
PMLABCU	T: Cubic: Percent Males 16+ Years Old In Labor Force		0.145						
IHISP145	Hispanic Interaction Of PMLABLN		-0.275						
IHISP146	Hispanic Interaction Of PMLABQU		-0.006						
IHISP147	Hispanic Interaction Of PMLABCU		-0.492						
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed								0.033
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed								0.184
PRENTLN	T: Linear: Percent Housing Rented							0.178	0.617
PRENTQU	T: Quadratic: Percent Housing Rented							0.164	
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases							0.062	
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases							0.193	
V18FLN	C: Linear: Marijuana Posession Arrest Rate			-0.374					
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate			0.560				-0.056	
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate			-0.143				-0.078	
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate			0.113				-0.070	
V18EQR	C: Quartic: Opium/Cocaine & Deriv Posession Arrest Rate			0.059				-0.068	
IHISP195	Hispanic Interaction Of V18ELN			-0.211					
IHISP196	Hispanic Interaction Of V18EQU			0.098					
IHISP197	Hispanic Interaction Of V18ECU			-0.420					
IHISP198	Hispanic Interaction Of V18EQR			0.187					
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate						0.013		
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate						-0.118		
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate								-0.271
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate								-0.238
VIOLCU	C: Cubic: Total Violent Offenses Arrest Rate								-0.404
RP80ALN	T: Linear: Recoded Median Household Income					-0.124			-0.540
RP80AQU	T: Quadratic: Recoded Median Household Income					0.259			-0.005
RP80ACU	T: Cubic: Recoded Median Household Income								-0.234
IFEM250	Female Interaction Of RP80ALN					-0.068			
IFEM251	Female Interaction Of RP80AQU					-0.224			
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992						0.118		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992						0.176		
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability			0.206					
BAGE34LN	B: Linear: Percent Persons 25-34 Years							-0.087	
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years							-0.070	
BAGE34CU	B: Cubic: Percent Persons 25-34 Years							0.090	
IFEM265	Female Interaction Of BAGE34LN							-0.066	
IFEM266	Female Interaction Of BAGE34QU							0.169	
IFEM267	Female Interaction Of BAGE34CU							-0.365	
BAGE44LN	B: Linear: Percent Persons 35-44 Years					-0.273		-0.023	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years							0.111	
IHISP270	Hispanic Interaction Of BAGE44LN							-0.032	
IHISP271	Hispanic Interaction Of BAGE44QU							-0.374	
BAGE54LN	B: Linear: Percent Persons 45-54 Years			-0.118		-0.079	0.132	-0.194	0.017
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years					-0.145	-0.083	-0.161	0.066
BAGE54CU	B: Cubic: Percent Persons 45-54 Years					0.110	-0.131	0.119	0.256
BAGE54QR	B: Quartic: Percent Persons 45-54 Years								0.114
IFEM275	Female Interaction Of BAGE54LN					-0.025		0.398	-0.334
IFEM276	Female Interaction Of BAGE54QU					0.295			-0.072
IFEM277	Female Interaction Of BAGE54CU								-0.491
IFEM278	Female Interaction Of BAGE54QR								-0.234
IBLK275	Black Interaction Of BAGE54LN							-0.331	
IBLK276	Black Interaction Of BAGE54QU							-0.018	
IBLK277	Black Interaction Of BAGE54CU							-0.313	
BAGE64LN	B: Linear: Percent Persons 55-64 Years			-0.055		-0.243	-0.009	0.051	
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years			-0.219			-0.128	-0.140	
BAGE64CU	B: Cubic: Percent Persons 55-64 Years			-0.118				-0.052	
IFEM280	Female Interaction Of BAGE64LN					0.406			
IBLK280	Black Interaction Of BAGE64LN			-0.010				-0.199	
IBLK281	Black Interaction Of BAGE64QU			0.326				0.076	
IBLK282	Black Interaction Of BAGE64CU			0.268				0.277	
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander							0.215	0.310
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander								0.056
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander								-0.270
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander								-0.011
IBLK285	Black Interaction Of BASIANLN								0.052
IBLK286	Black Interaction Of BASIANQU								0.235
IBLK287	Black Interaction Of BASIANCU								0.853
IBLK288	Black Interaction Of BASIANQR								-0.291
BCUBANLN	B: Linear: Percent Hispanics: Cuban				-0.198				
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban				0.300				
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed	0.223				0.119	-0.162		
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed	0.025				-0.103			
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed	-0.011				0.023			
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed	0.064				0.069			
IFEM295	Female Interaction Of BFNOTLN						0.304		
IBLK295	Black Interaction Of BFNOTLN					-0.139			
IBLK296	Black Interaction Of BFNOTQU					-0.072			
IBLK297	Black Interaction Of BFNOTCU					-0.475			
IHISP295	Hispanic Interaction Of BFNOTLN	-0.046							
IHISP296	Hispanic Interaction Of BFNOTQU	-0.084							
IHISP297	Hispanic Interaction Of BFNOTCU	0.048							
IHISP298	Hispanic Interaction Of BFNOTQR	-0.089							
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut			-0.153			0.082	0.222	0.533
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			0.071			-0.123	-0.164	
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut			0.230			0.076	0.063	
BINDIAQR	B: Quartic: Percent Pop: American Indian, Eskimo, Aleut						-0.095	0.048	
IBLK300	Black Interaction Of BINDIALN							-0.267	
IBLK301	Black Interaction Of BINDIAQU							0.346	
IBLK302	Black Interaction Of BINDIACU							-0.032	
IBLK303	Black Interaction Of BINDIAQR							-0.228	
IHISP300	Hispanic Interaction Of BINDIALN							-0.552	
IHISP301	Hispanic Interaction Of BINDIAQU							0.624	
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.145			-0.290	-0.040		-0.662
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.064				0.010		-0.342
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		0.017				-0.114		
IOTH310	Other Interaction Of BPOVERLN		-0.334						
IOTH311	Other Interaction Of BPOVERQU		0.406						
IOTH312	Other Interaction Of BPOVERCU		0.726						
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican					0.104	0.455		
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican					-0.306			
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					-0.041			
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican					-0.155			
IFEM315	Female Interaction Of BPRICALN					-0.625			
BSCHASLN	B: Linear: Percent Associates Degree				-0.083				
BSCHASQU	B: Quadratic: Percent Associates Degree				0.004				
BSCHASCU	B: Cubic: Percent Associates Degree				-0.308				
IHISP320	Hispanic Interaction Of BSCHASLN				-0.232				
IHISP321	Hispanic Interaction Of BSCHASQU				-0.188				
IHISP322	Hispanic Interaction Of BSCHASCU				0.568				
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander		0.133				0.241		
PCUBANLN	T: Linear: Percent Hispanics: Cuban	-0.068		0.057				-0.152	-0.483
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban	-0.011		0.189				0.173	0.293
PCUBANCU	T: Cubic: Percent Hispanics: Cuban							-0.114	0.183
PCUBANQR	T: Quartic: Percent Hispanics: Cuban							0.147	-0.178

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IFEM330	Female Interaction Of PCUBANLN	0.333							
IBLK330	Black Interaction Of PCUBANLN	0.103							-0.102
IBLK331	Black Interaction Of PCUBANQU	-0.800							-0.366
IBLK332	Black Interaction Of PCUBANCU								-0.242
IBLK333	Black Interaction Of PCUBANQR								0.480

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E10. Coefficients of Model Parameters for Past Year Arrested

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	-3.028	-3.002	-3.690	-4.853	-2.878	-2.294	-4.243	-5.665
FEMALE	O: Female Indicator	-1.145	-1.322	-0.960	-1.385	-1.128	-1.634	-1.813	-2.582
FEMBLCK	O: Black Interaction Of FEMALE	-0.180	-0.571	-0.083	-1.057	0.256	0.361	0.296	0.692
FEMHISP	O: Hispanic Interaction Of FEMALE	0.260	-1.209	-0.578	-0.624	0.246	-0.382	-0.074	-1.022
RACEBLCK	O: Race/Black Indicator	0.228	0.537	0.299	0.410	0.136	-0.746	0.333	0.006
RACEHISP	O: Race/Hispanic Indicator	0.185	0.050	0.249	0.099	0.369	-0.214	-0.315	0.722
RACEOTHR	O: Race/Other Indicator	-1.490	-1.138			-0.380	-0.543		
REGNOREA	O: Northeast Region Indicator					-0.208			-1.114
REGSOUTH	O: South Region Indicator					-0.434			-0.679
REGWEST	O: West Region Indicator					0.193			-0.372
PDENLEV1	O: Large MSA						-0.214	0.594	
PDENLEV2	O: Medium MSA						-0.162	0.676	
PDENLEV3	O: Small MSA						-0.005	1.159	
PDENLEV4	O: NonMSA, Urban						-0.154	1.026	
IBLK10	Black Interaction Of PDENLEV1						0.515		
IBLK11	Black Interaction Of PDENLEV2						0.547		
IBLK12	Black Interaction Of PDENLEV3						1.799		
IBLK13	Black Interaction Of PDENLEV4						0.850		
PHH1PLN	T: Linear: Percent One Person Households		-0.106				-0.088		
PHH1PQU	T: Quadratic: Percent One Person Households		0.032				0.049		
PHH1PCU	T: Cubic: Percent One Person Households		-0.020				-0.155		
PHH1PQR	T: Quartic: Percent One Person Households		-0.034				-0.039		
IBLK15	Black Interaction Of PHH1PLN		0.016				0.232		
IBLK16	Black Interaction Of PHH1PQU		0.009				0.043		
IBLK17	Black Interaction Of PHH1PCU		0.110				0.240		
IBLK18	Black Interaction Of PHH1PQR		0.089				0.076		
POPRMLN	T: Linear: Average Persons Per Room		-0.053				-0.143		
PAGE34LN	T: Linear: Percent Persons 25-34 Years								0.501
PAGE54LN	T: Linear: Percent Persons 45-54 Years							-0.063	
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years							0.045	
PAGE54CU	T: Cubic: Percent Persons 45-54 Years							-0.257	
IHISP45	Hispanic Interaction Of PAGE54LN							-0.196	
IHISP46	Hispanic Interaction Of PAGE54QU							-0.221	
IHISP47	Hispanic Interaction Of PAGE54CU							0.396	
PAGE64LN	T: Linear: Percent Persons 55-64 Years					0.205			
PSCH8LN	T: Linear: Percent 0-8 Years Of School	-0.100							
PSCH8QU	T: Quadratic: Percent 0-8 Years Of School	-0.017							
IOTH55	Other Interaction Of PSCH8LN	-0.661							
IOTH56	Other Interaction Of PSCH8QU	-0.773							
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma		0.187				0.157		
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma						0.061		
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma						-0.092		
IHISP60	Hispanic Interaction Of PSCH12LN						-0.351		
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree								0.222
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree								0.027
PSCHCOCU	T: Cubic: Bachelors, Graduate, Or Professional Degree								-0.315
IBLK75	Black Interaction Of PSCHCOLN								-0.227
IBLK76	Black Interaction Of PSCHCOQU								0.076
IBLK77	Black Interaction Of PSCHCOCU								0.608
PPUBASLN	T: Linear: Percent Households With Public Assist Income						0.199		
P64DISLN	T: Linear: Percent 16-64 With A Work Disability			0.200				0.261	
PBLACKLN	T: Linear: Percent Black Nonhispanic		-0.070						
PBLACKQU	T: Quadratic: Percent Black Nonhispanic		-0.094						
POTHLN	T: Linear: Percent Other Race/Hispanicity				0.071				
IHISP110	Hispanic Interaction Of POTHLN				0.287				
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force				0.284				0.066
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force								-0.333
IFEM125	Female Interaction Of PFLABLN				0.444				
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed							0.112	
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed							0.169	
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force				-0.038		0.137		-0.424
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force						-0.088		
IFEM145	Female Interaction Of PMLABLN				-0.636				
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed						-0.020		
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed						-0.052		
PMNOTCU	T: Cubic: Percent Males Separated, Divorced Or Widowed						-0.037		
PMNOTQR	T: Quartic: Percent Males Separated, Divorced Or Widowed						0.040		
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier	-0.145				-0.024			
POLDHUQU	T: Quadratic: Percent Housing Units Built 1939 Or Earlier					-0.027			
POLDHUCU	T: Cubic: Percent Housing Units Built 1939 Or Earlier					-0.119			
P40HULN	T: Linear: Percent Housing Units Built 1940-1949	0.201						-0.022	
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949							-0.152	
PRENTLN	T: Linear: Percent Housing Rented					0.362			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol		-0.012	-0.196					
IFEM180	Female Interaction Of ADRAT1LN		0.434						
V18FLN	C: Linear: Marijuana Posession Arrest Rate			-0.359				-0.201	
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate							-0.136	
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate			0.414	0.052			0.383	-0.015
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate				0.218				0.080
IFEM195	Female Interaction Of V18ELN								-0.540
IFEM196	Female Interaction Of V18EQU								-0.968
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate							-0.219	
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992	-0.091							
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992	-0.005							
DRATECU	C: Cubic: Mean Drug Client Treatment Rate 1991 & 1992	0.146							
RH43ALN	T: Linear: Recoded Median Rents For Rental Units				-0.254			-0.055	-0.700
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units							-0.031	
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units							0.053	
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units							0.056	
RH61ALN	T: Linear: Recoded Median Value Of Owner Occupied HUs		-0.007	-0.147			0.059	0.029	
RH61AQU	T: Quadratic: Recoded Median Value Of Owner Occupied HUs			-0.100			-0.061	-0.000	
RH61ACU	T: Cubic: Recoded Median Value Of Owner Occupied HUs						-0.099	0.132	
IHISP245	Hispanic Interaction Of RH61ALN							0.155	
IHISP246	Hispanic Interaction Of RH61AQU							0.166	
IHISP247	Hispanic Interaction Of RH61ACU							-0.282	
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992								0.010
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992								0.050
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992								-0.175
ARATEQR	C: Quartic: Mean A-Only Client Treatment Rate 1991 & 1992								-0.036
IBLK255	Black Interaction Of ARATELN								-0.037
IBLK256	Black Interaction Of ARATEQU								-0.264
IBLK257	Black Interaction Of ARATECU								0.089
IBLK258	Black Interaction Of ARATEQR								0.155
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability								0.086
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability								0.025
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability								0.085
B64DISQR	B: Quartic: Percent Persons 16-64 With A Work Disability								-0.076
BAGE34LN	B: Linear: Percent Persons 25-34 Years				0.128		-0.012		
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years				0.024		0.046		
BAGE34CU	B: Cubic: Percent Persons 25-34 Years						0.098		
BAGE34QR	B: Quartic: Percent Persons 25-34 Years						0.012		
IFEM265	Female Interaction Of BAGE34LN						0.114		
IFEM266	Female Interaction Of BAGE34QU						-0.132		
IFEM267	Female Interaction Of BAGE34CU						0.001		
IFEM268	Female Interaction Of BAGE34QR						-0.101		
IHISP265	Hispanic Interaction Of BAGE34LN				-0.008				
IHISP266	Hispanic Interaction Of BAGE34QU				-0.388				
BAGE44LN	B: Linear: Percent Persons 35-44 Years						-0.080	0.002	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years						-0.086	0.006	
BAGE44CU	B: Cubic: Percent Persons 35-44 Years						-0.035	-0.110	
BAGE44QR	B: Quartic: Percent Persons 35-44 Years						0.042		
IBLK270	Black Interaction Of BAGE44LN						0.094		
IBLK271	Black Interaction Of BAGE44QU						0.093		
IBLK272	Black Interaction Of BAGE44CU						-0.075		
IBLK273	Black Interaction Of BAGE44QR						-0.127		
IHISP270	Hispanic Interaction Of BAGE44LN							-0.127	
IHISP271	Hispanic Interaction Of BAGE44QU							-0.323	
BAGE54LN	B: Linear: Percent Persons 45-54 Years						-0.126		
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years						-0.030		
BAGE54CU	B: Cubic: Percent Persons 45-54 Years						-0.060		
BAGE54QR	B: Quartic: Percent Persons 45-54 Years						-0.040		
BAGE64LN	B: Linear: Percent Persons 55-64 Years		-0.139			-0.186	-0.040		
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years						-0.057		
BAGE64CU	B: Cubic: Percent Persons 55-64 Years						0.046		
IHISP280	Hispanic Interaction Of BAGE64LN		0.298				-0.033		
IHISP281	Hispanic Interaction Of BAGE64QU						-0.086		
IHISP282	Hispanic Interaction Of BAGE64CU						-0.240		
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut		-0.005			0.170			
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut					-0.046			
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut					0.184			
IOTH300	Other Interaction Of BINDIALN		0.671						
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed	0.043							
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed	-0.139							
BPOVERLN	B: Linear: Percent Families Below Poverty Level							0.057	
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level							0.049	
BPOVERCU	B: Cubic: Percent Families Below Poverty Level							-0.171	
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican					0.120			
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican					0.167			
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					-0.174			
IFEM315	Female Interaction Of BPRICALN					-0.283			
IFEM316	Female Interaction Of BPRICAQU					-0.125			
IFEM317	Female Interaction Of BPRICACU					0.403			
BSCHASLN	B: Linear: Percent Associates Degree			-0.168		-0.041		0.001	
BSCHASQU	B: Quadratic: Percent Associates Degree					0.018		0.018	
BSCHASCU	B: Cubic: Percent Associates Degree					0.172		0.023	
BSCHASQR	B: Quartic: Percent Associates Degree							-0.059	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander	-0.083				0.061			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander					0.048			
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander					-0.152			
IFEM325	Female Interaction Of PASIANLN					0.118			
IFEM326	Female Interaction Of PASIANQU					-0.029			
IFEM327	Female Interaction Of PASIANCU					0.334			
PCUBANLN	T: Linear: Percent Hispanics: Cuban							0.026	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban							-0.023	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban							-0.311	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

E11. Coefficients of Model Parameters for Past Year Treatment Needed for Drug Abuse

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	-3.441	-2.427	-2.751	-3.791	-3.445	-2.226	-3.181	-4.922
FEMALE	O: Female Indicator	0.246	-0.602	-0.654	-0.524	-0.011	-0.583	-0.279	-0.990
FEMBLCK	O: Black Interaction Of FEMALE	-0.318	-0.109	0.236	-0.776		0.303	-0.658	-0.173
FEMHISP	O: Hispanic Interaction Of FEMALE	-0.570	0.186	0.405	0.028		0.069	-0.771	0.355
FEMOTHR	O: Other Interaction Of FEMALE	0.479	0.422	1.042			-2.145	0.083	1.541
RACEBLCK	O: Race/Black Indicator	-0.331	-0.299	0.010	0.689		-1.157	0.615	0.939
RACEHISP	O: Race/Hispanic Indicator	0.264	-0.626	-0.642	-0.281		-0.463	-0.207	0.132
RACEOTHR	O: Race/Other Indicator	-4.288	-1.315	-2.825	-0.586		-1.072	-3.867	-0.274
PDENLEV1	O: Large MSA						-0.235		
PDENLEV2	O: Medium MSA						-0.171		
PDENLEV3	O: Small MSA						-0.056		
PDENLEV4	O: NonMSA, Urban						-0.014		
IBLK10	Black Interaction Of PDENLEV1						0.319		
IBLK11	Black Interaction Of PDENLEV2						0.508		
IBLK12	Black Interaction Of PDENLEV3						1.306		
IBLK13	Black Interaction Of PDENLEV4						1.185		
PHH1PLN	T: Linear: Percent One Person Households				0.210				-0.175
PHH1PQU	T: Quadratic: Percent One Person Households				0.071				0.059
PHH1PCU	T: Cubic: Percent One Person Households				-0.016				-0.088
PHH1PQR	T: Quartic: Percent One Person Households				-0.017				-0.086
PAGE18LN	T: Linear: Percent Persons 0-18 Years						0.021		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years						0.078		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years						0.054		
IOTH25	Other Interaction Of PAGE18LN						0.052		
IOTH26	Other Interaction Of PAGE18QU						-0.794		
PAGE24LN	T: Linear: Percent Persons 19-24 Years								-0.107
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years								0.107
PAGE24CU	T: Cubic: Percent Persons 19-24 Years								-0.054
PAGE24QR	T: Quartic: Percent Persons 19-24 Years								0.074
PAGE44LN	T: Linear: Percent Persons 35-44 Years				0.054				
PAGE54LN	T: Linear: Percent Persons 45-54 Years				-0.023		-0.130		-0.155
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years								-0.042
PAGE54CU	T: Cubic: Percent Persons 45-54 Years								-0.091
PAGE54QR	T: Quartic: Percent Persons 45-54 Years								-0.042
IFEM45	Female Interaction Of PAGE54LN				0.205				
IHISP45	Hispanic Interaction Of PAGE54LN						0.274		
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma							0.176	
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma							-0.084	
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma							0.043	
IFEM60	Female Interaction Of PSCH12LN							0.025	
IFEM61	Female Interaction Of PSCH12QU							0.161	
PPOVERLN	T: Linear: Percent Families Below Poverty Level					-0.026			
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level					0.020			
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					-0.133			
PBLACKLN	T: Linear: Percent Black Nonhispanic							-0.144	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic							-0.129	
IOTH100	Other Interaction Of PBLACKLN							-2.087	
IOTH101	Other Interaction Of PBLACKQU							-1.492	
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force					-0.108	0.055		
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force					-0.096	-0.043		
PFLABCU	T: Cubic: Percent Females 16+ Years Old In Labor Force						0.012		
PFLABQR	T: Quartic: Percent Females 16+ Years Old In Labor Force						0.031		
IBLK125	Black Interaction Of PFLABLN						-0.006		
IBLK126	Black Interaction Of PFLABQU						0.203		
PFNEVLN	T: Linear: Percent Females Never Married							0.092	
PFNEVQU	T: Quadratic: Percent Females Never Married							0.126	
IFEM130	Female Interaction Of PFNEVLN							0.036	
IFEM131	Female Interaction Of PFNEVQU							-0.107	
PMNEVLN	T: Linear: Percent Males Never Married				-0.005				0.223
PMNEVQU	T: Quadratic: Percent Males Never Married				-0.100				0.062
PMNEVCU	T: Cubic: Percent Males Never Married								-0.268
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force								0.032
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force								-0.040
PMLABCU	T: Cubic: Percent Males 16+ Years Old In Labor Force								-0.073
PMLABQR	T: Quartic: Percent Males 16+ Years Old In Labor Force								-0.040
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed							0.149	
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier					-0.094			
P40HULN	T: Linear: Percent Housing Units Built 1940-1949	0.026				-0.050			
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949	0.044				-0.023			
P40HUCU	T: Cubic: Percent Housing Units Built 1940-1949					-0.013			
IHISP160	Hispanic Interaction Of P40HULN	-0.049							
IHISP161	Hispanic Interaction Of P40HUQU	-0.096							
PRENTLN	T: Linear: Percent Housing Rented			0.097				0.035	0.590
PRENTQU	T: Quadratic: Percent Housing Rented							0.055	
PRENTCU	T: Cubic: Percent Housing Rented							0.010	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
PRENTQR	T: Quartic: Percent Housing Rented							0.011	
IHISP165	Hispanic Interaction Of PRENTLN							-0.150	
IHISP166	Hispanic Interaction Of PRENTQU							0.085	
IHISP167	Hispanic Interaction Of PRENTCU							-0.203	
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases						-0.023		
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases						0.117		
ADRATECU	C: Cubic: Death Rate For All Alcohol-Related Cases						-0.030		
ADRATEQR	C: Quartic: Death Rate For All Alcohol-Related Cases						0.027		
IBLK175	Black Interaction Of ADRATELN						0.058		
IBLK176	Black Interaction Of ADRATEQU						-0.084		
IBLK177	Black Interaction Of ADRATECU						0.133		
IBLK178	Black Interaction Of ADRATEQR						-0.046		
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol								0.249
ADRAT1QU	C: Quadratic: Death Rate With Explicit Mention Of Alcohol								0.083
IFEM180	Female Interaction Of ADRAT1LN								-0.354
IFEM181	Female Interaction Of ADRAT1QU								-0.570
V18FLN	C: Linear: Marijuana Posession Arrest Rate				-0.224				
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate			0.089				-0.059	-0.007
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate							0.037	0.089
V18BCU	C: Cubic: Marijuana Sale/Manufacture Arrest Rate							0.111	0.358
V18BQR	C: Quartic: Marijuana Sale/Manufacture Arrest Rate								0.082
IFEM190	Female Interaction Of V18BLN								-0.051
IFEM191	Female Interaction Of V18BQU								-0.167
IFEM192	Female Interaction Of V18BCU								-0.570
IFEM193	Female Interaction Of V18BQR								-0.137
IBLK190	Black Interaction Of V18BLN							0.092	
IBLK191	Black Interaction Of V18BQU							0.093	
IBLK192	Black Interaction Of V18BCU							-0.211	
IHISP190	Hispanic Interaction Of V18BLN			-0.143					
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate								-0.264
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate								-0.275
IHISP200	Hispanic Interaction Of V18ALN								0.135
IHISP201	Hispanic Interaction Of V18AQU								0.432
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate						0.133		
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate						-0.073		
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992						-0.051		
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						-0.024		
IBLK235	Black Interaction Of DRATELN						0.066		
IBLK236	Black Interaction Of DRATEQU						-0.035		
RH43ALN	T: Linear: Recoded Median Rents For Rental Units								0.009
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units								0.301
IBLK240	Black Interaction Of RH43ALN								0.165
IBLK241	Black Interaction Of RH43AQU								-0.379
RP80ALN	T: Linear: Recoded Median Household Income		0.132						
IBLK250	Black Interaction Of RP80ALN		-0.368						
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability						0.073	0.189	
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability							-0.038	
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability							-0.088	
IHISP260	Hispanic Interaction Of B64DISLN							-0.063	
IHISP261	Hispanic Interaction Of B64DISQU							-0.008	
IHISP262	Hispanic Interaction Of B64DISCU							0.210	
IOTH260	Other Interaction Of B64DISLN							-0.344	
IOTH261	Other Interaction Of B64DISQU							0.652	
BAGE34LN	B: Linear: Percent Persons 25-34 Years						0.012		-0.054
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years								0.008
BAGE34CU	B: Cubic: Percent Persons 25-34 Years								0.050
BAGE34QR	B: Quartic: Percent Persons 25-34 Years								-0.014
IBLK265	Black Interaction Of BAGE34LN								-0.067
IBLK266	Black Interaction Of BAGE34QU								-0.040
IBLK267	Black Interaction Of BAGE34CU								-0.039
IBLK268	Black Interaction Of BAGE34QR								0.056
BAGE44LN	B: Linear: Percent Persons 35-44 Years	0.013				-0.164		-0.011	-0.012
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years								-0.001
BAGE44CU	B: Cubic: Percent Persons 35-44 Years								0.125
IOTH270	Other Interaction Of BAGE44LN	2.045						0.855	
BAGE54LN	B: Linear: Percent Persons 45-54 Years					-0.003	0.116		
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years					-0.040			
BAGE54CU	B: Cubic: Percent Persons 45-54 Years					0.039			
BAGE54QR	B: Quartic: Percent Persons 45-54 Years					0.034			
IOTH275	Other Interaction Of BAGE54LN						-1.094		
BAGE64LN	B: Linear: Percent Persons 55-64 Years		-0.090				-0.025		
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		-0.016				-0.059		
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		0.068				0.038		
IBLK280	Black Interaction Of BAGE64LN						0.179		
IBLK281	Black Interaction Of BAGE64QU						0.025		
IBLK282	Black Interaction Of BAGE64CU						-0.136		
IHISP280	Hispanic Interaction Of BAGE64LN		0.194						
IHISP281	Hispanic Interaction Of BAGE64QU		-0.088						
IHISP282	Hispanic Interaction Of BAGE64CU		-0.173						
BCUBANLN	B: Linear: Percent Hispanics: Cuban								-0.193
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban								-0.388

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed	0.033	0.076			0.044	0.061	0.064	
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed		0.010			-0.062	-0.054	-0.097	
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed					0.011	0.042	-0.051	
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed					0.028	0.005	0.007	
IBLK295	Black Interaction Of BFNOTLN						-0.144		
IBLK296	Black Interaction Of BFNOTQU						0.204		
IBLK297	Black Interaction Of BFNOTCU						-0.085		
IBLK298	Black Interaction Of BFNOTQR						-0.070		
IHISP295	Hispanic Interaction Of BFNOTLN		-0.123				-0.187		
IHISP296	Hispanic Interaction Of BFNOTQU		0.118				-0.162		
IHISP297	Hispanic Interaction Of BFNOTCU						-0.185		
IOTH295	Other Interaction Of BFNOTLN	0.657						0.544	
IOTH296	Other Interaction Of BFNOTQU							0.523	
IOTH297	Other Interaction Of BFNOTCU							0.228	
IOTH298	Other Interaction Of BFNOTQR							0.352	
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut						0.137		-0.024
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut						-0.094		0.253
IFEM300	Female Interaction Of BINDIALN						-0.061		
IFEM301	Female Interaction Of BINDIAQU						0.127		
IHISP300	Hispanic Interaction Of BINDIALN								-0.160
IHISP301	Hispanic Interaction Of BINDIAQU								-0.371
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed							-0.045	
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed							0.011	
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed							-0.021	
BMNOTQR	B: Quartic: Percent Males Separated, Divorced or Widowed							0.021	
IHISP305	Hispanic Interaction Of BMNOTLN							0.081	
IHISP306	Hispanic Interaction Of BMNOTQU							0.009	
IHISP307	Hispanic Interaction Of BMNOTCU							0.236	
IHISP308	Hispanic Interaction Of BMNOTQR							-0.092	
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.134		0.256		-0.051		-0.195
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.106		0.104				-0.194
BPOVERCU	B: Cubic: Percent Families Below Poverty Level								-0.034
BPOVERQR	B: Quartic: Percent Families Below Poverty Level								-0.042
IFEM310	Female Interaction Of BPOVERLN				-0.040		0.192		0.034
IFEM311	Female Interaction Of BPOVERQU				-0.199				0.233
IFEM312	Female Interaction Of BPOVERCU								-0.165
IFEM313	Female Interaction Of BPOVERQR								0.102
IBLK310	Black Interaction Of BPOVERLN		-0.299						
IBLK311	Black Interaction Of BPOVERQU		-0.123						
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican						0.231		
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican						0.036		
IBLK315	Black Interaction Of BPRICALN						-0.332		
IBLK316	Black Interaction Of BPRICAQU						0.479		
BSCHASLN	B: Linear: Percent Associates Degree			0.029				0.069	0.144
BSCHASQU	B: Quadratic: Percent Associates Degree							0.016	
BSCHASCU	B: Cubic: Percent Associates Degree							0.074	
BSCHASQR	B: Quartic: Percent Associates Degree							-0.037	
IFEM320	Female Interaction Of BSCHASLN							-0.001	
IFEM321	Female Interaction Of BSCHASQU							0.018	
IFEM322	Female Interaction Of BSCHASCU							0.047	
IFEM323	Female Interaction Of BSCHASQR							0.059	
IBLK320	Black Interaction Of BSCHASLN			-0.279					
IOTH320	Other Interaction Of BSCHASLN							-0.410	
IOTH321	Other Interaction Of BSCHASQU							0.971	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander				-0.032			0.101	
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander							0.034	
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander							-0.085	
IBLK325	Black Interaction Of PASIANLN							-0.083	
IBLK326	Black Interaction Of PASIANQU							0.104	
IBLK327	Black Interaction Of PASIANCU							0.321	
IOTH325	Other Interaction Of PASIANLN								-0.664
PCUBANLN	T: Linear: Percent Hispanics: Cuban			-0.021				0.022	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban			0.127				0.139	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban			0.100				-0.093	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban							0.037	
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut					0.064	0.096	0.061	
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut					-0.084	0.093	0.026	
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut					0.097		0.113	
IFEM335	Female Interaction Of PINDIALN						-0.047		
IFEM336	Female Interaction Of PINDIAQU						-0.110		
IOTH335	Other Interaction Of PINDIALN							-0.255	
IOTH336	Other Interaction Of PINDIAQU							0.425	
IOTH337	Other Interaction Of PINDIACU							-0.934	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Appendix F: SignificanceProbabilities for Fixed Effect Coefficients

F1. Significance Probabilities for Fixed Effect Coefficients for Past Month Alcohol Use by Age

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	29.20	0.00	0.00	0.00	0.12	0.00	0.00	0.00
FEMBLCK	O: Black Interaction Of FEMALE	15.10	5.21	0.00	0.00	46.56	27.70	4.71	0.00
FEMHISP	O: Hispanic Interaction Of FEMALE	0.36	0.00	0.00	0.00	98.60	0.04	0.04	0.00
FEMOTHR	O: Other Interaction Of FEMALE	41.04	30.60	0.00	0.00	2.09	95.46	44.24	0.00
RACEBLCK	O: Race/Black Indicator	0.00	0.00	0.00	0.00	5.11	0.00	0.00	0.00
RACEHISP	O: Race/Hispanic Indicator	95.35	0.00	66.62	58.11	99.80	37.20	43.40	0.53
RACEOTHR	O: Race/Other Indicator	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
REGNOREA	O: Northeast Region Indicator						32.62		30.40
REGSOUTH	O: South Region Indicator						2.01		0.06
REGWEST	O: West Region Indicator						50.63		39.18
IOTH7	Other Interaction Of REGNOREA						0.00		0.00
IOTH8	Other Interaction Of REGSOUTH						0.00		0.00
IOTH9	Other Interaction Of REGWEST						0.00		0.00
UCLASS9	T: Underclass Indicator								0.00
PAGE18LN	T: Linear: Percent Persons 0-18 Years				0.00				0.19
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years								0.00
IOTH25	Other Interaction Of PAGE18LN								0.00
IOTH26	Other Interaction Of PAGE18QU								0.00
PAGE24LN	T: Linear: Percent Persons 19-24 Years						7.83		
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years						0.00		
IFEM30	Female Interaction Of PAGE24LN						69.15		
IFEM31	Female Interaction Of PAGE24QU						0.00		
PAGE34LN	T: Linear: Percent Persons 25-34 Years	0.01							
PAGE44LN	T: Linear: Percent Persons 35-44 Years				0.00				0.00
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years								0.00
PAGE44CU	T: Cubic: Percent Persons 35-44 Years								15.17
PAGE44QR	T: Quartic: Percent Persons 35-44 Years								0.00
PAGE64LN	T: Linear: Percent Persons 55-64 Years	1.90							
PSCH8LN	T: Linear: Percent 0-8 Years Of School		0.00						
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree		0.00				0.00	0.00	
IBLK75	Black Interaction Of PSCHCOLN						0.00	0.01	
PSCHSCLN	T: Linear: Percent Some College And No Degree		0.00						
PPOVERLN	T: Linear: Percent Families Below Poverty Level					2.48			
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level					22.11			
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					0.00			
PPOVERQR	T: Quartic: Percent Families Below Poverty Level					0.27			
IBLK85	Black Interaction Of PPOVERLN					2.49			
IBLK86	Black Interaction Of PPOVERQU					50.49			
IBLK87	Black Interaction Of PPOVERCU					36.10			
IBLK88	Black Interaction Of PPOVERQR					3.40			
PBLACKLN	T: Linear: Percent Black Nonhispanic		0.00						
POTHLN	T: Linear: Percent Other Race/Hispanicity						7.76		
POTHQU	T: Quadratic: Percent Other Race/Hispanicity						0.00		
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force	6.05							
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force	0.07							
PFNEVLN	T: Linear: Percent Females Never Married						0.00		
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed				0.00				
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed				20.07				
PFNOTCU	T: Cubic: Percent Females Separated, Divorced Or Widowed				0.00				
PFNOTQR	T: Quartic: Percent Females Separated, Divorced Or Widowed				0.00				
PMNEVLN	T: Linear: Percent Males Never Married			0.00		0.22		0.21	0.00
PMNEVQU	T: Quadratic: Percent Males Never Married							6.51	0.00
PMNEVCU	T: Cubic: Percent Males Never Married							0.94	54.64
PMNEVQR	T: Quartic: Percent Males Never Married								0.00
IFEM140	Female Interaction Of PMNEVLN								0.00
IFEM141	Female Interaction Of PMNEVQU								0.00
IBLK140	Black Interaction Of PMNEVLN					0.53			
IHISP140	Hispanic Interaction Of PMNEVLN			0.00					
IOTH140	Other Interaction Of PMNEVLN			0.00				24.42	
IOTH141	Other Interaction Of PMNEVQU							42.50	
IOTH142	Other Interaction Of PMNEVCU							0.00	
P40HULN	T: Linear: Percent Housing Units Built 1940-1949					0.01			
IFEM160	Female Interaction Of P40HULN					0.04			
PRENTLN	T: Linear: Percent Housing Rented			0.48		80.55			
IBLK165	Black Interaction Of PRENTLN			0.01					

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

		BIG CITY				REMAINDER			
Variable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases		86.86						
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases		0.00						
V18FLN	C: Linear: Marijuana Posession Arrest Rate								0.00
IOTH185	Other Interaction Of V18FLN								0.00
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate		86.81		0.00				0.00
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate		0.01		0.00				0.00
V18ACU	C: Cubic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate				0.00				4.60
V18AQR	C: Quartic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate								0.00
IHISP200	Hispanic Interaction Of V18ALN		0.00						
IHISP201	Hispanic Interaction Of V18AQU		0.00						
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate								0.00
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992					1.22			
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992					49.97			
DRATECU	C: Cubic: Mean Drug Client Treatment Rate 1991 & 1992					22.78			
IOTH235	Other Interaction Of DRATELN					86.06			
IOTH236	Other Interaction Of DRATEQU					55.24			
IOTH237	Other Interaction Of DRATECU					0.05			
RH61ALN	T: Linear: Recoded Median Value Of Owner Occupied HUs		83.25						0.00
RH61AQU	T: Quadratic: Recoded Median Value Of Owner Occupied HUs		52.81						
IHISP245	Hispanic Interaction Of RH61ALN		0.01						
IHISP246	Hispanic Interaction Of RH61AQU		0.00						
RP80ALN	T: Linear: Recoded Median Household Income				0.00				
RP80AQU	T: Quadratic: Recoded Median Household Income				0.00				
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992		47.83						
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992		37.47						
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992		48.01						
IFEM255	Female Interaction Of ARATELN		0.12						
IFEM256	Female Interaction Of ARATEQU		5.10						
IFEM257	Female Interaction Of ARATECU		0.00						
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability				0.00		26.54		0.00
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability				0.00		53.81		
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability						0.00		
IHISP260	Hispanic Interaction Of B64DISLN						45.33		0.00
IHISP261	Hispanic Interaction Of B64DISQU						2.01		
IHISP262	Hispanic Interaction Of B64DISCU						0.61		
BAGE34LN	B: Linear: Percent Persons 25-34 Years							1.59	
IOTH265	Other Interaction Of BAGE34LN							0.00	
BAGE44LN	B: Linear: Percent Persons 35-44 Years								0.00
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years								7.53
IHISP270	Hispanic Interaction Of BAGE44LN								0.00
IHISP271	Hispanic Interaction Of BAGE44QU								0.00
BAGE54LN	B: Linear: Percent Persons 45-54 Years				0.00				0.00
IBLK275	Black Interaction Of BAGE54LN								0.00
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander							90.86	0.00
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander							57.39	29.56
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander							0.02	22.18
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander							57.06	
IOTH285	Other Interaction Of BASIANLN							0.05	0.00
IOTH286	Other Interaction Of BASIANQU							97.29	0.00
IOTH287	Other Interaction Of BASIANCU							0.01	0.00
IOTH288	Other Interaction Of BASIANQR							0.00	
BCUBANLN	B: Linear: Percent Hispanics: Cuban		7.86			27.48			
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban					49.56			
BCUBANCU	B: Cubic: Percent Hispanics: Cuban					39.14			
IHISP290	Hispanic Interaction Of BCUBANLN		0.00						
IOTH290	Other Interaction Of BCUBANLN					53.03			
IOTH291	Other Interaction Of BCUBANQU					86.43			
IOTH292	Other Interaction Of BCUBANCU					5.64			
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed		0.05				3.85		
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed		91.65				0.19		
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed		0.22						
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed		0.02						
IOTH295	Other Interaction Of BFNOTLN		0.13				22.41		
IOTH296	Other Interaction Of BFNOTQU		97.61				0.00		
IOTH297	Other Interaction Of BFNOTCU		0.00						
IOTH298	Other Interaction Of BFNOTQR		0.00						
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut					1.69	26.85		
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut						43.34		
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut						0.00		
IFEM300	Female Interaction Of BINDIALN					0.04	73.67		
IFEM301	Female Interaction Of BINDIAQU						80.16		
IFEM302	Female Interaction Of BINDIACU						0.00		
IBLK300	Black Interaction Of BINDIALN					2.38			
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed	4.45							0.00
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed								39.55
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed								0.00

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IHISP305	Hispanic Interaction Of BMNOTLN	0.12							0.24
IHISP306	Hispanic Interaction Of BMNOTQU								59.03
IHISP307	Hispanic Interaction Of BMNOTCU								0.00
BPOVERLN	B: Linear: Percent Families Below Poverty Level			0.00					
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	5.52				73.88	37.62		
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican	60.48				99.91	0.05		
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican	0.11				46.06	49.01		
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican						0.00		
IBLK315	Black Interaction Of BPRICALN						18.30		
IBLK316	Black Interaction Of BPRICAQU						0.05		
BSCHASLN	B: Linear: Percent Associates Degree				0.00				0.00
BSCHASQU	B: Quadratic: Percent Associates Degree				0.00				
BSCHASCU	B: Cubic: Percent Associates Degree				0.00				
IBLK320	Black Interaction Of BSCHASLN				0.00				
IBLK321	Black Interaction Of BSCHASQU				0.00				
IBLK322	Black Interaction Of BSCHASCU				0.00				
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander			0.01					
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander			0.00					
IBLK325	Black Interaction Of PASIANLN			0.07					
IBLK326	Black Interaction Of PASIANQU			0.00					
PCUBANLN	T: Linear: Percent Hispanics: Cuban				0.00				
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban				1.19				
PCUBANCU	T: Cubic: Percent Hispanics: Cuban				0.00				
IFEM330	Female Interaction Of PCUBANLN				0.41				
IFEM331	Female Interaction Of PCUBANQU				0.00				
IFEM332	Female Interaction Of PCUBANCU				0.00				
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut			0.00					30.06
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			0.00					0.00
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut			9.34					0.00
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut			0.00					0.00
IFEM335	Female Interaction Of PINDIALN								0.00
IFEM336	Female Interaction Of PINDIAQU								0.00
IFEM337	Female Interaction Of PINDIACU								0.00
IFEM338	Female Interaction Of PINDIAQR								0.00

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F2. Significance Probabilities for Fixed Effect Coefficients for Past Month Any Illicit Drug Use by Age

Varibable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
UCLASS9	T: Underclass Indicator						0.55		
FEMALE	O: Female Indicator	64.64	0.00	0.00	0.00	48.67	0.00	0.00	0.00
FEMBLCK	O: Black Interaction Of FEMALE	0.41	0.00	0.01	9.18	88.02	31.63	43.68	1.24
FEMHISP	O: Hispanic Interaction Of FEMALE	10.40	27.91	0.00	0.00		28.86	40.07	59.64
FEMOTHR	O: Other Interaction Of FEMALE	23.94	69.98	0.09			56.72	0.00	0.01
RACEBLCK	O: Race/Black Indicator	59.06	4.39	0.00	0.00	18.51	0.31	90.64	0.00
RACEHISP	O: Race/Hispanic Indicator	4.56	0.00	0.00	0.00	48.56	0.14	24.39	0.15
RACEOTHR	O: Race/Other Indicator	0.02	0.00	0.00	0.00		0.20	0.00	0.00
REGNOREA	O: Northeast Region Indicator						4.96	1.92	
REGSOUTH	O: South Region Indicator						12.14	22.19	
REGWEST	O: West Region Indicator						0.01	0.00	
PDENLEV1	O: Large MSA					72.37	0.13	51.18	
PDENLEV2	O: Medium MSA					29.66	1.67	97.05	
PDENLEV3	O: Small MSA					16.25	0.88	84.57	
PDENLEV4	O: NonMSA, Urban					53.05	0.39	64.75	
IBLK10	Black Interaction Of PDENLEV1						4.41		
IBLK11	Black Interaction Of PDENLEV2						17.62		
IBLK12	Black Interaction Of PDENLEV3						1.87		
IBLK13	Black Interaction Of PDENLEV4						0.98		
IHISP10	Hispanic Interaction Of PDENLEV1							2.08	
IHISP11	Hispanic Interaction Of PDENLEV2							0.38	
IHISP12	Hispanic Interaction Of PDENLEV3							12.70	
IHISP13	Hispanic Interaction Of PDENLEV4							13.06	
PHH1PLN	T: Linear: Percent One Person Households						9.71		
PHH1PQU	T: Quadratic: Percent One Person Households						0.03		
POPRMLN	T: Linear: Average Persons Per Room			38.68				0.00	
POPRMQU	T: Quadratic: Average Persons Per Room							0.00	
IFEM20	Female Interaction Of POPRMLN							0.93	
IFEM21	Female Interaction Of POPRMQU							0.04	
PAGE18LN	T: Linear: Percent Persons 0-18 Years						4.17		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years						37.39		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years						0.00		
IFEM25	Female Interaction Of PAGE18LN						0.09		
IFEM26	Female Interaction Of PAGE18QU						98.64		
IFEM27	Female Interaction Of PAGE18CU						0.01		
PAGE24LN	T: Linear: Percent Persons 19-24 Years							0.00	
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years							4.50	
PAGE24CU	T: Cubic: Percent Persons 19-24 Years							0.00	
PAGE24QR	T: Quartic: Percent Persons 19-24 Years							37.40	
IHISP30	Hispanic Interaction Of PAGE24LN							66.58	
IHISP31	Hispanic Interaction Of PAGE24QU							64.69	
IHISP32	Hispanic Interaction Of PAGE24CU							81.90	
IHISP33	Hispanic Interaction Of PAGE24QR							62.74	
PAGE34LN	T: Linear: Percent Persons 25-34 Years					96.79	41.91		
PAGE34QU	T: Quadratic: Percent Persons 25-34 Years					0.02	2.16		
IHISP35	Hispanic Interaction Of PAGE34LN						48.67		
IHISP36	Hispanic Interaction Of PAGE34QU						1.50		
PAGE44LN	T: Linear: Percent Persons 35-44 Years							0.63	
PAGE54LN	T: Linear: Percent Persons 45-54 Years				0.00		0.00		0.00
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years				0.68		6.44		0.00
PAGE54CU	T: Cubic: Percent Persons 45-54 Years						0.16		
IBLK45	Black Interaction Of PAGE54LN				0.00				
IBLK46	Black Interaction Of PAGE54QU				0.00				
IHISP45	Hispanic Interaction Of PAGE54LN						1.71		
PSCH8LN	T: Linear: Percent 0-8 Years Of School							2.32	
PSCH8QU	T: Quadratic: Percent 0-8 Years Of School							14.81	
IBLK55	Black Interaction Of PSCH8LN							0.16	
IBLK56	Black Interaction Of PSCH8QU							0.03	
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma	95.60			0.00	18.22			4.37
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma	0.46				0.64			0.00
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma	18.91							
PSCH12QR	T: Quartic: Percent 9-12 Years & No High School Diploma	0.48							
IFEM60	Female Interaction Of PSCH12LN					35.76			
IFEM61	Female Interaction Of PSCH12QU					1.18			
PSCHASLN	T: Linear: Percent Associates Degree							44.93	
PSCHASQU	T: Quadratic: Percent Associates Degree							0.00	
IBLK65	Black Interaction Of PSCHASLN							30.45	
IBLK66	Black Interaction Of PSCHASQU							0.80	
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree					11.51			87.93

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Varibable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree					88.61			0.00
PSCHCOCU	T: Cubic: Bachelors, Graduate, Or Professional Degree					1.19			
IFEM75	Female Interaction Of PSCHCOLN					83.45			
IFEM76	Female Interaction Of PSCHCOQU					73.24			
PPOVERLN	T: Linear: Percent Families Below Poverty Level	0.58				15.37	0.00	0.09	
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level	72.28				88.13		25.01	
PPOVERCU	T: Cubic: Percent Families Below Poverty Level	0.00				15.02		2.96	
PPOVERQR	T: Quartic: Percent Families Below Poverty Level					3.83		0.01	
IFEM85	Female Interaction Of PPOVERLN						28.49		
PBLACKLN	T: Linear: Percent Black Nonhispanic			0.00			1.80	5.57	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic			73.24			34.80	1.16	
PBLACKCU	T: Cubic: Percent Black Nonhispanic			0.00			40.82	37.41	
PBLACKQR	T: Quartic: Percent Black Nonhispanic						0.01		
PHISPLN	T: Linear: Percent Hispanic								0.66
PHISPQU	T: Quadratic: Percent Hispanic								0.23
PHISPCU	T: Cubic: Percent Hispanic								0.00
PHISPQR	T: Quartic: Percent Hispanic								0.05
POTHLN	T: Linear: Percent Other Race/Hispanicity					15.11			
POTHQU	T: Quadratic: Percent Other Race/Hispanicity					96.69			
POTHCU	T: Cubic: Percent Other Race/Hispanicity					53.72			
PHHF18LN	T: Linear: % Female-Headed HH W/No Spouse & Chld Under 18					47.63			29.88
PHHF18QU	T: Quadratic: % Female-Headed HH W/No Spouse & Chld Under 18					59.87			
PHHF18CU	T: Cubic: % Female-Headed HH W/No Spouse & Chld Under 18					69.18			
PHHF18QR	T: Quartic: % Female-Headed HH W/No Spouse & Chld Under 18					25.38			
IFEM120	Female Interaction Of PHHF18LN								0.22
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed			0.00				92.70	
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed							4.77	
IHISP135	Hispanic Interaction Of PFNOTLN							7.57	
IHISP136	Hispanic Interaction Of PFNOTQU							40.16	
PMNEVLN	T: Linear: Percent Males Never Married				0.00				
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed							0.21	
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed							2.95	
IHISP150	Hispanic Interaction Of PMNOTLN							75.31	
IHISP151	Hispanic Interaction Of PMNOTQU							6.65	
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier			0.00			0.01	81.76	
POLDHUQU	T: Quadratic: Percent Housing Units Built 1939 Or Earlier			3.75			35.15	0.01	
POLDHUCU	T: Cubic: Percent Housing Units Built 1939 Or Earlier			0.00					
P40HULN	T: Linear: Percent Housing Units Built 1940-1949					55.89		0.00	
PRENTLN	T: Linear: Percent Housing Rented					11.66	32.47	0.00	0.00
PRENTQU	T: Quadratic: Percent Housing Rented							0.13	0.00
PRENTCU	T: Cubic: Percent Housing Rented							3.96	46.39
PRENTQR	T: Quartic: Percent Housing Rented							0.00	0.00
IFEM165	Female Interaction Of PRENTLN					0.14			
IBLK165	Black Interaction Of PRENTLN							0.31	
IBLK166	Black Interaction Of PRENTQU							37.29	
IBLK167	Black Interaction Of PRENTCU							2.03	
IHISP165	Hispanic Interaction Of PRENTLN						51.22		
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases		82.11			54.19	0.37		
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases		80.43			44.92	39.73		
ADRATECU	C: Cubic: Death Rate For All Alcohol-Related Cases		0.01			9.55	89.07		
ADRATEQR	C: Quartic: Death Rate For All Alcohol-Related Cases						0.00		
IBLK175	Black Interaction Of ADRATELN						47.11		
IBLK176	Black Interaction Of ADRATEQU						11.73		
IBLK177	Black Interaction Of ADRATECU						43.82		
IBLK178	Black Interaction Of ADRATEQR						0.12		
V18FLN	C: Linear: Marijuana Posession Arrest Rate					71.95			
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate						19.71	0.01	0.00
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate						17.38		0.00
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate						34.23		0.05
V18EQR	C: Quartic: Opium/Cocaine & Deriv Posession Arrest Rate						0.00		76.84
IFEM195	Female Interaction Of V18ELN						0.80		
IFEM196	Female Interaction Of V18EQU						10.92		
IFEM197	Female Interaction Of V18ECU						43.77		
IFEM198	Female Interaction Of V18EQR						0.00		
IBLK195	Black Interaction Of V18ELN						0.24		
IHISP195	Hispanic Interaction Of V18ELN						0.16		
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate	32.64	0.00			33.40		85.95	
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate	1.42				17.74			
V18ACU	C: Cubic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate	19.69				0.16			
IFEM200	Female Interaction Of V18ALN	99.67							
IFEM201	Female Interaction Of V18AQU	27.95							
IFEM202	Female Interaction Of V18ACU	1.11							
IBLK200	Black Interaction Of V18ALN					19.64			
IBLK201	Black Interaction Of V18AQU					55.11			
IBLK202	Black Interaction Of V18ACU					6.06			
IHISP200	Hispanic Interaction Of V18ALN		0.00						
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate		8.26	84.55	0.26			0.07	16.24

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Varibable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate			1.59				0.05	0.00
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate			57.21				57.05	
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate			0.02				0.02	
IFEM225	Female Interaction Of V18LN							30.93	
IFEM226	Female Interaction Of V18QU							4.42	
IFEM227	Female Interaction Of V18CU							64.77	
IFEM228	Female Interaction Of V18QR							3.73	
IHISP225	Hispanic Interaction Of V18LN		86.80						
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate				40.38				0.00
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate				1.52				0.00
IBLK230	Black Interaction Of VIOLLN				0.00				
IBLK231	Black Interaction Of VIOLQU				3.40				
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992					0.68	45.77	37.47	
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						0.34	0.00	
IFEM235	Female Interaction Of DRATELN							4.44	
IFEM236	Female Interaction Of DRATEQU							0.00	
RH43ALN	T: Linear: Recoded Median Rents For Rental Units	58.26						99.05	0.00
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units	53.33							0.00
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units	41.72							
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units	0.06							
IBLK240	Black Interaction Of RH43ALN								80.90
IBLK241	Black Interaction Of RH43AQU								0.00
RH61ALN	T: Linear: Recoded Median Value Of Owner Occupied HUs						60.84		
RH61AQU	T: Quadratic: Recoded Median Value Of Owner Occupied HUs						87.63		
RH61ACU	T: Cubic: Recoded Median Value Of Owner Occupied HUs						0.00		
RP80ALN	T: Linear: Recoded Median Household Income					8.20		20.29	
RP80AQU	T: Quadratic: Recoded Median Household Income					70.51			
RP80ACU	T: Cubic: Recoded Median Household Income					11.44			
RP80AQR	T: Quartic: Recoded Median Household Income					34.09			
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992	2.33				89.33			
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992					56.76			
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992					94.43			
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability					94.94	87.27		
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability					20.83	13.54		
IBLK260	Black Interaction Of B64DISLN						2.11		
BAGE34LN	B: Linear: Percent Persons 25-34 Years						0.01		
BAGE44LN	B: Linear: Percent Persons 35-44 Years					0.19			
BAGE54LN	B: Linear: Percent Persons 45-54 Years						59.98	51.67	0.09
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years						32.73	0.00	0.00
BAGE54CU	B: Cubic: Percent Persons 45-54 Years						0.25	4.57	0.00
BAGE54QR	B: Quartic: Percent Persons 45-54 Years							0.00	
IFEM275	Female Interaction Of BAGE54LN						65.53	94.23	22.39
IFEM276	Female Interaction Of BAGE54QU						99.78	91.20	0.00
IFEM277	Female Interaction Of BAGE54CU						0.00	20.49	0.00
IFEM278	Female Interaction Of BAGE54QR							0.00	
IHISP275	Hispanic Interaction Of BAGE54LN							20.63	
IHISP276	Hispanic Interaction Of BAGE54QU							77.50	
IHISP277	Hispanic Interaction Of BAGE54CU							5.36	
BAGE64LN	B: Linear: Percent Persons 55-64 Years	34.53	0.00	39.60	0.00	27.42	0.17		0.00
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		0.00	6.11	0.00	79.11	2.60		0.01
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		53.23	0.00		9.93	0.61		0.00
IFEM280	Female Interaction Of BAGE64LN	0.34			0.00	52.09			0.00
IFEM281	Female Interaction Of BAGE64QU				0.00	71.10			0.00
IFEM282	Female Interaction Of BAGE64CU					1.46			26.41
IBLK280	Black Interaction Of BAGE64LN								0.06
IBLK281	Black Interaction Of BAGE64QU								32.25
IHISP280	Hispanic Interaction Of BAGE64LN		0.12	71.49			87.64		
IHISP281	Hispanic Interaction Of BAGE64QU		56.69	12.48			72.95		
IHISP282	Hispanic Interaction Of BAGE64CU		0.00	0.00			32.52		
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander		74.24	51.28			3.98	33.12	0.08
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander		3.18	93.90				0.00	2.25
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander			0.01				85.32	0.06
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander			0.59					0.00
IBLK285	Black Interaction Of BASIANLN								26.24
IBLK286	Black Interaction Of BASIANQU								0.00
IBLK287	Black Interaction Of BASIANCU								0.14
IBLK288	Black Interaction Of BASIANQR								0.00
IHISP285	Hispanic Interaction Of BASIANLN			10.80				81.83	
IHISP286	Hispanic Interaction Of BASIANQU			0.08				48.99	
IHISP287	Hispanic Interaction Of BASIANCU			0.20				7.31	
IHISP288	Hispanic Interaction Of BASIANQR			0.00					
BCUBANLN	B: Linear: Percent Hispanics: Cuban						0.00	18.70	40.55
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban								0.00
BCUBANCU	B: Cubic: Percent Hispanics: Cuban								0.00
IFEM290	Female Interaction Of BCUBANLN							3.34	
IBLK290	Black Interaction Of BCUBANLN								90.40
IBLK291	Black Interaction Of BCUBANQU								0.20
IBLK292	Black Interaction Of BCUBANCU								0.00
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed	1.80		90.00			28.43	0.00	0.00

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Varibable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed	23.28		0.08			73.10	3.52	
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed	19.05					0.09		
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed	0.46					0.48		
IFEM295	Female Interaction Of BFNOTLN						88.46		
IFEM296	Female Interaction Of BFNOTQU						9.85		
IFEM297	Female Interaction Of BFNOTCU						0.00		
IBLK295	Black Interaction Of BFNOTLN			4.48				0.08	
IBLK296	Black Interaction Of BFNOTQU			0.00				0.02	
IHISP295	Hispanic Interaction Of BFNOTLN						92.14		
IHISP296	Hispanic Interaction Of BFNOTQU						46.17		
IHISP297	Hispanic Interaction Of BFNOTCU						79.42		
IHISP298	Hispanic Interaction Of BFNOTQR						24.17		
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut					6.79	1.99	0.00	1.04
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut					6.55			2.54
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut					0.33			
IBLK300	Black Interaction Of BINDIALN					0.24	45.47		
IHISP300	Hispanic Interaction Of BINDIALN								1.62
IHISP301	Hispanic Interaction Of BINDIAQU								0.00
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed				0.00	11.87	42.26	76.21	
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed				16.19	51.44	36.06	70.91	
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed				0.00			1.10	
IFEM305	Female Interaction Of BMNOTLN				0.00			2.44	
IFEM306	Female Interaction Of BMNOTQU				99.19			81.91	
IFEM307	Female Interaction Of BMNOTCU				0.00			0.10	
IHISP305	Hispanic Interaction Of BMNOTLN						67.77		
IHISP306	Hispanic Interaction Of BMNOTQU						17.44		
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.00				36.84		0.00
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		48.17						0.81
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		26.14						0.03
BPOVERQR	B: Quartic: Percent Families Below Poverty Level		0.73						0.00
IFEM310	Female Interaction Of BPOVERLN						0.51		1.45
IFEM311	Female Interaction Of BPOVERQU								10.93
IFEM312	Female Interaction Of BPOVERCU								0.00
IFEM313	Female Interaction Of BPOVERQR								0.00
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	9.63	0.00			12.61	80.94	23.33	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican		0.00			7.30	89.16	7.76	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					86.81	2.71	48.66	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican					76.78			
IFEM315	Female Interaction Of BPRICALN					66.56			
IFEM316	Female Interaction Of BPRICAQU					25.37			
IFEM317	Female Interaction Of BPRICACU					61.94			
IFEM318	Female Interaction Of BPRICAQR					5.13			
IBLK315	Black Interaction Of BPRICALN		36.24				36.64	16.42	
IBLK316	Black Interaction Of BPRICAQU		0.00				5.82	48.41	
IBLK317	Black Interaction Of BPRICACU							7.66	
IHISP315	Hispanic Interaction Of BPRICALN						33.73		
IHISP316	Hispanic Interaction Of BPRICAQU						88.73		
IHISP317	Hispanic Interaction Of BPRICACU						5.73		
BSCHASLN	B: Linear: Percent Associates Degree		0.20			29.12	42.56	21.96	0.00
BSCHASQU	B: Quadratic: Percent Associates Degree		15.77			47.22	2.05	6.83	0.15
BSCHASCU	B: Cubic: Percent Associates Degree					77.76	9.98	0.01	20.91
BSCHASQR	B: Quartic: Percent Associates Degree					12.90			
IFEM320	Female Interaction Of BSCHASLN					33.96			
IFEM321	Female Interaction Of BSCHASQU					15.65			
IFEM322	Female Interaction Of BSCHASCU					41.80			
IFEM323	Female Interaction Of BSCHASQR					0.79			
IBLK320	Black Interaction Of BSCHASLN					73.42		64.11	
IBLK321	Black Interaction Of BSCHASQU					19.99		27.97	
IBLK322	Black Interaction Of BSCHASCU					2.19		15.99	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander				0.00		3.35	94.00	
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander						99.45	0.50	
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander						0.43	0.00	
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander						0.52		
IFEM325	Female Interaction Of PASIANLN						42.13		
IFEM326	Female Interaction Of PASIANQU						11.45		
IFEM327	Female Interaction Of PASIANCU						67.28		
IFEM328	Female Interaction Of PASIANQR						0.07		
IHISP325	Hispanic Interaction Of PASIANLN						75.83		
IHISP326	Hispanic Interaction Of PASIANQU						26.60		
IHISP327	Hispanic Interaction Of PASIANCU						8.32		
PCUBANLN	T: Linear: Percent Hispanics: Cuban	26.15		0.16		22.66		63.25	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban	0.38		21.33		5.21		5.02	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban	5.07				16.69		68.25	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban							47.64	
IFEM330	Female Interaction Of PCUBANLN					38.74			
IFEM331	Female Interaction Of PCUBANQU					20.49			
IFEM332	Female Interaction Of PCUBANCU					0.49			
IBLK330	Black Interaction Of PCUBANLN	85.50							
IBLK331	Black Interaction Of PCUBANQU	41.03							
IBLK332	Black Interaction Of PCUBANCU	57.84							
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut	24.01		57.41			44.62		0.00
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut	0.56		0.05					70.93
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut	0.02							25.32

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

		BIG CITY				REMAINDER			
Varibable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut								0.00
IFEM335	Female Interaction Of PINDIALN								0.70
IFEM336	Female Interaction Of PINDIAQU								0.18
IFEM337	Female Interaction Of PINDIACU								1.03
IFEM338	Female Interaction Of PINDIAQR								0.00
IHISP335	Hispanic Interaction Of PINDIALN			0.11					
IHISP336	Hispanic Interaction Of PINDIAQU			0.00					

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F3. Significance Probabilities for Fixed Effect Coefficients for Past Month Cigarette Use by Age

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	18.66	0.00	0.00	0.00	69.46	10.47	0.01	0.00
FEMBLCK	O: Black Interaction Of FEMALE	0.08	22.30	0.00	0.00	95.57	12.13	0.24	6.07
FEMHISP	O: Hispanic Interaction Of FEMALE	22.71	0.00	0.00	0.00			0.00	0.26
FEMOTHR	O: Other Interaction Of FEMALE	61.50	0.00	0.00	0.00	63.05	0.00	0.00	33.73
RACEBLCK	O: Race/Black Indicator	0.00	0.00	16.43	0.00	0.01	0.00	42.15	0.00
RACEHISP	O: Race/Hispanic Indicator	0.01	0.00	0.00	0.00			51.45	14.68
RACEOTHR	O: Race/Other Indicator	0.00	0.00	26.30	0.00	0.04	96.08	12.02	0.00
REGNOREA	O: Northeast Region Indicator						0.69		
REGSOUTH	O: South Region Indicator						45.28		
REGWEST	O: West Region Indicator						62.65		
PDENLEV1	O: Large MSA					27.60	18.67	10.88	
PDENLEV2	O: Medium MSA					0.35	32.48	4.90	
PDENLEV3	O: Small MSA					4.30	26.97	88.49	
PDENLEV4	O: NonMSA, Urban					0.50	95.89	83.20	
IHISP10	Hispanic Interaction Of PDENLEV1							22.94	
IHISP11	Hispanic Interaction Of PDENLEV2							40.45	
POPRMLN	T: Linear: Average Persons Per Room			0.00		79.47		48.67	
POPRMQU	T: Quadratic: Average Persons Per Room					26.43			
POPRMCU	T: Cubic: Average Persons Per Room					3.06			
IFEM20	Female Interaction Of POPRMLN							0.09	
IBLK20	Black Interaction Of POPRMLN					58.00			
IBLK21	Black Interaction Of POPRMQU					65.78			
IBLK22	Black Interaction Of POPRMCU					2.41			
IOTH20	Other Interaction Of POPRMLN					1.82			
IOTH21	Other Interaction Of POPRMQU					57.88			
IOTH22	Other Interaction Of POPRMCU					0.33			
PAGE18LN	T: Linear: Percent Persons 0-18 Years		0.00				0.00		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years						27.41		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years						0.00		
IFEM25	Female Interaction Of PAGE18LN		0.00				0.88		
IFEM26	Female Interaction Of PAGE18QU						2.96		
IFEM27	Female Interaction Of PAGE18CU						0.00		
PAGE24LN	T: Linear: Percent Persons 19-24 Years					6.32		20.06	21.56
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years					0.62		0.15	0.00
PAGE24CU	T: Cubic: Percent Persons 19-24 Years							0.95	0.17
PAGE24QR	T: Quartic: Percent Persons 19-24 Years							5.18	0.00
IHISP30	Hispanic Interaction Of PAGE24LN							1.23	
IHISP31	Hispanic Interaction Of PAGE24QU							8.94	
IHISP32	Hispanic Interaction Of PAGE24CU							85.50	
IHISP33	Hispanic Interaction Of PAGE24QR							36.06	
IOTH30	Other Interaction Of PAGE24LN								0.00
IOTH31	Other Interaction Of PAGE24QU								0.02
IOTH32	Other Interaction Of PAGE24CU								0.42
IOTH33	Other Interaction Of PAGE24QR								0.00
PAGE44LN	T: Linear: Percent Persons 35-44 Years							14.74	
PAGE54LN	T: Linear: Percent Persons 45-54 Years				0.00	51.19			0.00
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years				72.44	1.40			0.00
IBLK45	Black Interaction Of PAGE54LN				0.00	91.08			
IBLK46	Black Interaction Of PAGE54QU				0.00	6.12			
PAGE64LN	T: Linear: Percent Persons 55-64 Years					15.22			
PAGE64QU	T: Quadratic: Percent Persons 55-64 Years					41.91			
PAGE64CU	T: Cubic: Percent Persons 55-64 Years					74.00			
PAGE64QR	T: Quartic: Percent Persons 55-64 Years					0.18			
PSCH8LN	T: Linear: Percent 0-8 Years Of School					47.40			
IOTH55	Other Interaction Of PSCH8LN					0.24			
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma					55.63			
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma					30.15			
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma					10.93			
PSCHASLN	T: Linear: Percent Associates Degree					56.16		53.02	
PSCHASQU	T: Quadratic: Percent Associates Degree					22.97		77.20	
PSCHASCU	T: Cubic: Percent Associates Degree					41.91			
PSCHASQR	T: Quartic: Percent Associates Degree					0.00			
IFEM65	Female Interaction Of PSCHASLN					35.64			
IFEM66	Female Interaction Of PSCHASQU					8.51			
IFEM67	Female Interaction Of PSCHASCU					91.64			
IFEM68	Female Interaction Of PSCHASQR					0.02			
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree					14.46			0.00
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree					18.11			0.00
IFEM75	Female Interaction Of PSCHCOLN					6.27			
IFEM76	Female Interaction Of PSCHCOQU					0.82			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
PPOVERLN	T: Linear: Percent Families Below Poverty Level					77.08			0.00
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level					34.98			0.00
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					9.53			
PPOVERQR	T: Quartic: Percent Families Below Poverty Level					89.49			
IFEM85	Female Interaction Of PPOVERLN					14.84			0.00
IFEM86	Female Interaction Of PPOVERQU					19.79			0.00
IFEM87	Female Interaction Of PPOVERCU					80.14			
IFEM88	Female Interaction Of PPOVERQR					0.03			
IOTH85	Other Interaction Of PPOVERLN					31.27			
IOTH86	Other Interaction Of PPOVERQU					14.67			
IOTH87	Other Interaction Of PPOVERCU					0.98			
PPUBASLN	T: Linear: Percent Households With Public Assist Income					11.10			
IFEM90	Female Interaction Of PPUBASLN					1.79			
P64DISLN	T: Linear: Percent 16-64 With A Work Disability	0.68				35.32			
IOTH95	Other Interaction Of P64DISLN	0.59				6.36			
PBLACKLN	T: Linear: Percent Black Nonhispanic			0.87				9.66	0.01
PBLACKQU	T: Quadratic: Percent Black Nonhispanic			0.59				0.04	0.00
PBLACKCU	T: Cubic: Percent Black Nonhispanic			0.03				0.77	
PHISPLN	T: Linear: Percent Hispanic					36.24			0.00
PHISPQU	T: Quadratic: Percent Hispanic					37.78			0.00
PHISPCU	T: Cubic: Percent Hispanic					26.20			0.00
PHISPQR	T: Quartic: Percent Hispanic								0.00
POTHLN	T: Linear: Percent Other Race/Hispanicity					0.02			0.00
POTHQU	T: Quadratic: Percent Other Race/Hispanicity					84.46			
POTHCU	T: Cubic: Percent Other Race/Hispanicity					88.27			
IOTH110	Other Interaction Of POTHLN								0.00
PHHF18LN	T: Linear: % Female-Headed HH W/No Spouse & Chld Under 18					50.39			0.00
PHHF18QU	T: Quadratic: % Female-Headed HH W/No Spouse & Chld Under 18					99.10			
PHHF18CU	T: Cubic: % Female-Headed HH W/No Spouse & Chld Under 18					26.23			
PHHF18QR	T: Quartic: % Female-Headed HH W/No Spouse & Chld Under 18					81.69			
IFEM120	Female Interaction Of PHHF18LN								0.00
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force						2.11		
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force						78.75		
PFLABCU	T: Cubic: Percent Females 16+ Years Old In Labor Force						0.02		
PFLABQR	T: Quartic: Percent Females 16+ Years Old In Labor Force						0.02		
IBLK125	Black Interaction Of PFLABLN						93.27		
IBLK126	Black Interaction Of PFLABQU						55.82		
IBLK127	Black Interaction Of PFLABCU						43.02		
IBLK128	Black Interaction Of PFLABQR						1.47		
PFNEVLN	T: Linear: Percent Females Never Married								0.00
PFNEVQU	T: Quadratic: Percent Females Never Married								0.00
PFNEVCU	T: Cubic: Percent Females Never Married								0.01
PFNEVQR	T: Quartic: Percent Females Never Married								0.00
IHISP130	Hispanic Interaction Of PFNEVLN								96.53
IHISP131	Hispanic Interaction Of PFNEVQU								9.53
IHISP132	Hispanic Interaction Of PFNEVCU								88.95
IHISP133	Hispanic Interaction Of PFNEVQR								0.00
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed		0.00	91.65			32.64	5.49	
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed						78.97	18.06	
PFNOTCU	T: Cubic: Percent Females Separated, Divorced Or Widowed						28.17	0.07	
PFNOTQR	T: Quartic: Percent Females Separated, Divorced Or Widowed						8.59	0.03	
IHISP135	Hispanic Interaction Of PFNOTLN							79.07	
IHISP136	Hispanic Interaction Of PFNOTQU							79.51	
IHISP137	Hispanic Interaction Of PFNOTCU							58.54	
IHISP138	Hispanic Interaction Of PFNOTQR							7.09	
IOTH135	Other Interaction Of PFNOTLN						36.03		
IOTH136	Other Interaction Of PFNOTQU						23.14		
IOTH137	Other Interaction Of PFNOTCU						0.10		
IOTH138	Other Interaction Of PFNOTQR						0.00		
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force					56.43			
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force					1.69			
IOTH145	Other Interaction Of PMLABLN					49.23			
IOTH146	Other Interaction Of PMLABQU					3.23			
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed						61.54	0.00	
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed						52.04	62.39	
IFEM150	Female Interaction Of PMNOTLN						1.01		
IFEM151	Female Interaction Of PMNOTQU						0.47		
IHISP150	Hispanic Interaction Of PMNOTLN							15.03	
IHISP151	Hispanic Interaction Of PMNOTQU							70.62	
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier				0.00				0.00
P40HULN	T: Linear: Percent Housing Units Built 1940-1949			10.21		65.83		32.36	
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949			18.70		63.79		24.99	
P40HUCU	T: Cubic: Percent Housing Units Built 1940-1949			99.98		7.17		85.11	
P40HUQR	T: Quartic: Percent Housing Units Built 1940-1949							0.86	
IBLK160	Black Interaction Of P40HULN			5.44				6.17	
IBLK161	Black Interaction Of P40HUQU			0.15				98.27	
IBLK162	Black Interaction Of P40HUCU			0.00				50.56	
IBLK163	Black Interaction Of P40HUQR							0.85	
IOTH160	Other Interaction Of P40HULN					7.60			
IOTH161	Other Interaction Of P40HUQU					0.22			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
IOTH162	Other Interaction Of P40HUCU					6.37			
PRENTLN	T: Linear: Percent Housing Rented						1.89	0.00	9.86
PRENTQU	T: Quadratic: Percent Housing Rented						24.93		50.22
PRENTCU	T: Cubic: Percent Housing Rented						0.10		7.39
PRENTQR	T: Quartic: Percent Housing Rented								48.40
IOTH165	Other Interaction Of PRENTLN						68.28		
IOTH166	Other Interaction Of PRENTQU						0.01		
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases					42.91			
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases					17.55			
ADRATECU	C: Cubic: Death Rate For All Alcohol-Related Cases					2.84			
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol		1.34				67.84		0.00
ADRAT1QU	C: Quadratic: Death Rate With Explicit Mention Of Alcohol		78.46				34.11		0.00
ADRAT1CU	C: Cubic: Death Rate With Explicit Mention Of Alcohol		82.61				72.98		0.00
ADRAT1QR	C: Quartic: Death Rate With Explicit Mention Of Alcohol		97.97				0.11		0.00
IBLK180	Black Interaction Of ADRAT1LN		10.99				19.41		
IBLK181	Black Interaction Of ADRAT1QU		12.54				6.46		
IBLK182	Black Interaction Of ADRAT1CU		3.63				44.45		
IBLK183	Black Interaction Of ADRAT1QR		0.00				0.16		
V18FLN	C: Linear: Marijuana Posession Arrest Rate			0.00		64.41		1.52	
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate							52.39	
V18FCU	C: Cubic: Marijuana Posession Arrest Rate							38.64	
V18FQR	C: Quartic: Marijuana Posession Arrest Rate							0.00	
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate				0.00			40.04	
IBLK190	Black Interaction Of V18BLN				0.00			0.01	
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate						80.34		0.00
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate						98.09		60.42
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate						0.21		15.33
V18EQR	C: Quartic: Opium/Cocaine & Deriv Posession Arrest Rate								0.00
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate		25.97				1.57	63.12	0.00
IBLK200	Black Interaction Of V18ALN								0.00
IHISP200	Hispanic Interaction Of V18ALN		0.00						
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate		45.24	0.44			6.19	7.80	
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate			0.00			27.56	0.08	
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate			0.00				0.00	
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate			0.19				6.22	
IFEM225	Female Interaction Of V18LN							0.02	
IFEM226	Female Interaction Of V18QU							2.36	
IFEM227	Female Interaction Of V18CU							6.57	
IFEM228	Female Interaction Of V18QR							0.43	
IHISP225	Hispanic Interaction Of V18LN		5.53						
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate				0.00				0.00
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate				75.90				0.00
IBLK230	Black Interaction Of VIOLLN				0.00				0.00
IBLK231	Black Interaction Of VIOLQU				0.00				
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992					19.26	0.41		0.89
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						90.70		6.47
IFEM235	Female Interaction Of DRATELN						15.73		
IFEM236	Female Interaction Of DRATEQU						22.31		
IOTH235	Other Interaction Of DRATELN								0.00
IOTH236	Other Interaction Of DRATEQU								0.00
RH43ALN	T: Linear: Recoded Median Rents For Rental Units	66.83				14.95		0.32	
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units	1.16				4.50			
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units	97.17				0.00			
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units	90.74							
RP80ALN	T: Linear: Recoded Median Household Income		0.09	0.00		47.91	91.18	0.00	2.10
RP80AQU	T: Quadratic: Recoded Median Household Income					82.14	1.25		0.00
RP80ACU	T: Cubic: Recoded Median Household Income					38.84	8.12		
RP80AQR	T: Quartic: Recoded Median Household Income					0.45			
IBLK250	Black Interaction Of RP80ALN						56.45		
IBLK251	Black Interaction Of RP80AQU						84.11		
IBLK252	Black Interaction Of RP80ACU						93.86		
IHISP250	Hispanic Interaction Of RP80ALN		0.00						28.82
IHISP251	Hispanic Interaction Of RP80AQU								0.00
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992	14.49				29.55			
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992					24.32			
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992					32.00			
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability			0.00	0.00	87.78	0.00	0.00	
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability					6.06			
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability					46.04			
B64DISQR	B: Quartic: Percent Persons 16-64 With A Work Disability					5.35			
IFEM260	Female Interaction Of B64DISLN					75.29			
IFEM261	Female Interaction Of B64DISQU					9.25			
IFEM262	Female Interaction Of B64DISCU					16.93			
IFEM263	Female Interaction Of B64DISQR					0.02			
BAGE34LN	B: Linear: Percent Persons 25-34 Years				0.00	88.83		0.28	0.00
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years					2.08		0.03	0.00
BAGE34CU	B: Cubic: Percent Persons 25-34 Years							6.87	0.00
IFEM265	Female Interaction Of BAGE34LN								0.00
IFEM266	Female Interaction Of BAGE34QU								0.00

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IFEM267	Female Interaction Of BAGE34CU								0.00
IBLK265	Black Interaction Of BAGE34LN								0.00
IOTH265	Other Interaction Of BAGE34LN					1.63		0.00	
IOTH266	Other Interaction Of BAGE34QU					0.50		0.00	
IOTH267	Other Interaction Of BAGE34CU							0.00	
BAGE44LN	B: Linear: Percent Persons 35-44 Years							5.70	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years							0.00	
IHISP270	Hispanic Interaction Of BAGE44LN							25.74	
IHISP271	Hispanic Interaction Of BAGE44QU							1.72	
BAGE54LN	B: Linear: Percent Persons 45-54 Years					64.51	95.57	0.00	1.65
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years					90.79	0.17	0.88	0.00
BAGE54CU	B: Cubic: Percent Persons 45-54 Years					14.56	0.00	5.08	0.00
BAGE54QR	B: Quartic: Percent Persons 45-54 Years							7.17	
IFEM275	Female Interaction Of BAGE54LN						37.51		85.96
IFEM276	Female Interaction Of BAGE54QU						2.09		0.00
IFEM277	Female Interaction Of BAGE54CU						0.00		14.88
IBLK275	Black Interaction Of BAGE54LN						94.22		
IBLK276	Black Interaction Of BAGE54QU						80.74		
IOTH275	Other Interaction Of BAGE54LN					16.67		38.98	
IOTH276	Other Interaction Of BAGE54QU					80.34		36.63	
IOTH277	Other Interaction Of BAGE54CU					0.77		0.10	
IOTH278	Other Interaction Of BAGE54QR							0.00	
BAGE64LN	B: Linear: Percent Persons 55-64 Years		36.79	56.55	0.00	0.91	29.28		0.00
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		0.00	19.32	0.00		62.11		68.56
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		49.19	58.04			5.82		0.00
BAGE64QR	B: Quartic: Percent Persons 55-64 Years						11.28		
IFEM280	Female Interaction Of BAGE64LN				0.00				0.00
IFEM281	Female Interaction Of BAGE64QU				24.61				0.00
IFEM282	Female Interaction Of BAGE64CU								92.42
IBLK280	Black Interaction Of BAGE64LN								1.21
IBLK281	Black Interaction Of BAGE64QU								0.76
IHISP280	Hispanic Interaction Of BAGE64LN		89.38	0.00					
IHISP281	Hispanic Interaction Of BAGE64QU		67.45	1.13					
IHISP282	Hispanic Interaction Of BAGE64CU		2.47	53.65					
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander		0.08	69.16			95.06	0.00	0.00
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander		9.72	32.30			1.60	91.23	7.44
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander			35.49			0.05	1.46	0.00
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander			30.23			17.13	0.00	
IFEM285	Female Interaction Of BASIANLN						34.28		0.07
IFEM286	Female Interaction Of BASIANQU						33.07		45.23
IFEM287	Female Interaction Of BASIANCU						12.38		0.00
IFEM288	Female Interaction Of BASIANQR						0.03		
IHISP285	Hispanic Interaction Of BASIANLN			86.33				1.20	
IHISP286	Hispanic Interaction Of BASIANQU			0.00					
IHISP287	Hispanic Interaction Of BASIANCU			86.51					
IHISP288	Hispanic Interaction Of BASIANQR			0.00					
BCUBANLN	B: Linear: Percent Hispanics: Cuban						12.06	20.62	
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban						37.38		
BCUBANCU	B: Cubic: Percent Hispanics: Cuban						7.35		
IFEM290	Female Interaction Of BCUBANLN							7.19	
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed					62.65		79.84	
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed					82.60		0.00	
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed					10.60			
IFEM295	Female Interaction Of BFNOTLN					91.77		21.84	
IFEM296	Female Interaction Of BFNOTQU					35.85		0.00	
IFEM297	Female Interaction Of BFNOTCU					0.08			
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut		0.00	80.41			1.26		
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			0.77			42.08		
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut						85.33		
BINDIAQR	B: Quartic: Percent Pop: American Indian, Eskimo, Aleut						45.55		
IBLK300	Black Interaction Of BINDIALN						5.29		
IHISP300	Hispanic Interaction Of BINDIALN			0.00					
IHISP301	Hispanic Interaction Of BINDIAQU			0.00					
IOTH300	Other Interaction Of BINDIALN						0.01		
IOTH301	Other Interaction Of BINDIAQU						0.00		
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed			0.00	0.00	64.82	0.38		0.00
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed			0.63	0.00	0.18	98.04		0.00
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed			3.27	0.00		8.19		
BMNOTQR	B: Quartic: Percent Males Separated, Divorced or Widowed			0.00					
IFEM305	Female Interaction Of BMNOTLN				0.00	68.58			
IFEM306	Female Interaction Of BMNOTQU				0.00	0.15			
IFEM307	Female Interaction Of BMNOTCU				0.29				
IHISP305	Hispanic Interaction Of BMNOTLN			56.38					11.40
IHISP306	Hispanic Interaction Of BMNOTQU			42.61					0.00
IHISP307	Hispanic Interaction Of BMNOTCU			6.37					
IHISP308	Hispanic Interaction Of BMNOTQR			0.00					
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.00			91.59	10.32		0.00
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.15			64.98	23.28		0.00
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		0.00			19.25	0.13		0.00
BPOVERQR	B: Quartic: Percent Families Below Poverty Level		0.98				43.51		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IFEM310	Female Interaction Of BPOVERLN								0.00
IFEM311	Female Interaction Of BPOVERQU								9.04
IFEM312	Female Interaction Of BPOVERCU								0.00
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	30.32				33.42		0.00	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican					41.14		64.40	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					0.19		1.61	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican							0.00	
IFEM315	Female Interaction Of BPRICALN					23.63			
IBLK315	Black Interaction Of BPRICALN							7.33	
IBLK316	Black Interaction Of BPRICAQU							32.76	
IBLK317	Black Interaction Of BPRICACU							77.74	
IBLK318	Black Interaction Of BPRICAQR							2.99	
BSCHASLN	B: Linear: Percent Associates Degree		0.49			72.59	1.36	86.58	0.14
BSCHASQU	B: Quadratic: Percent Associates Degree		0.00			14.45	76.34	2.27	0.00
BSCHASCU	B: Cubic: Percent Associates Degree						22.74	48.17	0.00
BSCHASQR	B: Quartic: Percent Associates Degree						0.84		
IFEM320	Female Interaction Of BSCHASLN					3.88	0.60		
IFEM321	Female Interaction Of BSCHASQU					6.02			
IBLK320	Black Interaction Of BSCHASLN							62.08	
IBLK321	Black Interaction Of BSCHASQU							0.17	
IBLK322	Black Interaction Of BSCHASCU							94.56	
IHISP320	Hispanic Interaction Of BSCHASLN		0.03						
IHISP321	Hispanic Interaction Of BSCHASQU		0.00						
IOTH320	Other Interaction Of BSCHASLN								0.00
IOTH321	Other Interaction Of BSCHASQU								0.00
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander			74.91		9.33	46.98		
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander			3.31		90.98	38.53		
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander					90.90	2.13		
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander						0.03		
IBLK325	Black Interaction Of PASIANLN			0.01					
IBLK326	Black Interaction Of PASIANQU			0.00					
IOTH325	Other Interaction Of PASIANLN					39.46			
IOTH326	Other Interaction Of PASIANQU					65.04			
IOTH327	Other Interaction Of PASIANCU					4.16			
PCUBANLN	T: Linear: Percent Hispanics: Cuban	83.27		16.22	0.00	91.85	0.00	5.44	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban	8.02		0.00		19.63		57.08	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban	5.83		8.52		21.13		19.76	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban			0.00				44.74	
IFEM330	Female Interaction Of PCUBANLN	30.49				44.89			
IFEM331	Female Interaction Of PCUBANQU	0.01				59.93			
IFEM332	Female Interaction Of PCUBANCU					0.61			
IBLK330	Black Interaction Of PCUBANLN	0.23					3.24	65.59	
IBLK331	Black Interaction Of PCUBANQU	50.61						9.17	
IBLK332	Black Interaction Of PCUBANCU	14.67						40.84	
IBLK333	Black Interaction Of PCUBANQR							1.82	
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut			0.00		63.26		14.08	0.01
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut							74.48	0.00
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut							1.46	20.53
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut							19.00	0.00
IFEM335	Female Interaction Of PINDIALN							74.73	
IFEM336	Female Interaction Of PINDIAQU							97.48	
IFEM337	Female Interaction Of PINDIACU							48.90	
IFEM338	Female Interaction Of PINDIAQR							0.00	
IBLK335	Black Interaction Of PINDIALN					1.57		17.96	
IBLK336	Black Interaction Of PINDIAQU							1.05	
IHISP335	Hispanic Interaction Of PINDIALN							52.44	
IHISP336	Hispanic Interaction Of PINDIAQU							35.86	
IHISP337	Hispanic Interaction Of PINDIACU							82.11	
IHISP338	Hispanic Interaction Of PINDIAQR							0.72	
IOTH335	Other Interaction Of PINDIALN							0.00	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F4. Significance Probabilities for Fixed Effect Coefficients for Past Month Cocaine Use by Age

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	1.88	0.00	0.00	0.00	68.84	0.00	10.68	47.23
FEMBLCK	O: Black Interaction Of FEMALE	0.25	64.10	0.98	0.00	88.15	44.96	30.22	0.00
FEMHISP	O: Hispanic Interaction Of FEMALE	28.95	31.40	21.27	0.01	76.91	48.80	31.16	1.47
RACEBLCK	O: Race/Black Indicator	30.25	10.25	6.69	0.00	61.05	47.58	51.66	0.00
RACEHISP	O: Race/Hispanic Indicator	2.71	0.00	0.00	0.00	8.08	36.90	1.74	0.00
RACEOTHR	O: Race/Other Indicator	17.11	0.00				46.97	88.74	
PDENLEV1	O: Large MSA						51.01	62.66	
PDENLEV2	O: Medium MSA						28.49	28.80	
PDENLEV3	O: Small MSA						9.48	29.55	
PDENLEV4	O: NonMSA, Urban						0.14	22.58	
POPRMLN	T: Linear: Average Persons Per Room			0.00				0.00	
IFEM20	Female Interaction Of POPRMLN							0.00	
PAGE24LN	T: Linear: Percent Persons 19-24 Years							6.30	
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years							95.50	
PAGE24CU	T: Cubic: Percent Persons 19-24 Years							0.00	
PAGE24QR	T: Quartic: Percent Persons 19-24 Years							0.00	
IHISP30	Hispanic Interaction Of PAGE24LN							56.68	
IHISP31	Hispanic Interaction Of PAGE24QU							95.46	
IHISP32	Hispanic Interaction Of PAGE24CU							5.38	
IHISP33	Hispanic Interaction Of PAGE24QR							7.99	
PAGE44LN	T: Linear: Percent Persons 35-44 Years							0.00	
PAGE54LN	T: Linear: Percent Persons 45-54 Years				0.00				17.83
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years				0.00				0.00
IBLK45	Black Interaction Of PAGE54LN				0.00				
IBLK46	Black Interaction Of PAGE54QU				0.00				
PSCHASLN	T: Linear: Percent Associates Degree							53.87	
PSCHASQU	T: Quadratic: Percent Associates Degree							0.00	
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree					52.13			1.98
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree					5.87			0.00
IFEM75	Female Interaction Of PSCHCOLN					15.29			
IFEM76	Female Interaction Of PSCHCOQU					3.00			
PBLACKLN	T: Linear: Percent Black Nonhispanic			0.00				7.45	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic			9.18				22.48	
PBLACKCU	T: Cubic: Percent Black Nonhispanic			0.00				0.00	
PHISPLN	T: Linear: Percent Hispanic								10.76
PHISPQU	T: Quadratic: Percent Hispanic								1.98
PHISPCU	T: Cubic: Percent Hispanic								18.05
PHISPQR	T: Quartic: Percent Hispanic								0.00
POTHLN	T: Linear: Percent Other Race/Hispanicity					81.78			
POTHQU	T: Quadratic: Percent Other Race/Hispanicity					27.78			
POTHCU	T: Cubic: Percent Other Race/Hispanicity					13.38			
PHHF18LN	T: Linear: % Female-Headed HH W/No Spouse & Chld Under 18					31.22			0.00
PHHF18QU	T: Quadratic: % Female-Headed HH W/No Spouse & Chld Under 18					65.13			
PHHF18CU	T: Cubic: % Female-Headed HH W/No Spouse & Chld Under 18					47.40			
PHHF18QR	T: Quartic: % Female-Headed HH W/No Spouse & Chld Under 18					2.90			
IFEM120	Female Interaction Of PHHF18LN								0.00
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed			0.00				0.74	
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed							0.00	
IHISP135	Hispanic Interaction Of PFNOTLN							21.71	
IHISP136	Hispanic Interaction Of PFNOTQU							3.36	
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed						2.29	34.05	
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed						0.04	0.32	
IFEM150	Female Interaction Of PMNOTLN						0.00		
IFEM151	Female Interaction Of PMNOTQU						0.06		
IHISP150	Hispanic Interaction Of PMNOTLN							82.93	
IHISP151	Hispanic Interaction Of PMNOTQU							7.03	
PRENTLN	T: Linear: Percent Housing Rented								0.00
PRENTQU	T: Quadratic: Percent Housing Rented								0.00
PRENTCU	T: Cubic: Percent Housing Rented								0.00
PRENTQR	T: Quartic: Percent Housing Rented								0.00
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate							7.03	
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate							65.75	
V18BCU	C: Cubic: Marijuana Sale/Manufacture Arrest Rate							0.00	
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate								0.50
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate								0.00
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate								10.98
V18EQR	C: Quartic: Opium/Cocaine & Deriv Posession Arrest Rate								0.00
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate		0.00					0.33	
IHISP200	Hispanic Interaction Of V18ALN		0.01						

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	18-25	26-34	35+	REMAINDER 12-17	18-25	26-34	35+
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate		0.00	3.35			67.38	0.00	
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate			86.91			6.64	79.65	
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate			0.10				39.64	
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate			24.78				0.00	
IFEM225	Female Interaction Of V18LN							18.05	
IFEM226	Female Interaction Of V18QU							67.85	
IFEM227	Female Interaction Of V18CU							21.59	
IFEM228	Female Interaction Of V18QR							0.31	
IHISP225	Hispanic Interaction Of V18LN		0.00						
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate				0.00				0.03
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate				0.00				0.00
IBLK230	Black Interaction Of VIOLLN				0.00				
IBLK231	Black Interaction Of VIOLQU				0.00				
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992						42.31		
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						0.04		
IFEM235	Female Interaction Of DRATELN						57.92		
IFEM236	Female Interaction Of DRATEQU						0.01		
RH43ALN	T: Linear: Recoded Median Rents For Rental Units	48.95						0.00	
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units	48.97							
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units	1.63							
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units	1.27							
RP80ALN	T: Linear: Recoded Median Household Income					91.39	71.40	0.00	
RP80AQU	T: Quadratic: Recoded Median Household Income					21.16	8.19		
RP80ACU	T: Cubic: Recoded Median Household Income					92.55	0.01		
RP80AQR	T: Quartic: Recoded Median Household Income					40.19			
IBLK250	Black Interaction Of RP80ALN						30.82		
IBLK251	Black Interaction Of RP80AQU						40.66		
IBLK252	Black Interaction Of RP80ACU						5.03		
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992	0.58							
BAGE54LN	B: Linear: Percent Persons 45-54 Years						77.16		0.00
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years						0.00		0.00
BAGE54CU	B: Cubic: Percent Persons 45-54 Years								0.00
IFEM275	Female Interaction Of BAGE54LN								0.00
IFEM276	Female Interaction Of BAGE54QU								0.00
IFEM277	Female Interaction Of BAGE54CU								0.00
IBLK275	Black Interaction Of BAGE54LN						93.61		
IBLK276	Black Interaction Of BAGE54QU						0.34		
BAGE64LN	B: Linear: Percent Persons 55-64 Years		0.87	36.83	0.00		0.05		95.01
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		0.01	79.44	0.00		39.57		0.00
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		18.22	0.45			0.95		0.00
BAGE64QR	B: Quartic: Percent Persons 55-64 Years						86.67		
IFEM280	Female Interaction Of BAGE64LN				36.61				0.00
IFEM281	Female Interaction Of BAGE64QU				0.00				0.00
IFEM282	Female Interaction Of BAGE64CU								0.00
IBLK280	Black Interaction Of BAGE64LN								1.49
IBLK281	Black Interaction Of BAGE64QU								0.00
IHISP280	Hispanic Interaction Of BAGE64LN		97.76	5.34					
IHISP281	Hispanic Interaction Of BAGE64QU		31.03	31.50					
IHISP282	Hispanic Interaction Of BAGE64CU		0.00	0.00					
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander		29.95				21.50		
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander		0.00				8.64		
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander						66.63		
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander						0.06		
BCUBANLN	B: Linear: Percent Hispanics: Cuban						83.19	0.03	
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban						48.05		
BCUBANCU	B: Cubic: Percent Hispanics: Cuban						0.19		
IFEM290	Female Interaction Of BCUBANLN							0.99	
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed				0.00	12.75	4.44		
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed				0.00	4.15	11.50		
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed				0.00		0.35		
IFEM305	Female Interaction Of BMNOTLN				0.00				
IFEM306	Female Interaction Of BMNOTQU				0.00				
IFEM307	Female Interaction Of BMNOTCU				0.00				
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.00				26.03		
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		3.93				63.23		
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		4.10				8.48		
BPOVERQR	B: Quartic: Percent Families Below Poverty Level		0.00				0.03		
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	16.63				84.59		0.41	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican							2.28	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican							0.07	
IFEM315	Female Interaction Of BPRICALN					51.12			
IBLK315	Black Interaction Of BPRICALN							94.69	
IBLK316	Black Interaction Of BPRICAQU							34.82	
IBLK317	Black Interaction Of BPRICACU							0.07	
BSCHASLN	B: Linear: Percent Associates Degree		0.16			51.87	3.29	0.44	17.63
BSCHASQU	B: Quadratic: Percent Associates Degree		0.00			10.31	44.14	34.23	22.26
BSCHASCU	B: Cubic: Percent Associates Degree							0.04	0.00
IFEM320	Female Interaction Of BSCHASLN					49.80	0.02		
IFEM321	Female Interaction Of BSCHASQU					5.39			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IBLK320	Black Interaction Of BSCHASLN							35.71	
IBLK321	Black Interaction Of BSCHASQU							28.54	
IBLK322	Black Interaction Of BSCHASCU							2.37	
PCUBANLN	T: Linear: Percent Hispanics: Cuban	6.13		0.00				21.02	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban	3.78		0.00				25.58	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban	21.45						5.33	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban							0.00	
IBLK330	Black Interaction Of PCUBANLN	60.28							
IBLK331	Black Interaction Of PCUBANQU	5.32							
IBLK332	Black Interaction Of PCUBANCU	4.58							

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F5. Significance Probabilities for Fixed Effect Coefficients for Past Month Any Illicit Drug But Marijuana by Age

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	29.69	0.00	0.00	0.00	19.09	0.00	0.05	15.97
FEMBLCK	O: Black Interaction Of FEMALE	7.24	27.46	16.55	0.00	15.03	42.92	27.75	0.00
FEMHISP	O: Hispanic Interaction Of FEMALE	71.27	0.24	0.02	0.00		99.81	92.54	77.59
FEMOTHR	O: Other Interaction Of FEMALE	48.96	0.00	95.28		37.39	23.59		0.00
RACEBLCK	O: Race/Black Indicator	5.49	0.00	0.00	0.00	6.81	2.52	11.62	0.00
RACEHISP	O: Race/Hispanic Indicator	90.88	0.00	0.00	0.00		17.13	30.45	13.65
RACEOTHR	O: Race/Other Indicator	1.31	0.00	0.00		33.65	0.00	0.04	0.00
REGNOREA	O: Northeast Region Indicator							0.05	
REGSOUTH	O: South Region Indicator							0.54	
REGWEST	O: West Region Indicator						0.02	0.00	
PDENLEV1	O: Large MSA								0.00
PDENLEV2	O: Medium MSA								62.05
PDENLEV3	O: Small MSA								0.00
PDENLEV4	O: NonMSA, Urban								76.27
PHH1PLN	T: Linear: Percent One Person Households					0.22			
PAGE24LN	T: Linear: Percent Persons 19-24 Years							0.70	0.00
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years							34.90	0.00
PAGE24CU	T: Cubic: Percent Persons 19-24 Years							0.00	
PAGE24QR	T: Quartic: Percent Persons 19-24 Years							0.00	
PAGE34LN	T: Linear: Percent Persons 25-34 Years					27.92			
PAGE34QU	T: Quadratic: Percent Persons 25-34 Years					0.04			
PAGE44LN	T: Linear: Percent Persons 35-44 Years					60.65			
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years					95.50			
PAGE54LN	T: Linear: Percent Persons 45-54 Years								0.00
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years								0.00
IFEM45	Female Interaction Of PAGE54LN								70.27
IFEM46	Female Interaction Of PAGE54QU								0.00
PSCH8LN	T: Linear: Percent 0-8 Years Of School							0.00	
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma	51.88				32.65			
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma	1.30				9.49			
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma	90.69				83.93			
PSCH12QR	T: Quartic: Percent 9-12 Years & No High School Diploma	0.02				1.53			
IFEM60	Female Interaction Of PSCH12LN	11.21				14.46			
IFEM61	Female Interaction Of PSCH12QU	0.10				1.89			
PSCHASLN	T: Linear: Percent Associates Degree					41.46			
PSCHASQU	T: Quadratic: Percent Associates Degree					8.40			
IBLK65	Black Interaction Of PSCHASLN					19.76			
IBLK66	Black Interaction Of PSCHASQU					1.30			
PSCHSCLN	T: Linear: Percent Some College And No Degree					87.11			
PSCHSCQU	T: Quadratic: Percent Some College And No Degree					0.96			
PPOVERLN	T: Linear: Percent Families Below Poverty Level							9.97	
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level							14.90	
PPOVERCU	T: Cubic: Percent Families Below Poverty Level							89.27	
PPOVERQR	T: Quartic: Percent Families Below Poverty Level							0.00	
P64DISLN	T: Linear: Percent 16-64 With A Work Disability	18.53				16.27			
P64DISQU	T: Quadratic: Percent 16-64 With A Work Disability	11.22				8.68			
P64DISCU	T: Cubic: Percent 16-64 With A Work Disability	3.37				12.12			
IBLK95	Black Interaction Of P64DISLN	2.12				68.77			
IBLK96	Black Interaction Of P64DISQU	92.69				18.19			
IBLK97	Black Interaction Of P64DISCU	0.14				1.25			
PBLACKLN	T: Linear: Percent Black Nonhispanic			0.24				0.39	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic			78.57				60.13	
PBLACKCU	T: Cubic: Percent Black Nonhispanic			0.00				0.00	
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed			0.00				0.00	
PMNEVLN	T: Linear: Percent Males Never Married					1.85			
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed					84.60			
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed					93.42			
PMNOTCU	T: Cubic: Percent Males Separated, Divorced Or Widowed					46.88			
PMNOTQR	T: Quartic: Percent Males Separated, Divorced Or Widowed					4.29			
P40HULN	T: Linear: Percent Housing Units Built 1940-1949					2.32		0.00	
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949							11.39	
P40HUCU	T: Cubic: Percent Housing Units Built 1940-1949							0.07	
P40HUQR	T: Quartic: Percent Housing Units Built 1940-1949							0.00	
PRENTLN	T: Linear: Percent Housing Rented							7.03	0.00
PRENTQU	T: Quadratic: Percent Housing Rented							0.00	
ADRATEQR	C: Quartic: Death Rate For All Alcohol-Related Cases						0.00		
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol						0.00		
V18FLN	C: Linear: Marijuana Posession Arrest Rate					20.21		0.04	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate					9.03		18.82	
V18FCU	C: Cubic: Marijuana Posession Arrest Rate							0.10	
V18FQR	C: Quartic: Marijuana Posession Arrest Rate							0.00	
IHISP185	Hispanic Interaction Of V18FLN							8.45	
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate								39.46
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate								0.00
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate					74.11			
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate					51.23			
V18ACU	C: Cubic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate					2.58			
V18AQR	C: Quartic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate					2.06			
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate		0.00						
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992						0.00		
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						0.00		
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability					39.13			
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability					8.72			
BAGE34LN	B: Linear: Percent Persons 25-34 Years					90.45	0.00	0.00	
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years					45.41		0.02	
BAGE34CU	B: Cubic: Percent Persons 25-34 Years					22.41		0.07	
IFEM265	Female Interaction Of BAGE34LN							0.78	
IFEM266	Female Interaction Of BAGE34QU							1.18	
IFEM267	Female Interaction Of BAGE34CU							0.02	
IOTH265	Other Interaction Of BAGE34LN					9.69		0.00	
IOTH266	Other Interaction Of BAGE34QU					8.78			
IOTH267	Other Interaction Of BAGE34CU					8.97			
BAGE44LN	B: Linear: Percent Persons 35-44 Years					0.37	28.15	89.84	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years						0.18	3.71	
IBLK270	Black Interaction Of BAGE44LN							77.59	
IBLK271	Black Interaction Of BAGE44QU							0.38	
IOTH270	Other Interaction Of BAGE44LN						3.83		
IOTH271	Other Interaction Of BAGE44QU						0.00		
BAGE54LN	B: Linear: Percent Persons 45-54 Years							28.21	
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years							0.00	
BAGE64LN	B: Linear: Percent Persons 55-64 Years		0.00	15.99		1.04	0.00		
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		0.01	85.44			1.25		
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		1.52	0.00			0.00		
BAGE64QR	B: Quartic: Percent Persons 55-64 Years						0.03		
IHISP280	Hispanic Interaction Of BAGE64LN		0.95	27.69			28.65		
IHISP281	Hispanic Interaction Of BAGE64QU		97.41	0.08			45.45		
IHISP282	Hispanic Interaction Of BAGE64CU		0.00	0.00			50.76		
IHISP283	Hispanic Interaction Of BAGE64QR						3.55		
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander						77.41	20.12	
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander						0.15	0.00	
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander						25.67		
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander						0.01		
IOTH285	Other Interaction Of BASIANLN						0.07		
BCUBANLN	B: Linear: Percent Hispanics: Cuban				0.00	22.52			
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban				0.00	17.04			
BCUBANCU	B: Cubic: Percent Hispanics: Cuban					47.13			
IFEM290	Female Interaction Of BCUBANLN					82.47			
IFEM291	Female Interaction Of BCUBANQU					97.63			
IFEM292	Female Interaction Of BCUBANCU					18.24			
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed	38.88				94.70	38.07		
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed	64.59				70.01	16.14		
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed	82.97				25.05	22.98		
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed	0.45				18.27			
IFEM295	Female Interaction Of BFNOTLN						3.10		
IFEM296	Female Interaction Of BFNOTQU						53.21		
IFEM297	Female Interaction Of BFNOTCU						0.01		
IBLK295	Black Interaction Of BFNOTLN					66.87	82.80		
IBLK296	Black Interaction Of BFNOTQU					86.29	0.27		
IBLK297	Black Interaction Of BFNOTCU					20.55			
IBLK298	Black Interaction Of BFNOTQR					8.56			
IHISP295	Hispanic Interaction Of BFNOTLN						7.58		
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut					22.06			
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut					35.55			
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut					0.16			
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed				0.00				14.90
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed								0.00
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed								0.00
IFEM305	Female Interaction Of BMNOTLN				0.00				
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.03				12.49		
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		1.67				0.01		
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		9.25				0.00		
BPOVERQR	B: Quartic: Percent Families Below Poverty Level		0.00				0.02		
IFEM310	Female Interaction Of BPOVERLN		12.14				0.22		
IFEM311	Female Interaction Of BPOVERQU		0.04				0.72		
IFEM312	Female Interaction Of BPOVERCU		0.00				0.00		
IHISP310	Hispanic Interaction Of BPOVERLN		0.00				91.32		
IHISP311	Hispanic Interaction Of BPOVERQU						2.41		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican	0.42				24.31		12.64	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican					6.37		0.06	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					26.72		0.00	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican					65.32			
IFEM315	Female Interaction Of BPRICALN					90.15		2.46	
IFEM316	Female Interaction Of BPRICAQU					3.78		26.43	
IFEM317	Female Interaction Of BPRICACU					50.82		0.01	
IFEM318	Female Interaction Of BPRICAQR					0.18			
BSCHASLN	B: Linear: Percent Associates Degree	61.98	0.00			63.54	88.36	0.73	
BSCHASQU	B: Quadratic: Percent Associates Degree	38.21				2.16	0.03	81.98	
BSCHASCU	B: Cubic: Percent Associates Degree	5.57				15.09		0.00	
BSCHASQR	B: Quartic: Percent Associates Degree	0.18				9.29			
IFEM320	Female Interaction Of BSCHASLN					74.76			
IFEM321	Female Interaction Of BSCHASQU					0.11			
IFEM322	Female Interaction Of BSCHASCU					3.39			
IFEM323	Female Interaction Of BSCHASQR					6.10			
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander						92.01	33.52	0.00
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander						19.27	7.88	
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander						0.01	0.00	
IHISP325	Hispanic Interaction Of PASIANLN						24.64		
IHISP326	Hispanic Interaction Of PASIANQU						42.91		
IHISP327	Hispanic Interaction Of PASIANCU						8.42		
IOTH325	Other Interaction Of PASIANLN						0.01		
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut						7.72	0.03	
IHISP335	Hispanic Interaction Of PINDIALN							1.37	
IOTH335	Other Interaction Of PINDIALN						0.00		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F6. Significance Probabilities for Fixed Effect Coefficients for Past Year Alcohol Treatment by Age

		BIG CITY				REMAINDER			
Variable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	99.00	0.00	0.00	0.00	85.66	0.34	0.00	0.00
FEMBLCK	O: Black Interaction Of FEMALE	94.17		21.49	0.00	81.87	63.81	58.55	0.01
FEMHISP	O: Hispanic Interaction Of FEMALE	37.43	9.98	0.53	0.00	79.23	34.38	27.93	5.21
RACEBLCK	O: Race/Black Indicator	18.02	0.00	0.00	0.00	27.57	12.09	9.89	11.02
RACEHISP	O: Race/Hispanic Indicator	83.50	77.95	0.00	0.00	81.83	48.59	23.48	0.70
REGNOREA	O: Northeast Region Indicator						89.75		
REGWEST	O: West Region Indicator						87.02		
PAGE18LN	T: Linear: Percent Persons 0-18 Years		0.00				3.42		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years						0.04		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years						23.42		
PAGE18QR	T: Quartic: Percent Persons 0-18 Years						0.25		
IHISP25	Hispanic Interaction Of PAGE18LN						93.64		
IHISP26	Hispanic Interaction Of PAGE18QU						35.64		
IHISP27	Hispanic Interaction Of PAGE18CU						63.08		
IHISP28	Hispanic Interaction Of PAGE18QR						19.87		
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma		0.01				45.61		
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma		0.00				0.00		
IFEM60	Female Interaction Of PSCH12LN						39.05		
IFEM61	Female Interaction Of PSCH12QU						0.00		
IHISP60	Hispanic Interaction Of PSCH12LN						65.20		
IHISP61	Hispanic Interaction Of PSCH12QU						7.25		
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree							0.05	0.00
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree							0.00	83.37
PSCHCOCU	T: Cubic: Bachelors, Graduate, Or Professional Degree								0.00
PSCHCOQR	T: Quartic: Bachelors, Graduate, Or Professional Degree								0.00
PSCHSCLN	T: Linear: Percent Some College And No Degree	6.84				54.91			
PSCHSCQU	T: Quadratic: Percent Some College And No Degree	58.91				8.79			
PSCHSCCU	T: Cubic: Percent Some College And No Degree					17.73			
IHISP80	Hispanic Interaction Of PSCHSCLN	24.70							
IHISP81	Hispanic Interaction Of PSCHSCQU	5.15							
PPOVERLN	T: Linear: Percent Families Below Poverty Level	34.82				38.89			
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level	0.29				13.76			
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					0.13			
PPUBASLN	T: Linear: Percent Households With Public Assist Income		0.00				0.96		
PPUBASQU	T: Quadratic: Percent Households With Public Assist Income		7.25				1.00		
PPUBASCU	T: Cubic: Percent Households With Public Assist Income		0.00						
IHISP90	Hispanic Interaction Of PPUBASLN		0.19				71.66		
IHISP91	Hispanic Interaction Of PPUBASQU						10.49		
PBLACKLN	T: Linear: Percent Black Nonhispanic								19.46
PBLACKQU	T: Quadratic: Percent Black Nonhispanic								0.00
POTHLN	T: Linear: Percent Other Race/Hispanicity						57.98		
POTHQU	T: Quadratic: Percent Other Race/Hispanicity						20.25		
POTHCU	T: Cubic: Percent Other Race/Hispanicity						0.18		
IFEM110	Female Interaction Of POTHLN						16.77		
IFEM111	Female Interaction Of POTHQU						28.30		
IFEM112	Female Interaction Of POTHCU						0.00		
PFNEVLN	T: Linear: Percent Females Never Married						0.00		
PMNEVLN	T: Linear: Percent Males Never Married		0.00		12.69		0.01		0.00
PMNEVQU	T: Quadratic: Percent Males Never Married						0.12		25.00
PMNEVCU	T: Cubic: Percent Males Never Married						1.89		0.00
IBLK140	Black Interaction Of PMNEVLN						94.45		35.73
IBLK141	Black Interaction Of PMNEVQU						67.45		22.42
IBLK142	Black Interaction Of PMNEVCU						6.56		0.00
IHISP140	Hispanic Interaction Of PMNEVLN		0.00		0.00		20.05		0.00
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force				53.29	98.58			0.00
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force				0.00	68.70			3.14
PMLABCU	T: Cubic: Percent Males 16+ Years Old In Labor Force					72.12			26.22
PMLABQR	T: Quartic: Percent Males 16+ Years Old In Labor Force					3.66			
IFEM145	Female Interaction Of PMLABLN				0.00				68.96
IFEM146	Female Interaction Of PMLABQU				0.00				0.00
IFEM147	Female Interaction Of PMLABCU								0.00
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier				0.00	59.32			
POLDHUQU	T: Quadratic: Percent Housing Units Built 1939 Or Earlier					16.03			
POLDHUCU	T: Cubic: Percent Housing Units Built 1939 Or Earlier					34.04			
IFEM155	Female Interaction Of POLDHULN				0.00	40.32			
IFEM156	Female Interaction Of POLDHUQU					15.54			
IFEM157	Female Interaction Of POLDHUCU					4.60			
P40HULN	T: Linear: Percent Housing Units Built 1940-1949								0.00
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol			57.85			2.41		
ADRAT1QU	C: Quadratic: Death Rate With Explicit Mention Of Alcohol			6.40			0.35		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
ADRAT1CU	C: Cubic: Death Rate With Explicit Mention Of Alcohol			0.97			0.00		
ADRAT1QR	C: Quartic: Death Rate With Explicit Mention Of Alcohol			0.00					
V18FLN	C: Linear: Marijuana Posession Arrest Rate						0.08		
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate						0.01		
IBLK185	Black Interaction Of V18FLN						33.22		
IBLK186	Black Interaction Of V18FQU						13.51		
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate						35.71		
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate						68.20		
IFEM195	Female Interaction Of V18ELN						2.75		
IFEM196	Female Interaction Of V18EQU						0.00		
IBLK195	Black Interaction Of V18ELN					64.77			
IBLK198	Black Interaction Of V18EQR					46.33			
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate					71.34			
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate					97.77			
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate					5.59			
IFEM225	Female Interaction Of V18LN					88.51			
IFEM226	Female Interaction Of V18QU					91.88			
IFEM227	Female Interaction Of V18CU					3.69			
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate					98.51			
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate					82.82			
VIOLCU	C: Cubic: Total Violent Offenses Arrest Rate					22.64			
VIOLQR	C: Quartic: Total Violent Offenses Arrest Rate					6.80			
RH43ALN	T: Linear: Recoded Median Rents For Rental Units				0.00				0.07
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units				0.00				0.00
RH61ALN	T: Linear: Recoded Median Value Of Owner Occupied HUs				0.00				
IBLK245	Black Interaction Of RH61ALN				0.00				
RP80ALN	T: Linear: Recoded Median Household Income					18.43	5.66		
RP80AQU	T: Quadratic: Recoded Median Household Income					58.81	0.05		
RP80ACU	T: Cubic: Recoded Median Household Income					2.31			
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992			14.28		3.01		50.20	
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992			0.01				1.89	
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992			0.00				0.60	
IHISP255	Hispanic Interaction Of ARATELN			0.00				7.02	
IHISP256	Hispanic Interaction Of ARATEQU			0.00				68.23	
IHISP257	Hispanic Interaction Of ARATECU			0.00				4.47	
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability						0.01		
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability						8.40		
IBLK260	Black Interaction Of B64DISLN						51.68		
IBLK261	Black Interaction Of B64DISQU						3.84		
BAGE34LN	B: Linear: Percent Persons 25-34 Years				0.00	62.42			33.53
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years				0.00	15.40			0.00
BAGE34CU	B: Cubic: Percent Persons 25-34 Years				0.00				0.00
BAGE34QR	B: Quartic: Percent Persons 25-34 Years								6.03
IFEM265	Female Interaction Of BAGE34LN				0.00				0.00
IFEM266	Female Interaction Of BAGE34QU				0.00				42.83
IFEM267	Female Interaction Of BAGE34CU				0.00				0.00
IFEM268	Female Interaction Of BAGE34QR								0.00
IBLK265	Black Interaction Of BAGE34LN					86.13			
IBLK266	Black Interaction Of BAGE34QU					19.44			
IHISP265	Hispanic Interaction Of BAGE34LN				20.02				93.37
IHISP266	Hispanic Interaction Of BAGE34QU				0.00				0.60
IHISP267	Hispanic Interaction Of BAGE34CU								31.11
BAGE44LN	B: Linear: Percent Persons 35-44 Years						2.06		
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years						69.76		
BAGE44CU	B: Cubic: Percent Persons 35-44 Years						69.73		
BAGE44QR	B: Quartic: Percent Persons 35-44 Years						0.03		
BAGE54LN	B: Linear: Percent Persons 45-54 Years					4.68		6.29	
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years					64.43		0.00	
BAGE54CU	B: Cubic: Percent Persons 45-54 Years					19.15			
BAGE54QR	B: Quartic: Percent Persons 45-54 Years					25.03			
IFEM275	Female Interaction Of BAGE54LN					14.02			
IFEM276	Female Interaction Of BAGE54QU					14.42			
IFEM277	Female Interaction Of BAGE54CU					99.82			
IFEM278	Female Interaction Of BAGE54QR					3.05			
BAGE64LN	B: Linear: Percent Persons 55-64 Years				0.00				0.00
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years								8.39
BAGE64CU	B: Cubic: Percent Persons 55-64 Years								0.00
BAGE64QR	B: Quartic: Percent Persons 55-64 Years								0.00
IBLK280	Black Interaction Of BAGE64LN				0.00				0.00
IBLK282	Black Interaction Of BAGE64CU								22.10
IBLK283	Black Interaction Of BAGE64QR								0.00
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander	24.24				91.31			0.22
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander	29.06				44.69			0.00
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander	62.49				15.03			
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander	23.64				23.27			
IFEM285	Female Interaction Of BASIANLN	11.70				53.48			
IFEM286	Female Interaction Of BASIANQU	32.74				79.62			
IFEM287	Female Interaction Of BASIANCU	97.25				87.99			
IFEM288	Female Interaction Of BASIANQR	5.71				4.69			
IHISP285	Hispanic Interaction Of BASIANLN								11.49

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
IHISP286	Hispanic Interaction Of BASIANQU								0.17
BCUBANLN	B: Linear: Percent Hispanics: Cuban				0.00				8.21
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban								0.00
IHISP290	Hispanic Interaction Of BCUBANLN				0.00				16.37
IHISP291	Hispanic Interaction Of BCUBANQU								1.55
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed				0.00	78.51			0.20
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed					66.11			76.24
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed					0.25			0.00
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed								0.00
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut			40.17	98.08			3.98	0.00
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			19.35	0.00			0.56	0.00
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut				0.00				0.00
BINDIAQR	B: Quartic: Percent Pop: American Indian, Eskimo, Aleut				0.00				
IHISP300	Hispanic Interaction Of BINDIALN			38.45				96.34	
IHISP301	Hispanic Interaction Of BINDIAQU			0.00				2.61	
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed	69.97				86.27			
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed	18.54				31.79			
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed					18.19			
BMNOTQR	B: Quartic: Percent Males Separated, Divorced or Widowed					24.53			
IFEM305	Female Interaction Of BMNOTLN	23.89				70.83			
IFEM306	Female Interaction Of BMNOTQU	3.24				31.76			
IFEM307	Female Interaction Of BMNOTCU					97.78			
IFEM308	Female Interaction Of BMNOTQR					2.40			
BPOVERLN	B: Linear: Percent Families Below Poverty Level	32.43						46.73	
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level	60.68						60.11	
BPOVERCU	B: Cubic: Percent Families Below Poverty Level	84.12						0.00	
BPOVERQR	B: Quartic: Percent Families Below Poverty Level	5.34							
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican							17.52	7.43
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican							25.91	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican							4.87	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican							4.47	
IHISP315	Hispanic Interaction Of BPRICALN							94.62	0.73
IHISP316	Hispanic Interaction Of BPRICAQU							12.72	
IHISP317	Hispanic Interaction Of BPRICACU							30.37	
IHISP318	Hispanic Interaction Of BPRICAQR							14.08	
BSCHASLN	B: Linear: Percent Associates Degree				0.00				14.87
BSCHASQU	B: Quadratic: Percent Associates Degree				0.00				0.00
BSCHASCU	B: Cubic: Percent Associates Degree				0.00				0.00
BSCHASQR	B: Quartic: Percent Associates Degree				0.00				0.00
IBLK320	Black Interaction Of BSCHASLN				0.00				21.52
IBLK321	Black Interaction Of BSCHASQU								0.00
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander				0.00			0.03	0.00
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander							0.00	0.00
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander								3.07
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander								0.00
IFEM325	Female Interaction Of PASIANLN				0.00			0.37	0.00
IFEM326	Female Interaction Of PASIANQU							0.00	
IHISP325	Hispanic Interaction Of PASIANLN							5.15	
PCUBANLN	T: Linear: Percent Hispanics: Cuban						69.17	0.00	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban						1.46	0.56	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban						42.87	50.48	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban						18.70	0.00	
IFEM330	Female Interaction Of PCUBANLN						64.02		
IFEM331	Female Interaction Of PCUBANQU						18.59		
IFEM332	Female Interaction Of PCUBANCU						59.22		
IFEM333	Female Interaction Of PCUBANQR						1.97		
IBLK330	Black Interaction Of PCUBANLN							28.02	
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut			0.00				0.03	
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			0.04					
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut			0.00					
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut			0.00					

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F7. Significance Probabilities for Fixed Effect Coefficients for Past Year Illicit Drug Use Treatment by Age

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	0.14	4.25	0.00	68.27	3.58	0.79	0.01	0.00
FEMBLCK	O: Black Interaction Of FEMALE	63.75	76.89	64.53	0.00	44.15	46.19	42.27	14.83
FEMHISP	O: Hispanic Interaction Of FEMALE	1.08	0.19	0.01	0.12	54.30	70.89	46.50	25.73
FEMOTHR	O: Other Interaction Of FEMALE							12.59	
RACEBLCK	O: Race/Black Indicator	0.17	0.84	0.16	0.00	37.33	9.46	25.89	0.00
RACEHISP	O: Race/Hispanic Indicator	80.64	88.44	1.06	0.00	85.61	5.98	23.33	16.68
RACEOTHR	O: Race/Other Indicator		0.02			21.58		60.38	0.00
REGNOREA	O: Northeast Region Indicator					3.15			
REGSOUTH	O: South Region Indicator					51.98			
IHISP7	Hispanic Interaction Of REGNOREA			0.07					
IHISP8	Hispanic Interaction Of REGSOUTH			1.33					
IHISP9	Hispanic Interaction Of REGWEST			4.74					
PHH1PLN	T: Linear: Percent One Person Households							0.01	0.00
PHH1PQU	T: Quadratic: Percent One Person Households							0.01	11.18
PHH1PCU	T: Cubic: Percent One Person Households							1.33	0.01
PHH1PQR	T: Quartic: Percent One Person Households							0.00	0.00
POPRMLN	T: Linear: Average Persons Per Room							0.00	
PAGE34LN	T: Linear: Percent Persons 25-34 Years					61.80			
IFEM35	Female Interaction Of PAGE34LN					0.61			
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma				0.00				0.00
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma								0.00
PSCHSCLN	T: Linear: Percent Some College And No Degree		0.00				2.09		0.00
PSCHSCQU	T: Quadratic: Percent Some College And No Degree		0.78				0.19		0.00
IFEM80	Female Interaction Of PSCHSCLN		0.02						0.00
IFEM81	Female Interaction Of PSCHSCQU		0.00						0.00
IHISP80	Hispanic Interaction Of PSCHSCLN						38.73		
IHISP81	Hispanic Interaction Of PSCHSCQU						8.65		
PPOVERLN	T: Linear: Percent Families Below Poverty Level				0.00				0.00
PBLACKLN	T: Linear: Percent Black Nonhispanic							0.64	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic							0.06	
PHISPLN	T: Linear: Percent Hispanic			93.38				64.08	
PHISPQU	T: Quadratic: Percent Hispanic			0.00				0.03	
PHISPCU	T: Cubic: Percent Hispanic			0.00				0.00	
PHISPQR	T: Quartic: Percent Hispanic			0.00				0.16	
POTHLN	T: Linear: Percent Other Race/Hispanicity						34.18		
POTHQU	T: Quadratic: Percent Other Race/Hispanicity						42.63		
POTHCU	T: Cubic: Percent Other Race/Hispanicity						0.17		
IHISP110	Hispanic Interaction Of POTHLN						77.49		
IHISP111	Hispanic Interaction Of POTHQU						15.30		
IHISP112	Hispanic Interaction Of POTHCU						24.84		
PHHF18LN	T: Linear: % Female-Headed HH W/No Spouse & Chld Under 18					42.41			
PHHF18QU	T: Quadratic: % Female-Headed HH W/No Spouse & Chld Under 18					78.22			
PHHF18CU	T: Cubic: % Female-Headed HH W/No Spouse & Chld Under 18					16.55			
PHHF18QR	T: Quartic: % Female-Headed HH W/No Spouse & Chld Under 18					0.82			
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force	76.03				16.18			
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force	0.29				56.85			
PFLABCU	T: Cubic: Percent Females 16+ Years Old In Labor Force					83.02			
PFLABQR	T: Quartic: Percent Females 16+ Years Old In Labor Force					0.05			
PFNEVLN	T: Linear: Percent Females Never Married			0.00				66.06	
PFNEVQU	T: Quadratic: Percent Females Never Married							11.79	
PFNEVCU	T: Cubic: Percent Females Never Married							24.54	
PFNEVQR	T: Quartic: Percent Females Never Married							0.00	
PMNEVLN	T: Linear: Percent Males Never Married						0.00		0.15
PMNEVQU	T: Quadratic: Percent Males Never Married								0.00
IBLK140	Black Interaction Of PMNEVLN						2.39		
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force								2.51
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force								9.23
PMLABCU	T: Cubic: Percent Males 16+ Years Old In Labor Force								0.00
PMLABQR	T: Quartic: Percent Males 16+ Years Old In Labor Force								0.00
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases						59.54		
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases						0.00		
V18FLN	C: Linear: Marijuana Posession Arrest Rate				0.00		42.28		0.00
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate						1.06		0.00
V18FCU	C: Cubic: Marijuana Posession Arrest Rate						9.02		0.00
V18FQR	C: Quartic: Marijuana Posession Arrest Rate						0.11		
IHISP185	Hispanic Interaction Of V18FLN								9.25
IHISP186	Hispanic Interaction Of V18FQU								74.08
IHISP187	Hispanic Interaction Of V18FCU								0.06
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate		0.12				0.00		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

		BIG CITY				REMAINDER			
Variable	Label	12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate						0.00		
IBLK190	Black Interaction Of V18BLN						99.33		
IBLK191	Black Interaction Of V18BQU						3.64		
IHISP190	Hispanic Interaction Of V18BLN		0.00				21.06		
IHISP191	Hispanic Interaction Of V18BQU						2.42		
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate						0.00		
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate				0.00				0.00
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992	8.52							
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992	0.43							
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992	0.02							
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability			0.00	13.00			0.00	
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability			1.38	0.00			0.53	
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability			56.16				84.63	
B64DISQR	B: Quartic: Percent Persons 16-64 With A Work Disability			0.00				0.06	
IHISP260	Hispanic Interaction Of B64DISLN							81.65	
IHISP261	Hispanic Interaction Of B64DISQU							82.66	
IHISP262	Hispanic Interaction Of B64DISCU							21.84	
IHISP263	Hispanic Interaction Of B64DISQR							24.37	
BAGE34LN	B: Linear: Percent Persons 25-34 Years					61.08			
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years					82.90			
BAGE34CU	B: Cubic: Percent Persons 25-34 Years					94.80			
BAGE44LN	B: Linear: Percent Persons 35-44 Years			0.00		55.25	34.50	78.58	0.01
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years			46.69		68.67	59.71	3.34	0.00
BAGE44CU	B: Cubic: Percent Persons 35-44 Years			0.00		6.31	19.32	1.85	0.00
BAGE44QR	B: Quartic: Percent Persons 35-44 Years						8.37	0.00	
IFEM270	Female Interaction Of BAGE44LN						94.21		
IFEM271	Female Interaction Of BAGE44QU						18.10		
IFEM272	Female Interaction Of BAGE44CU						0.05		
IFEM273	Female Interaction Of BAGE44QR						0.23		
BAGE54LN	B: Linear: Percent Persons 45-54 Years	0.36				10.96	0.23		
IFEM275	Female Interaction Of BAGE54LN						0.00		
IBLK275	Black Interaction Of BAGE54LN	0.22							
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander		0.00	0.00			2.35		
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander			0.00			63.36		
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander						27.65		
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander						19.50		
IFEM285	Female Interaction Of BASIANLN						13.19		
IFEM286	Female Interaction Of BASIANQU						1.90		
IFEM287	Female Interaction Of BASIANCU						58.98		
IFEM288	Female Interaction Of BASIANQR						0.00		
IBLK285	Black Interaction Of BASIANLN		0.03				1.72		
BCUBANLN	B: Linear: Percent Hispanics: Cuban	2.09					11.54		
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban	57.86					34.75		
BCUBANCU	B: Cubic: Percent Hispanics: Cuban	0.03							
IFEM290	Female Interaction Of BCUBANLN						59.81		
IFEM291	Female Interaction Of BCUBANQU						1.40		
IBLK290	Black Interaction Of BCUBANLN	1.33							
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed					7.18	11.48	0.10	0.00
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed						83.72	61.42	0.00
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed						0.03	21.01	
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed						0.02	0.01	
IFEM295	Female Interaction Of BFNOTLN					5.88			0.77
IFEM296	Female Interaction Of BFNOTQU								0.00
IBLK295	Black Interaction Of BFNOTLN						72.52		
IBLK297	Black Interaction Of BFNOTCU						36.26		
IBLK298	Black Interaction Of BFNOTQR						10.42		
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut		89.05		0.00		0.03		0.00
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut		14.74		0.00		43.39		0.00
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut		0.00		0.00		0.01		0.00
BINDIAQR	B: Quartic: Percent Pop: American Indian, Eskimo, Aleut		0.39				0.06		0.00
IBLK300	Black Interaction Of BINDIALN		0.68		0.00		99.37		6.75
IBLK301	Black Interaction Of BINDIAQU		40.46		0.00		35.31		0.73
IBLK302	Black Interaction Of BINDIACU		16.63		0.00		89.33		0.00
IBLK303	Black Interaction Of BINDIAQR		3.12				19.93		0.00
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed								0.00
BPOVERLN	B: Linear: Percent Families Below Poverty Level		29.37						0.00
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.02						0.00
IOTH311	Other Interaction Of BPOVERQU		9.68						
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican		0.00				2.12		0.00
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican						22.65		
IHISP315	Hispanic Interaction Of BPRICALN						51.89		
IHISP316	Hispanic Interaction Of BPRICAQU						1.94		
BSCHASLN	B: Linear: Percent Associates Degree							0.63	
BSCHASQU	B: Quadratic: Percent Associates Degree							13.65	
BSCHASCU	B: Cubic: Percent Associates Degree							0.09	
IBLK320	Black Interaction Of BSCHASLN							7.22	
IBLK321	Black Interaction Of BSCHASQU							29.60	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
IBLK322	Black Interaction Of BSCHASCU							0.42	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander					34.34		76.52	
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander							0.09	
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander							48.88	
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander							0.00	
PCUBANLN	T: Linear: Percent Hispanics: Cuban						94.41		
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban						14.84		
PCUBANCU	T: Cubic: Percent Hispanics: Cuban						81.24		
PCUBANQR	T: Quartic: Percent Hispanics: Cuban						4.25		
IFEM330	Female Interaction Of PCUBANLN						0.58		
IFEM331	Female Interaction Of PCUBANQU						17.81		
IFEM332	Female Interaction Of PCUBANCU						62.35		
IFEM333	Female Interaction Of PCUBANQR						0.01		
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut							0.01	
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut							2.07	
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut							0.00	
IBLK335	Black Interaction Of PINDIALN							69.10	
IBLK336	Black Interaction Of PINDIAQU							63.35	
IBLK337	Black Interaction Of PINDIACU							1.95	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F8. Significance Probabilities for Fixed Effect Coefficients for Past Year Dependency on Alcohol Only by Age

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	4.96	0.00	0.00	0.00	65.85	0.00	0.00	0.00
FEMBLCK	O: Black Interaction Of FEMALE	0.74	89.92	50.43	0.00	25.55	84.93	94.45	20.97
FEMHISP	O: Hispanic Interaction Of FEMALE	95.05	0.00	0.00	0.00	86.83	83.81	26.07	0.52
FEMOTHR	O: Other Interaction Of FEMALE	64.23	19.71	32.50	0.00		0.00	50.36	
RACEBLCK	O: Race/Black Indicator	54.46	0.00	0.00	0.00	23.20	0.41	87.95	0.00
RACEHISP	O: Race/Hispanic Indicator	1.90	35.01	0.06	52.63	91.23	25.78	0.75	4.51
RACEOTHR	O: Race/Other Indicator	0.73	0.00	0.00	0.00		0.00	0.99	15.38
REGNOREA	O: Northeast Region Indicator					70.41			
REGSOUTH	O: South Region Indicator					45.32			
REGWEST	O: West Region Indicator					51.64			
IBLK7	Black Interaction Of REGNOREA					28.15			
IBLK8	Black Interaction Of REGSOUTH					30.03			
IBLK9	Black Interaction Of REGWEST					18.21			
PDENLEV1	O: Large MSA					9.37			2.73
PDENLEV2	O: Medium MSA					16.12			70.98
PDENLEV3	O: Small MSA					81.48			1.63
PDENLEV4	O: NonMSA, Urban					53.42			0.00
IBLK10	Black Interaction Of PDENLEV1					66.79			0.00
IBLK11	Black Interaction Of PDENLEV2					92.18			0.00
IBLK12	Black Interaction Of PDENLEV3					23.52			0.54
IBLK13	Black Interaction Of PDENLEV4					46.49			6.10
PHH1PLN	T: Linear: Percent One Person Households								0.00
PHH1PQU	T: Quadratic: Percent One Person Households								0.00
IFEM15	Female Interaction Of PHH1PLN								22.53
IFEM16	Female Interaction Of PHH1PQU								0.00
POPRMLN	T: Linear: Average Persons Per Room		0.00				0.01		
POPRMQU	T: Quadratic: Average Persons Per Room						0.01		
POPRMCU	T: Cubic: Average Persons Per Room						0.00		
IBLK20	Black Interaction Of POPRMLN						78.61		
IBLK21	Black Interaction Of POPRMQU						4.50		
PAGE44LN	T: Linear: Percent Persons 35-44 Years					95.95			
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years					3.94			
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma				68.91				0.00
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma				0.00				0.00
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma								0.00
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree								0.00
PSCHSCLN	T: Linear: Percent Some College And No Degree			0.32				11.60	
PSCHSCQU	T: Quadratic: Percent Some College And No Degree			63.96				0.01	
PSCHSCCU	T: Cubic: Percent Some College And No Degree			0.00				73.54	
IFEM80	Female Interaction Of PSCHSCLN							29.01	
IFEM81	Female Interaction Of PSCHSCQU							63.88	
IFEM82	Female Interaction Of PSCHSCCU							0.00	
PPOVERLN	T: Linear: Percent Families Below Poverty Level					92.01		0.00	
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level					49.93		0.01	
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					34.87		23.15	
PPOVERQR	T: Quartic: Percent Families Below Poverty Level					98.40		0.00	
IFEM85	Female Interaction Of PPOVERLN					78.71			
IFEM86	Female Interaction Of PPOVERQU					9.66			
IFEM87	Female Interaction Of PPOVERCU					51.85			
IFEM88	Female Interaction Of PPOVERQR					1.34			
IHISP85	Hispanic Interaction Of PPOVERLN					24.62		68.85	
IHISP86	Hispanic Interaction Of PPOVERQU					18.44		0.52	
P64DISLN	T: Linear: Percent 16-64 With A Work Disability		0.35				0.01		
P64DISQU	T: Quadratic: Percent 16-64 With A Work Disability		9.42				2.99		
P64DISCU	T: Cubic: Percent 16-64 With A Work Disability		11.69				15.47		
P64DISQR	T: Quartic: Percent 16-64 With A Work Disability		15.40				0.00		
IBLK95	Black Interaction Of P64DISLN		3.64				12.94		
IBLK96	Black Interaction Of P64DISQU		0.00				43.73		
IBLK97	Black Interaction Of P64DISCU		88.81				16.76		
IBLK98	Black Interaction Of P64DISQR		0.00				2.58		
PHISPLN	T: Linear: Percent Hispanic								0.00
PHISPQU	T: Quadratic: Percent Hispanic								0.00
IBLK105	Black Interaction Of PHISPLN								33.38
IBLK106	Black Interaction Of PHISPQU								0.00
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed						36.75		
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed						40.18		
IOTH135	Other Interaction Of PFNOTLN						41.02		
IOTH136	Other Interaction Of PFNOTQU						0.28		
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed								0.07
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed								0.00
PMNOTCU	T: Cubic: Percent Males Separated, Divorced Or Widowed								0.00

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier								0.55
POLDHUQU	T: Quadratic: Percent Housing Units Built 1939 Or Earlier								16.50
POLDHUCU	T: Cubic: Percent Housing Units Built 1939 Or Earlier								0.00
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol	30.62				34.36			
ADRAT1QU	C: Quadratic: Death Rate With Explicit Mention Of Alcohol	77.02				93.74			
ADRAT1CU	C: Cubic: Death Rate With Explicit Mention Of Alcohol	59.08				90.66			
ADRAT1QR	C: Quartic: Death Rate With Explicit Mention Of Alcohol	0.83				1.38			
V18FLN	C: Linear: Marijuana Posession Arrest Rate					85.01			
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate					25.85			
IFEM185	Female Interaction Of V18FLN					93.74			
IFEM186	Female Interaction Of V18FQU					0.85			
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate			0.00	72.83		0.84	6.69	67.52
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate			0.00	0.00			0.05	0.64
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate				0.00				0.00
IHISP195	Hispanic Interaction Of V18ELN								13.91
IHISP196	Hispanic Interaction Of V18EQU								0.05
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate			0.64				70.37	36.12
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate			0.42				0.02	0.00
IHISP200	Hispanic Interaction Of V18ALN								1.59
IHISP201	Hispanic Interaction Of V18AQU								0.00
IOTH200	Other Interaction Of V18ALN			0.12					
IOTH201	Other Interaction Of V18AQU			0.00					
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate	11.07		0.00	43.44	11.66		0.01	42.18
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate			0.01	74.10	19.02		0.01	0.05
V18CU	C: Cubic: Total Drug Abuse Violations Arrest Rate				0.00	23.24			0.00
V18QR	C: Quartic: Total Drug Abuse Violations Arrest Rate					0.48			
IBLK225	Black Interaction Of V18LN								0.62
IBLK226	Black Interaction Of V18QU								0.00
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992	7.10				2.29			
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992	6.13				2.25			
DRATECU	C: Cubic: Mean Drug Client Treatment Rate 1991 & 1992	85.42							
DRATEQR	C: Quartic: Mean Drug Client Treatment Rate 1991 & 1992	3.80							
IBLK235	Black Interaction Of DRATELN	61.09							
IBLK236	Black Interaction Of DRATEQU	7.02							
IBLK237	Black Interaction Of DRATECU	0.76							
IBLK238	Black Interaction Of DRATEQR	0.30							
RP80ALN	T: Linear: Recoded Median Household Income		3.73				0.01		
RP80AQU	T: Quadratic: Recoded Median Household Income		0.48				57.07		
RP80ACU	T: Cubic: Recoded Median Household Income		11.70				66.61		
RP80AQR	T: Quartic: Recoded Median Household Income		32.55				0.00		
IBLK250	Black Interaction Of RP80ALN		36.53				77.22		
IBLK251	Black Interaction Of RP80AQU		0.07				1.83		
IBLK252	Black Interaction Of RP80ACU		1.45				4.15		
IBLK253	Black Interaction Of RP80AQR		0.00				18.41		
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992					0.20			
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability		0.12				82.29		
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability		64.85				44.26		
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability		33.49				3.11		
B64DISQR	B: Quartic: Percent Persons 16-64 With A Work Disability		45.70						
IHISP260	Hispanic Interaction Of B64DISLN		1.75						
IHISP261	Hispanic Interaction Of B64DISQU		64.92						
IHISP262	Hispanic Interaction Of B64DISCU		4.41						
IHISP263	Hispanic Interaction Of B64DISQR		0.00						
BAGE34LN	B: Linear: Percent Persons 25-34 Years		5.81				50.12		0.00
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years		32.45				0.16		0.00
BAGE34CU	B: Cubic: Percent Persons 25-34 Years						0.00		0.00
BAGE34QR	B: Quartic: Percent Persons 25-34 Years								0.00
IFEM265	Female Interaction Of BAGE34LN		0.00						
IFEM266	Female Interaction Of BAGE34QU		0.00						
BAGE44LN	B: Linear: Percent Persons 35-44 Years			0.00		68.94		7.09	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years			3.56		20.51		0.63	
BAGE44CU	B: Cubic: Percent Persons 35-44 Years			0.00		1.93		8.29	
BAGE44QR	B: Quartic: Percent Persons 35-44 Years							0.00	
IOTH270	Other Interaction Of BAGE44LN			5.75				0.75	
IOTH271	Other Interaction Of BAGE44QU			0.00					
BAGE54LN	B: Linear: Percent Persons 45-54 Years						39.63		
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years						0.81		
IOTH275	Other Interaction Of BAGE54LN						47.24		
IOTH276	Other Interaction Of BAGE54QU						0.58		
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander			66.77	0.00			35.64	
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander				0.00			48.54	
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander				89.60			74.27	
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander				0.00			0.67	
IHISP285	Hispanic Interaction Of BASIANLN			0.00				49.43	
IHISP286	Hispanic Interaction Of BASIANQU							17.81	
IHISP287	Hispanic Interaction Of BASIANCU							35.93	
IHISP288	Hispanic Interaction Of BASIANQR							0.94	
IOTH285	Other Interaction Of BASIANLN				37.09				
IOTH286	Other Interaction Of BASIANQU				0.00				
IOTH287	Other Interaction Of BASIANCU				0.00				
IOTH288	Other Interaction Of BASIANQR				0.00				

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
BCUBANLN	B: Linear: Percent Hispanics: Cuban		0.00	0.00			32.91		
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban		0.00	3.38			20.78		
BCUBANCU	B: Cubic: Percent Hispanics: Cuban						29.06		
BCUBANQR	B: Quartic: Percent Hispanics: Cuban						0.00		
IFEM290	Female Interaction Of BCUBANLN		0.00				63.31		
IFEM291	Female Interaction Of BCUBANQU		0.00				49.55		
IFEM292	Female Interaction Of BCUBANCU						0.04		
IFEM293	Female Interaction Of BCUBANQR						0.00		
IBLK290	Black Interaction Of BCUBANLN			10.19					
IBLK291	Black Interaction Of BCUBANQU			0.00					
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed							1.10	0.00
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed							92.94	24.93
IFEM295	Female Interaction Of BFNOTLN								0.06
IFEM296	Female Interaction Of BFNOTQU								0.00
IOTH295	Other Interaction Of BFNOTLN							0.16	
IOTH296	Other Interaction Of BFNOTQU							0.00	
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut			0.00		60.23		0.92	0.01
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut							10.17	0.00
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut							1.09	
IBLK300	Black Interaction Of BINDIALN			0.00		2.59		50.83	
IBLK301	Black Interaction Of BINDIAQU							63.91	
IBLK302	Black Interaction Of BINDIACU							0.47	
IHISP300	Hispanic Interaction Of BINDIALN			0.00				35.26	
IHISP301	Hispanic Interaction Of BINDIAQU							3.21	
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed			0.00	0.00	55.18	9.17	8.02	0.14
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed				0.00	11.37	31.83	0.00	0.00
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed				46.81	88.51			0.00
BMNOTQR	B: Quartic: Percent Males Separated, Divorced or Widowed				0.00	4.51			17.98
IBLK305	Black Interaction Of BMNOTLN				0.00				0.00
IBLK306	Black Interaction Of BMNOTQU				0.00				53.28
IBLK307	Black Interaction Of BMNOTCU				0.00				0.00
IBLK308	Black Interaction Of BMNOTQR								0.00
IHISP305	Hispanic Interaction Of BMNOTLN					31.66			
IHISP306	Hispanic Interaction Of BMNOTQU					57.74			
IHISP307	Hispanic Interaction Of BMNOTCU					87.35			
IHISP308	Hispanic Interaction Of BMNOTQR					7.17			
BPOVERLN	B: Linear: Percent Families Below Poverty Level				0.00				0.00
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level				0.00				
BPOVERCU	B: Cubic: Percent Families Below Poverty Level				0.00				
IHISP310	Hispanic Interaction Of BPOVERLN								0.00
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican		0.00	0.02			0.13	3.05	
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican			2.63				16.50	
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican			0.00				12.62	
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican							0.00	
IBLK315	Black Interaction Of BPRICALN							14.04	
IBLK316	Black Interaction Of BPRICAQU							57.20	
IBLK317	Black Interaction Of BPRICACU							70.65	
IBLK318	Black Interaction Of BPRICAQR							0.87	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander	78.78				47.49	26.45		0.00
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander	0.77				11.47	16.96		0.00
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander	2.40				43.05	0.27		0.00
PASIANQR	T: Quartic: Percent Population: Asian, Pacific Islander					33.47			
IFEM325	Female Interaction Of PASIANLN	93.85				6.84			
IFEM326	Female Interaction Of PASIANQU	1.06				37.34			
IFEM327	Female Interaction Of PASIANCU	1.57				83.07			
IFEM328	Female Interaction Of PASIANQR					1.61			
IBLK325	Black Interaction Of PASIANLN						31.62		
IBLK326	Black Interaction Of PASIANQU						55.41		
IBLK327	Black Interaction Of PASIANCU						0.22		
PCUBANLN	T: Linear: Percent Hispanics: Cuban	56.65	0.00		0.03	82.30	26.17		0.00
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban				0.00		89.19		6.12
PCUBANCU	T: Cubic: Percent Hispanics: Cuban								0.00
IBLK330	Black Interaction Of PCUBANLN	1.54			0.00	14.80			7.06
IBLK331	Black Interaction Of PCUBANQU				0.00				0.69
IBLK332	Black Interaction Of PCUBANCU								0.00
IOTH330	Other Interaction Of PCUBANLN						8.87		
IOTH331	Other Interaction Of PCUBANQU						3.06		
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut	50.38			0.00	50.95			
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut	32.08			0.00	5.07			
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut	82.38			0.00	14.36			
PINDIAQR	T: Quartic: Percent Pop: American Indian, Eskimo, Aleut				0.00				
IFEM335	Female Interaction Of PINDIALN	26.50				95.94			
IFEM336	Female Interaction Of PINDIAQU	10.17				1.57			
IFEM337	Female Interaction Of PINDIACU	1.88				2.55			
IHISP335	Hispanic Interaction Of PINDIALN				0.00				
IHISP336	Hispanic Interaction Of PINDIAQU				0.00				
IHISP337	Hispanic Interaction Of PINDIACU				0.00				
IHISP338	Hispanic Interaction Of PINDIAQR				0.00				
IOTH335	Other Interaction Of PINDIALN				0.00				
IOTH336	Other Interaction Of PINDIAQU				0.00				

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F9. Significance Probabilities for Fixed Effect Coefficients for Past Year Illicit Dependency by Age

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	2.68	0.00	0.00	0.00	9.87	0.00	0.00	23.75
FEMBLCK	O: Black Interaction Of FEMALE	73.18	66.35	7.75	3.29	40.18	23.88	68.50	1.74
FEMHISP	O: Hispanic Interaction Of FEMALE	24.30	9.14	0.00	1.36	39.07	63.64	89.55	53.40
RACEBLCK	O: Race/Black Indicator	0.08	0.00	3.65	0.00	2.27	0.24	43.36	0.00
RACEHISP	O: Race/Hispanic Indicator	55.36	0.00	0.00	0.00	56.54	4.08	7.68	47.27
RACEOTHR	O: Race/Other Indicator		0.00				0.00		
REGNOREA	O: Northeast Region Indicator							5.24	
REGSOUTH	O: South Region Indicator							8.23	
REGWEST	O: West Region Indicator							0.00	
PHH1PLN	T: Linear: Percent One Person Households						0.00		
POPRMLN	T: Linear: Average Persons Per Room			0.00				0.05	
POPRMQU	T: Quadratic: Average Persons Per Room			0.00				37.41	
POPRMCU	T: Cubic: Average Persons Per Room							9.63	
POPRMQR	T: Quartic: Average Persons Per Room							0.00	
PAGE18LN	T: Linear: Percent Persons 0-18 Years		9.79				0.00		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years		0.33				0.13		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years		2.47				7.37		
PAGE18QR	T: Quartic: Percent Persons 0-18 Years		0.01				0.00		
IFEM25	Female Interaction Of PAGE18LN		0.00				11.06		
IFEM26	Female Interaction Of PAGE18QU		4.81				58.38		
IFEM27	Female Interaction Of PAGE18CU		0.13				25.26		
IFEM28	Female Interaction Of PAGE18QR		0.00				0.07		
PAGE34LN	T: Linear: Percent Persons 25-34 Years		1.33			3.30		3.46	
PAGE34QU	T: Quadratic: Percent Persons 25-34 Years					0.06		0.00	
IBLK35	Black Interaction Of PAGE34LN							0.00	
IHISP35	Hispanic Interaction Of PAGE34LN		0.00						
PAGE44LN	T: Linear: Percent Persons 35-44 Years								0.00
PAGE44QU	T: Quadratic: Percent Persons 35-44 Years								0.00
PAGE54LN	T: Linear: Percent Persons 45-54 Years				0.00	51.45			
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years				0.00	48.28			
PAGE54CU	T: Cubic: Percent Persons 45-54 Years				0.00	0.85			
IHISP45	Hispanic Interaction Of PAGE54LN				0.00				
IHISP46	Hispanic Interaction Of PAGE54QU				0.00				
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma	18.10							
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma	5.27							
IFEM60	Female Interaction Of PSCH12LN	91.76							
IFEM61	Female Interaction Of PSCH12QU	0.13							
PPOVERLN	T: Linear: Percent Families Below Poverty Level				0.00				
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level				35.49				
PPOVERCU	T: Cubic: Percent Families Below Poverty Level				0.00				
PMNEVLN	T: Linear: Percent Males Never Married			0.00					
PMNEVQU	T: Quadratic: Percent Males Never Married			2.37					
PMNEVCU	T: Cubic: Percent Males Never Married			0.00					
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force		0.01						
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force		0.15						
PMLABCU	T: Cubic: Percent Males 16+ Years Old In Labor Force		0.00						
IHISP145	Hispanic Interaction Of PMLABLN		0.10						
IHISP146	Hispanic Interaction Of PMLABQU		92.82						
IHISP147	Hispanic Interaction Of PMLABCU		0.00						
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed								11.82
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed								0.00
PRENTLN	T: Linear: Percent Housing Rented							0.08	0.00
PRENTQU	T: Quadratic: Percent Housing Rented							0.00	
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases							15.97	
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases							0.00	
V18FLN	C: Linear: Marijuana Posession Arrest Rate			0.00					
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate			0.00				29.72	
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate			0.15				3.28	
V18ECU	C: Cubic: Opium/Cocaine & Deriv Posession Arrest Rate			1.23				6.34	
V18EQR	C: Quartic: Opium/Cocaine & Deriv Posession Arrest Rate			0.01				0.00	
IHISP195	Hispanic Interaction Of V18ELN			21.79					
IHISP196	Hispanic Interaction Of V18EQU			43.45					
IHISP197	Hispanic Interaction Of V18ECU			0.71					
IHISP198	Hispanic Interaction Of V18EQR			0.00					
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate						76.05		
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate						0.01		
VIOLLN	C: Linear: Total Violent Offenses Arrest Rate								0.00
VIOLQU	C: Quadratic: Total Violent Offenses Arrest Rate								0.00
VIOLCU	C: Cubic: Total Violent Offenses Arrest Rate								0.00

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
RP80ALN	T: Linear: Recoded Median Household Income					21.98			0.00
RP80AQU	T: Quadratic: Recoded Median Household Income					0.01			76.48
RP80ACU	T: Cubic: Recoded Median Household Income								0.00
IFEM250	Female Interaction Of RP80ALN					59.34			
IFEM251	Female Interaction Of RP80AQU					2.43			
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992						0.02		
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992						0.00		
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability			0.00					
BAGE34LN	B: Linear: Percent Persons 25-34 Years							15.10	
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years							10.47	
BAGE34CU	B: Cubic: Percent Persons 25-34 Years							3.28	
IFEM265	Female Interaction Of BAGE34LN							46.07	
IFEM266	Female Interaction Of BAGE34QU							1.93	
IFEM267	Female Interaction Of BAGE34CU							0.00	
BAGE44LN	B: Linear: Percent Persons 35-44 Years					0.01		58.46	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years							0.06	
IHISP270	Hispanic Interaction Of BAGE44LN							91.16	
IHISP271	Hispanic Interaction Of BAGE44QU							12.09	
BAGE54LN	B: Linear: Percent Persons 45-54 Years			0.00		43.40	0.05	0.07	56.68
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years					4.50	0.25	0.00	0.21
BAGE54CU	B: Cubic: Percent Persons 45-54 Years					7.08	0.00	0.15	0.00
BAGE54QR	B: Quartic: Percent Persons 45-54 Years								0.00
IFEM275	Female Interaction Of BAGE54LN					84.69		0.00	0.00
IFEM276	Female Interaction Of BAGE54QU					0.32			4.12
IFEM277	Female Interaction Of BAGE54CU								0.00
IFEM278	Female Interaction Of BAGE54QR								0.00
IBLK275	Black Interaction Of BAGE54LN							6.13	
IBLK276	Black Interaction Of BAGE54QU							90.00	
IBLK277	Black Interaction Of BAGE54CU							2.12	
BAGE64LN	B: Linear: Percent Persons 55-64 Years			7.74		0.71	82.16	30.09	
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years			0.00			0.00	0.01	
BAGE64CU	B: Cubic: Percent Persons 55-64 Years			0.00				16.19	
IFEM280	Female Interaction Of BAGE64LN					0.21			
IBLK280	Black Interaction Of BAGE64LN			83.00				18.88	
IBLK281	Black Interaction Of BAGE64QU			0.00				50.12	
IBLK282	Black Interaction Of BAGE64CU			0.00				4.33	
BASIANLN	B: Linear: Percent Population: Asian, Pacific Islander							0.00	0.00
BASIANQU	B: Quadratic: Percent Population: Asian, Pacific Islander								1.30
BASIANCU	B: Cubic: Percent Population: Asian, Pacific Islander								0.00
BASIANQR	B: Quartic: Percent Population: Asian, Pacific Islander								26.99
IBLK285	Black Interaction Of BASIANLN								70.91
IBLK286	Black Interaction Of BASIANQU								6.88
IBLK287	Black Interaction Of BASIANCU								0.00
IBLK288	Black Interaction Of BASIANQR								0.00
BCUBANLN	B: Linear: Percent Hispanics: Cuban				0.00				
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban				0.00				
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed	0.08				12.10	0.03		
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed	58.29				5.56			
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed	86.25				72.29			
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed	0.26				0.12			
IFEM295	Female Interaction Of BFNOTLN						0.00		
IBLK295	Black Interaction Of BFNOTLN					74.70			
IBLK296	Black Interaction Of BFNOTQU					82.01			
IBLK297	Black Interaction Of BFNOTCU					24.76			
IHISP295	Hispanic Interaction Of BFNOTLN	71.13							
IHISP296	Hispanic Interaction Of BFNOTQU	37.92							
IHISP297	Hispanic Interaction Of BFNOTCU	64.52							
IHISP298	Hispanic Interaction Of BFNOTQR	1.22							
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut			0.00			13.06	0.16	0.00
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut			0.60			0.48	0.24	
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut			0.00			8.48	29.84	
BINDIAQR	B: Quartic: Percent Pop: American Indian, Eskimo, Aleut						0.00	1.39	
IBLK300	Black Interaction Of BINDIALN							28.21	
IBLK301	Black Interaction Of BINDIAQU							18.39	
IBLK302	Black Interaction Of BINDIACU							87.54	
IBLK303	Black Interaction Of BINDIAQR							3.85	
IHISP300	Hispanic Interaction Of BINDIALN							2.15	
IHISP301	Hispanic Interaction Of BINDIAQU							1.22	
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.00			0.16	35.74		0.00
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.16				74.12		0.00
BPOVERCU	B: Cubic: Percent Families Below Poverty Level		45.05				0.03		
IOTH310	Other Interaction Of BPOVERLN		0.59						
IOTH311	Other Interaction Of BPOVERQU		0.51						
IOTH312	Other Interaction Of BPOVERCU		0.00						
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican					62.71	0.00		
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican					2.66			
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					72.91			
BPRICAQR	B: Quartic: Percent Hispanics: Puerto Rican					0.10			
IFEM315	Female Interaction Of BPRICALN					3.55			
BSCHASLN	B: Linear: Percent Associates Degree				0.00				

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
BSCHASQU	B: Quadratic: Percent Associates Degree				75.20				
BSCHASCU	B: Cubic: Percent Associates Degree				0.00				
IHISP320	Hispanic Interaction Of BSCHASLN				0.00				
IHISP321	Hispanic Interaction Of BSCHASQU				0.00				
IHISP322	Hispanic Interaction Of BSCHASCU				0.00				
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander		0.00				0.00		
PCUBANLN	T: Linear: Percent Hispanics: Cuban	57.23		11.78				5.25	0.00
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban	87.26		0.00				0.25	0.00
PCUBANCU	T: Cubic: Percent Hispanics: Cuban							13.14	0.00
PCUBANQR	T: Quartic: Percent Hispanics: Cuban							0.00	0.00
IFEM330	Female Interaction Of PCUBANLN	1.97							
IBLK330	Black Interaction Of PCUBANLN	71.96							55.75
IBLK331	Black Interaction Of PCUBANQU	0.02							0.75
IBLK332	Black Interaction Of PCUBANCU								24.02
IBLK333	Black Interaction Of PCUBANQR								0.00

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F10. Significance Probabilities for Fixed Effect Coefficients for Past Year Arrested by Age

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMBLCK	O: Black Interaction Of FEMALE	48.80	0.03	56.74	0.00	57.42	22.89	41.07	0.09
FEMHISP	O: Hispanic Interaction Of FEMALE	29.73	0.00	0.05	0.00	62.42	49.28	90.88	5.05
RACEBLCK	O: Race/Black Indicator	12.36	0.00	0.03	0.00	58.00	23.82	4.25	97.92
RACEHISP	O: Race/Hispanic Indicator	22.85	55.04	0.11	5.57	20.39	32.54	32.31	0.00
RACEOTHR	O: Race/Other Indicator	0.51	0.00			22.45	2.82		
REGNOREA	O: Northeast Region Indicator					52.14			7.59
REGSOUTH	O: South Region Indicator					11.27			13.43
REGWEST	O: West Region Indicator					54.06			51.68
PDENLEV1	O: Large MSA						15.24	0.44	
PDENLEV2	O: Medium MSA						25.55	0.08	
PDENLEV3	O: Small MSA						97.72	0.00	
PDENLEV4	O: NonMSA, Urban						32.15	0.00	
IBLK10	Black Interaction Of PDENLEV1						44.04		
IBLK11	Black Interaction Of PDENLEV2						41.26		
IBLK12	Black Interaction Of PDENLEV3						1.16		
IBLK13	Black Interaction Of PDENLEV4						26.12		
PHH1PLN	T: Linear: Percent One Person Households		0.00				2.03		
PHH1PQU	T: Quadratic: Percent One Person Households		11.46				5.53		
PHH1PCU	T: Cubic: Percent One Person Households		39.48				0.00		
PHH1PQR	T: Quartic: Percent One Person Households		0.06				0.04		
IBLK15	Black Interaction Of PHH1PLN		69.17				8.32		
IBLK16	Black Interaction Of PHH1PQU		79.59				68.31		
IBLK17	Black Interaction Of PHH1PCU		1.78				2.22		
IBLK18	Black Interaction Of PHH1PQR		0.00				3.89		
POPRMLN	T: Linear: Average Persons Per Room		11.41				0.10		
PAGE34LN	T: Linear: Percent Persons 25-34 Years								0.00
PAGE54LN	T: Linear: Percent Persons 45-54 Years							8.43	
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years							10.86	
PAGE54CU	T: Cubic: Percent Persons 45-54 Years							0.00	
IHISP45	Hispanic Interaction Of PAGE54LN							43.22	
IHISP46	Hispanic Interaction Of PAGE54QU							21.15	
IHISP47	Hispanic Interaction Of PAGE54CU							1.71	
PAGE64LN	T: Linear: Percent Persons 55-64 Years					0.14			
PSCH8LN	T: Linear: Percent 0-8 Years Of School	4.50							
PSCH8QU	T: Quadratic: Percent 0-8 Years Of School	58.52							
IOTH55	Other Interaction Of PSCH8LN	10.17							
IOTH56	Other Interaction Of PSCH8QU	0.97							
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma		0.00				0.16		
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma						2.03		
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma						0.05		
IHISP60	Hispanic Interaction Of PSCH12LN						0.92		
PSCHCOLN	T: Linear: Bachelors, Graduate, Or Professional Degree								0.00
PSCHCOQU	T: Quadratic: Bachelors, Graduate, Or Professional Degree								12.60
PSCHCOCU	T: Cubic: Bachelors, Graduate, Or Professional Degree								0.00
IBLK75	Black Interaction Of PSCHCOLN								4.05
IBLK76	Black Interaction Of PSCHCOQU								33.42
IBLK77	Black Interaction Of PSCHCOCU								0.00
PPUBASLN	T: Linear: Percent Households With Public Assist Income						0.00		
P64DISLN	T: Linear: Percent 16-64 With A Work Disability			0.00				0.00	
PBLACKLN	T: Linear: Percent Black Nonhispanic		1.37						
PBLACKQU	T: Quadratic: Percent Black Nonhispanic		0.00						
POTHLN	T: Linear: Percent Other Race/Hispanicity				0.00				
IHISP110	Hispanic Interaction Of POTHLN				0.00				
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force				0.00				4.67
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force								0.00
IFEM125	Female Interaction Of PFLABLN				0.00				
PFNOTLN	T: Linear: Percent Females Separated, Divorced Or Widowed							0.28	
PFNOTQU	T: Quadratic: Percent Females Separated, Divorced Or Widowed							0.00	
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force				3.07		0.02		0.00
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force						0.02		
IFEM145	Female Interaction Of PMLABLN				0.00				
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed						63.85		
PMNOTQU	T: Quadratic: Percent Males Separated, Divorced Or Widowed						5.35		
PMNOTCU	T: Cubic: Percent Males Separated, Divorced Or Widowed						20.42		
PMNOTQR	T: Quartic: Percent Males Separated, Divorced Or Widowed						0.00		
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier	0.43				69.50			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
POLDHUQU	T: Quadratic: Percent Housing Units Built 1939 Or Earlier					52.75			
POLDHUCU	T: Cubic: Percent Housing Units Built 1939 Or Earlier					1.14			
P40HULN	T: Linear: Percent Housing Units Built 1940-1949	0.00						57.80	
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949							0.00	
PRENTLN	T: Linear: Percent Housing Rented					0.00			
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol		69.93	0.00					
IFEM180	Female Interaction Of ADRAT1LN		0.00						
V18FLN	C: Linear: Marijuana Posession Arrest Rate			0.00				0.00	
V18FQU	C: Quadratic: Marijuana Posession Arrest Rate							0.00	
V18ELN	C: Linear: Opium/Cocaine & Deriv Posession Arrest Rate			0.00	1.21			0.00	61.90
V18EQU	C: Quadratic: Opium/Cocaine & Deriv Posession Arrest Rate				0.00				0.00
IFEM195	Female Interaction Of V18ELN								0.00
IFEM196	Female Interaction Of V18EQU								0.00
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate							0.01	
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992	10.41							
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992	88.91							
DRATECU	C: Cubic: Mean Drug Client Treatment Rate 1991 & 1992	2.64							
RH43ALN	T: Linear: Recoded Median Rents For Rental Units				0.00			31.59	0.00
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units							32.80	
RH43ACU	T: Cubic: Recoded Median Rents For Rental Units							15.38	
RH43AQR	T: Quartic: Recoded Median Rents For Rental Units							0.00	
RH61ALN	T: Linear: Recoded Median Value Of Owner Occupied HUs		85.17	0.00			12.59	58.62	
RH61AQU	T: Quadratic: Recoded Median Value Of Owner Occupied HUs			0.00			1.43	99.01	
RH61ACU	T: Cubic: Recoded Median Value Of Owner Occupied HUs						0.03	0.01	
IHISP245	Hispanic Interaction Of RH61ALN							25.33	
IHISP246	Hispanic Interaction Of RH61AQU							17.71	
IHISP247	Hispanic Interaction Of RH61ACU							5.11	
ARATELN	C: Linear: Mean A-Only Client Treatment Rate 1991 & 1992								74.64
ARATEQU	C: Quadratic: Mean A-Only Client Treatment Rate 1991 & 1992								2.53
ARATECU	C: Cubic: Mean A-Only Client Treatment Rate 1991 & 1992								0.00
ARATEQR	C: Quartic: Mean A-Only Client Treatment Rate 1991 & 1992								0.42
IBLK255	Black Interaction Of ARATELN								80.25
IBLK256	Black Interaction Of ARATEQU								1.24
IBLK257	Black Interaction Of ARATECU								55.85
IBLK258	Black Interaction Of ARATEQR								0.02
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability								0.01
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability								13.60
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability								0.00
B64DISQR	B: Quartic: Percent Persons 16-64 With A Work Disability								0.00
BAGE34LN	B: Linear: Percent Persons 25-34 Years				0.00		71.42		
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years				1.30		6.91		
BAGE34CU	B: Cubic: Percent Persons 25-34 Years						0.07		
BAGE34QR	B: Quartic: Percent Persons 25-34 Years						28.51		
IFEM265	Female Interaction Of BAGE34LN						11.54		
IFEM266	Female Interaction Of BAGE34QU						3.06		
IFEM267	Female Interaction Of BAGE34CU						98.61		
IFEM268	Female Interaction Of BAGE34QR						0.00		
IHISP265	Hispanic Interaction Of BAGE34LN				80.41				
IHISP266	Hispanic Interaction Of BAGE34QU				0.00				
BAGE44LN	B: Linear: Percent Persons 35-44 Years						1.16	95.30	
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years						0.02	81.49	
BAGE44CU	B: Cubic: Percent Persons 35-44 Years						22.01	0.02	
BAGE44QR	B: Quartic: Percent Persons 35-44 Years						0.00		
IBLK270	Black Interaction Of BAGE44LN						39.46		
IBLK271	Black Interaction Of BAGE44QU						32.07		
IBLK272	Black Interaction Of BAGE44CU						41.39		
IBLK273	Black Interaction Of BAGE44QR						0.13		
IHISP270	Hispanic Interaction Of BAGE44LN							49.68	
IHISP271	Hispanic Interaction Of BAGE44QU							2.28	
BAGE54LN	B: Linear: Percent Persons 45-54 Years						0.01		
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years						21.09		
BAGE54CU	B: Cubic: Percent Persons 45-54 Years						1.70		
BAGE54QR	B: Quartic: Percent Persons 45-54 Years						0.00		
BAGE64LN	B: Linear: Percent Persons 55-64 Years		0.00			0.13	23.83		
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years						1.55		
BAGE64CU	B: Cubic: Percent Persons 55-64 Years						7.40		
IHISP280	Hispanic Interaction Of BAGE64LN		0.00				83.83		
IHISP281	Hispanic Interaction Of BAGE64QU						47.75		
IHISP282	Hispanic Interaction Of BAGE64CU						8.08		
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut		86.71			4.08			
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut					47.84			
BINDIACU	B: Cubic: Percent Pop: American Indian, Eskimo, Aleut					0.91			
IOTH300	Other Interaction Of BINDIALN		0.00						
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed	31.53							
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed	0.00							
BPOVERLN	B: Linear: Percent Families Below Poverty Level							20.47	
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level							7.60	
BPOVERCU	B: Cubic: Percent Families Below Poverty Level							0.00	
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican					28.41			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican					9.86			
BPRICACU	B: Cubic: Percent Hispanics: Puerto Rican					12.75			
IFEM315	Female Interaction Of BPRICALN					16.41			
IFEM316	Female Interaction Of BPRICAQU					49.28			
IFEM317	Female Interaction Of BPRICACU					4.19			
BSCHASLN	B: Linear: Percent Associates Degree			0.00		44.93		97.97	
BSCHASQU	B: Quadratic: Percent Associates Degree					65.53		50.87	
BSCHASCU	B: Cubic: Percent Associates Degree					0.03		40.06	
BSCHASQR	B: Quartic: Percent Associates Degree							0.00	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander	4.73				39.47			
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander					39.05			
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander					1.23			
IFEM325	Female Interaction Of PASIANLN					31.71			
IFEM326	Female Interaction Of PASIANQU					77.29			
IFEM327	Female Interaction Of PASIANCU					0.43			
PCUBANLN	T: Linear: Percent Hispanics: Cuban							70.18	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban							66.99	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban							0.00	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

F11. Significance Probabilities for Fixed Effect Coefficients for Past Year Treatment Needed for Drug Abuse by Age

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
DUMMY	O: Intercept Term	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FEMALE	O: Female Indicator	7.65	0.00	0.00	0.00	92.83	0.00	0.22	0.00
FEMBLCK	O: Black Interaction Of FEMALE	19.99	35.11	0.61	0.00		28.88	1.63	21.27
FEMHISP	O: Hispanic Interaction Of FEMALE	0.80	9.91	0.00	58.34		85.11	7.65	20.98
FEMOTHR	O: Other Interaction Of FEMALE	57.47	5.65	0.22			0.06	84.98	0.00
RACEBLCK	O: Race/Black Indicator	8.01	0.01	85.08	0.00		14.20	0.02	0.00
RACEHISP	O: Race/Hispanic Indicator	9.19	0.00	0.00	0.00		4.24	40.03	50.49
RACEOTHR	O: Race/Other Indicator	0.16	0.00	0.00	0.00		0.05	0.00	2.56
PDENLEV1	O: Large MSA						13.27		
PDENLEV2	O: Medium MSA						24.72		
PDENLEV3	O: Small MSA						75.63		
PDENLEV4	O: NonMSA, Urban						92.29		
IBLK10	Black Interaction Of PDENLEV1						69.86		
IBLK11	Black Interaction Of PDENLEV2						54.22		
IBLK12	Black Interaction Of PDENLEV3						14.95		
IBLK13	Black Interaction Of PDENLEV4						14.56		
PHH1PLN	T: Linear: Percent One Person Households				0.00				0.00
PHH1PQU	T: Quadratic: Percent One Person Households				0.00				0.00
PHH1PCU	T: Cubic: Percent One Person Households				1.43				0.00
PHH1PQR	T: Quartic: Percent One Person Households				0.00				0.00
PAGE18LN	T: Linear: Percent Persons 0-18 Years						48.37		
PAGE18QU	T: Quadratic: Percent Persons 0-18 Years						0.05		
PAGE18CU	T: Cubic: Percent Persons 0-18 Years						2.40		
IOTH25	Other Interaction Of PAGE18LN						79.97		
IOTH26	Other Interaction Of PAGE18QU						0.01		
PAGE24LN	T: Linear: Percent Persons 19-24 Years								0.00
PAGE24QU	T: Quadratic: Percent Persons 19-24 Years								0.00
PAGE24CU	T: Cubic: Percent Persons 19-24 Years								0.08
PAGE24QR	T: Quartic: Percent Persons 19-24 Years								0.00
PAGE44LN	T: Linear: Percent Persons 35-44 Years				0.00				
PAGE54LN	T: Linear: Percent Persons 45-54 Years				0.52		0.01		0.00
PAGE54QU	T: Quadratic: Percent Persons 45-54 Years								0.07
PAGE54CU	T: Cubic: Percent Persons 45-54 Years								0.00
PAGE54QR	T: Quartic: Percent Persons 45-54 Years								0.00
IFEM45	Female Interaction Of PAGE54LN				0.00				
IHISP45	Hispanic Interaction Of PAGE54LN						2.75		
PSCH12LN	T: Linear: Percent 9-12 Years & No High School Diploma							0.00	
PSCH12QU	T: Quadratic: Percent 9-12 Years & No High School Diploma							0.31	
PSCH12CU	T: Cubic: Percent 9-12 Years & No High School Diploma							6.24	
IFEM60	Female Interaction Of PSCH12LN							67.76	
IFEM61	Female Interaction Of PSCH12QU							0.04	
PPOVERLN	T: Linear: Percent Families Below Poverty Level					69.30			
PPOVERQU	T: Quadratic: Percent Families Below Poverty Level					64.31			
PPOVERCU	T: Cubic: Percent Families Below Poverty Level					0.24			
PBLACKLN	T: Linear: Percent Black Nonhispanic							0.04	
PBLACKQU	T: Quadratic: Percent Black Nonhispanic							0.00	
IOTH100	Other Interaction Of PBLACKLN							0.00	
IOTH101	Other Interaction Of PBLACKQU							0.01	
PFLABLN	T: Linear: Percent Females 16+ Years Old In Labor Force					7.88	9.40		
PFLABQU	T: Quadratic: Percent Females 16+ Years Old In Labor Force					2.02	4.63		
PFLABCU	T: Cubic: Percent Females 16+ Years Old In Labor Force						58.51		
PFLABQR	T: Quartic: Percent Females 16+ Years Old In Labor Force						0.01		
IBLK125	Black Interaction Of PFLABLN						95.71		
IBLK126	Black Interaction Of PFLABQU						2.17		
PFNEVLN	T: Linear: Percent Females Never Married							3.55	
PFNEVQU	T: Quadratic: Percent Females Never Married							0.00	
IFEM130	Female Interaction Of PFNEVLN							55.72	
IFEM131	Female Interaction Of PFNEVQU							1.73	
PMNEVLN	T: Linear: Percent Males Never Married				61.10				0.00
PMNEVQU	T: Quadratic: Percent Males Never Married				0.00				0.00
PMNEVCU	T: Cubic: Percent Males Never Married								0.00
PMLABLN	T: Linear: Percent Males 16+ Years Old In Labor Force								13.88
PMLABQU	T: Quadratic: Percent Males 16+ Years Old In Labor Force								0.26
PMLABCU	T: Cubic: Percent Males 16+ Years Old In Labor Force								0.00
PMLABQR	T: Quartic: Percent Males 16+ Years Old In Labor Force								0.00
PMNOTLN	T: Linear: Percent Males Separated, Divorced Or Widowed							0.02	
POLDHULN	T: Linear: Percent Housing Units Built 1939 Or Earlier					16.13			
P40HULN	T: Linear: Percent Housing Units Built 1940-1949	55.11				44.36			
P40HUQU	T: Quadratic: Percent Housing Units Built 1940-1949	20.64				61.03			
P40HUCU	T: Cubic: Percent Housing Units Built 1940-1949					76.71			

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
IHISP160	Hispanic Interaction Of P40HULN	50.65							
IHISP161	Hispanic Interaction Of P40HUQU	11.69							
PRENTLN	T: Linear: Percent Housing Rented			0.00				41.92	0.00
PRENTQU	T: Quadratic: Percent Housing Rented							2.13	
PRENTCU	T: Cubic: Percent Housing Rented							67.57	
PRENTQR	T: Quartic: Percent Housing Rented							20.14	
IHISP165	Hispanic Interaction Of PRENTLN							26.75	
IHISP166	Hispanic Interaction Of PRENTQU							43.50	
IHISP167	Hispanic Interaction Of PRENTCU							12.01	
ADRATELN	C: Linear: Death Rate For All Alcohol-Related Cases						49.93		
ADRATEQU	C: Quadratic: Death Rate For All Alcohol-Related Cases						0.00		
ADRATECU	C: Cubic: Death Rate For All Alcohol-Related Cases						32.60		
ADRATEQR	C: Quartic: Death Rate For All Alcohol-Related Cases						0.99		
IBLK175	Black Interaction Of ADRATELN						66.11		
IBLK176	Black Interaction Of ADRATEQU						37.95		
IBLK177	Black Interaction Of ADRATECU						23.28		
IBLK178	Black Interaction Of ADRATEQR						26.61		
ADRAT1LN	C: Linear: Death Rate With Explicit Mention Of Alcohol								0.00
ADRAT1QU	C: Quadratic: Death Rate With Explicit Mention Of Alcohol								0.11
IFEM180	Female Interaction Of ADRAT1LN								0.00
IFEM181	Female Interaction Of ADRAT1QU								0.00
V18FLN	C: Linear: Marijuana Posession Arrest Rate				0.00				
V18BLN	C: Linear: Marijuana Sale/Manufacture Arrest Rate			0.02				9.17	91.31
V18BQU	C: Quadratic: Marijuana Sale/Manufacture Arrest Rate							14.45	0.54
V18BCU	C: Cubic: Marijuana Sale/Manufacture Arrest Rate							0.00	0.00
V18BQR	C: Quartic: Marijuana Sale/Manufacture Arrest Rate								0.00
IFEM190	Female Interaction Of V18BLN								29.94
IFEM191	Female Interaction Of V18BQU								0.00
IFEM192	Female Interaction Of V18BCU								0.00
IFEM193	Female Interaction Of V18BQR								0.00
IBLK190	Black Interaction Of V18BLN							29.26	
IBLK191	Black Interaction Of V18BQU							19.27	
IBLK192	Black Interaction Of V18BCU							1.36	
IHISP190	Hispanic Interaction Of V18BLN			0.00					
V18ALN	C: Linear: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate								0.00
V18AQU	C: Quadratic: Opium/Cocaine & Deriv Sale/Manuf Arrest Rate								0.00
IHISP200	Hispanic Interaction Of V18ALN								17.14
IHISP201	Hispanic Interaction Of V18AQU								0.00
V18LN	C: Linear: Total Drug Abuse Violations Arrest Rate						0.05		
V18QU	C: Quadratic: Total Drug Abuse Violations Arrest Rate						0.10		
DRATELN	C: Linear: Mean Drug Client Treatment Rate 1991 & 1992						7.48		
DRATEQU	C: Quadratic: Mean Drug Client Treatment Rate 1991 & 1992						31.24		
IBLK235	Black Interaction Of DRATELN						54.88		
IBLK236	Black Interaction Of DRATEQU						70.51		
RH43ALN	T: Linear: Recoded Median Rents For Rental Units								71.29
RH43AQU	T: Quadratic: Recoded Median Rents For Rental Units								0.00
IBLK240	Black Interaction Of RH43ALN								0.50
IBLK241	Black Interaction Of RH43AQU								0.00
RP80ALN	T: Linear: Recoded Median Household Income		0.00						
IBLK250	Black Interaction Of RP80ALN		0.00						
B64DISLN	B: Linear: Percent Persons 16-64 With A Work Disability						0.98	0.00	
B64DISQU	B: Quadratic: Percent Persons 16-64 With A Work Disability							9.12	
B64DISCU	B: Cubic: Percent Persons 16-64 With A Work Disability							0.02	
IHISP260	Hispanic Interaction Of B64DISLN							72.92	
IHISP261	Hispanic Interaction Of B64DISQU							95.80	
IHISP262	Hispanic Interaction Of B64DISCU							14.57	
IOTH260	Other Interaction Of B64DISLN							1.66	
IOTH261	Other Interaction Of B64DISQU							0.00	
BAGE34LN	B: Linear: Percent Persons 25-34 Years						66.62		0.13
BAGE34QU	B: Quadratic: Percent Persons 25-34 Years								49.08
BAGE34CU	B: Cubic: Percent Persons 25-34 Years								0.04
BAGE34QR	B: Quartic: Percent Persons 25-34 Years								1.06
IBLK265	Black Interaction Of BAGE34LN								18.20
IBLK266	Black Interaction Of BAGE34QU								31.12
IBLK267	Black Interaction Of BAGE34CU								43.05
IBLK268	Black Interaction Of BAGE34QR								0.06
BAGE44LN	B: Linear: Percent Persons 35-44 Years	72.09				0.10		69.21	45.14
BAGE44QU	B: Quadratic: Percent Persons 35-44 Years								90.57
BAGE44CU	B: Cubic: Percent Persons 35-44 Years								0.00
IOTH270	Other Interaction Of BAGE44LN	0.15						0.01	
BAGE54LN	B: Linear: Percent Persons 45-54 Years					95.24	0.03		
BAGE54QU	B: Quadratic: Percent Persons 45-54 Years					27.71			
BAGE54CU	B: Cubic: Percent Persons 45-54 Years					38.39			
BAGE54QR	B: Quartic: Percent Persons 45-54 Years					2.77			
IOTH275	Other Interaction Of BAGE54LN						0.00		
BAGE64LN	B: Linear: Percent Persons 55-64 Years		0.00				40.18		
BAGE64QU	B: Quadratic: Percent Persons 55-64 Years		28.61				0.39		
BAGE64CU	B: Cubic: Percent Persons 55-64 Years		0.01				9.70		

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY 12-17	BIG CITY 18-25	BIG CITY 26-34	BIG CITY 35+	REMAINDER 12-17	REMAINDER 18-25	REMAINDER 26-34	REMAINDER 35+
IBLK280	Black Interaction Of BAGE64LN						9.66		
IBLK281	Black Interaction Of BAGE64QU						77.29		
IBLK282	Black Interaction Of BAGE64CU						19.48		
IHISP280	Hispanic Interaction Of BAGE64LN		0.00						
IHISP281	Hispanic Interaction Of BAGE64QU		0.67						
IHISP282	Hispanic Interaction Of BAGE64CU		0.00						
BCUBANLN	B: Linear: Percent Hispanics: Cuban								0.01
BCUBANQU	B: Quadratic: Percent Hispanics: Cuban								0.00
BFNOTLN	B: Linear: Percent Females Separated, Divorced or Widowed	43.24	0.06			40.71	3.68	5.21	
BFNOTQU	B: Quadratic: Percent Females Separated, Divorced or Widowed		53.26			10.52	1.17	0.00	
BFNOTCU	B: Cubic: Percent Females Separated, Divorced or Widowed					80.21	6.69	2.52	
BFNOTQR	B: Quartic: Percent Females Separated, Divorced or Widowed					7.38	56.30	36.06	
IBLK295	Black Interaction Of BFNOTLN						20.11		
IBLK296	Black Interaction Of BFNOTQU						4.19		
IBLK297	Black Interaction Of BFNOTCU						38.34		
IBLK298	Black Interaction Of BFNOTQR						12.24		
IHISP295	Hispanic Interaction Of BFNOTLN		0.49				21.23		
IHISP296	Hispanic Interaction Of BFNOTQU		0.04				18.49		
IHISP297	Hispanic Interaction Of BFNOTCU						11.71		
IOTH295	Other Interaction Of BFNOTLN	1.49						0.24	
IOTH296	Other Interaction Of BFNOTQU							0.00	
IOTH297	Other Interaction Of BFNOTCU							24.95	
IOTH298	Other Interaction Of BFNOTQR							0.00	
BINDIALN	B: Linear: Percent Pop: American Indian, Eskimo, Aleut						0.89		23.95
BINDIAQU	B: Quadratic: Percent Pop: American Indian, Eskimo, Aleut						0.76		0.00
IFEM300	Female Interaction Of BINDIALN						49.06		
IFEM301	Female Interaction Of BINDIAQU						3.44		
IHISP300	Hispanic Interaction Of BINDIALN								30.92
IHISP301	Hispanic Interaction Of BINDIAQU								0.45
BMNOTLN	B: Linear: Percent Males Separated, Divorced or Widowed							18.85	
BMNOTQU	B: Quadratic: Percent Males Separated, Divorced or Widowed							60.11	
BMNOTCU	B: Cubic: Percent Males Separated, Divorced or Widowed							37.47	
BMNOTQR	B: Quartic: Percent Males Separated, Divorced or Widowed							1.37	
IHISP305	Hispanic Interaction Of BMNOTLN							57.75	
IHISP306	Hispanic Interaction Of BMNOTQU							94.14	
IHISP307	Hispanic Interaction Of BMNOTCU							4.89	
IHISP308	Hispanic Interaction Of BMNOTQR							6.21	
BPOVERLN	B: Linear: Percent Families Below Poverty Level		0.00		0.00		17.78		0.00
BPOVERQU	B: Quadratic: Percent Families Below Poverty Level		0.00		0.00				0.00
BPOVERCU	B: Cubic: Percent Families Below Poverty Level								5.41
BPOVERQR	B: Quartic: Percent Families Below Poverty Level								0.00
IFEM310	Female Interaction Of BPOVERLN				0.95		0.04		30.50
IFEM311	Female Interaction Of BPOVERQU				0.00				0.00
IFEM312	Female Interaction Of BPOVERCU								0.00
IFEM313	Female Interaction Of BPOVERQR								0.00
IBLK310	Black Interaction Of BPOVERLN		0.00						
IBLK311	Black Interaction Of BPOVERQU		0.03						
BPRICALN	B: Linear: Percent Hispanics: Puerto Rican						0.00		
BPRICAQU	B: Quadratic: Percent Hispanics: Puerto Rican						35.86		
IBLK315	Black Interaction Of BPRICALN						3.36		
IBLK316	Black Interaction Of BPRICAQU						0.15		
BSCHASLN	B: Linear: Percent Associates Degree			7.11				3.48	0.00
BSCHASQU	B: Quadratic: Percent Associates Degree							53.40	
BSCHASCU	B: Cubic: Percent Associates Degree							0.61	
BSCHASQR	B: Quartic: Percent Associates Degree							0.08	
IFEM320	Female Interaction Of BSCHASLN							98.07	
IFEM321	Female Interaction Of BSCHASQU							68.29	
IFEM322	Female Interaction Of BSCHASCU							34.09	
IFEM323	Female Interaction Of BSCHASQR							0.19	
IBLK320	Black Interaction Of BSCHASLN			0.00					
IOTH320	Other Interaction Of BSCHASLN							0.92	
IOTH321	Other Interaction Of BSCHASQU							0.00	
PASIANLN	T: Linear: Percent Population: Asian, Pacific Islander				0.01			0.12	
PASIANQU	T: Quadratic: Percent Population: Asian, Pacific Islander							15.38	
PASIANCU	T: Cubic: Percent Population: Asian, Pacific Islander							0.14	
IBLK325	Black Interaction Of PASIANLN							41.19	
IBLK326	Black Interaction Of PASIANQU							23.63	
IBLK327	Black Interaction Of PASIANCU							0.12	
IOTH325	Other Interaction Of PASIANLN				0.00				
PCUBANLN	T: Linear: Percent Hispanics: Cuban			35.61				65.83	
PCUBANQU	T: Quadratic: Percent Hispanics: Cuban			0.00				0.02	
PCUBANCU	T: Cubic: Percent Hispanics: Cuban			0.00				4.15	
PCUBANQR	T: Quartic: Percent Hispanics: Cuban							2.37	
PINDIALN	T: Linear: Percent Pop: American Indian, Eskimo, Aleut					24.02	0.76	3.33	
PINDIAQU	T: Quadratic: Percent Pop: American Indian, Eskimo, Aleut					4.85	0.07	25.37	
PINDIACU	T: Cubic: Percent Pop: American Indian, Eskimo, Aleut					5.68		0.01	
IFEM335	Female Interaction Of PINDIALN						46.11		
IFEM336	Female Interaction Of PINDIAQU						2.05		
IOTH335	Other Interaction Of PINDIALN							14.52	
IOTH336	Other Interaction Of PINDIAQU							0.12	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Variable	Label	BIG CITY				REMAINDER			
		12-17	18-25	26-34	35+	12-17	18-25	26-34	35+
IOTH337	Other Interaction Of PINDIACU							0.00	

NOTE: T: Indicates a tract-level variable, C: Indicates a county-level variable, B: Indicates a block-level variable, O: Other

Appendix G: Details of Estimating Variances

Appendix G. Details of Estimating Variances

G.1 Derivation of Formulas

As discussed in Section 2.6, we want to estimate[1]

$$\text{MSE}_i = E(\hat{\pi}_i^W - \pi_i)^2 \quad , \tag{G.1}$$

where π_i is the sum of person-level propensities for area i. Expanding **(G.1)**, we have that it is equal to

$$\begin{aligned} \text{MSE}_i &= E\left[\sum_a \sum_d \sum_b N_{ib}^{ad} \hat{\pi}_{ib}^{ad} - \sum_a \sum_d \sum_b N_{ib}^{ad} \pi_{ib}^{ad}\right]^2 \\ &\doteq \sum_a E\left[\sum_d \sum_b N_{ib}^{ad}\left(\hat{\pi}_{ib}^{ad} - \pi_{ib}^{ad}\right)\right]^2 = \sum_a \text{MSE}_i^a \end{aligned} \tag{G.2}$$

This approximation ignores the covariances of the errors across the domains for which separate models were fit. Recall that there were four such domains, defined by age, as discussed in Section 2.5. We believe that these covariances are mostly positive since over- or underestimation of the substance abuse rate for one age domain in an area is probably accompanied by similar over- or underestimation of the rate for the other age domains in the same area, but the sizes and signs of the covariances are unknown. We were forced to ignore them by the computing constraints.

To simplify notation, we will now focus on a particular area *i* and age group *a*, dropping the *a* subscript since the age groups were modeled separately. We may then write the inner expectation in **(G.2)** as

$$\text{MSE}_i^a = E\left[\sum_d \sum_b N_{ib}^d (\hat{\pi}_{ib}^d - \pi_{ib}^d)\right]^2 = E\left[\tilde{N}^t (\hat{\pi} - \pi)\right]^2 \tag{G.3}$$

where $\tilde{N}$ is the column matrix of populations for age group *a* by domain *d* for the various block groups in the *i*-th area, $\hat{\pi}$ is the column matrix of estimated propensities to engage in the behavior

1 "Mean square error" is used to mean the same thing that is labeled "variance" in sections 1.6 and 2.6. Neither term is wholely satisfactory. "Mean square prediction error" would probably be more appropriate but is rather long.

of interest, and π is the column matrix of propensities given the manifestation of state and PSU effects.

Since $\tilde{N}$ is fixed, we may rewrite **(G.3)** as

$$\mathrm{MSE}_i^a = \tilde{N}^t \, \mathrm{E}\left[(\hat{\pi}-\pi)(\hat{\pi}-\pi)^t\right]\tilde{N} \tag{G.4}$$

Since $\hat{\pi}$ and π are nonlinear functions of $(\hat{\beta}, \hat{U})$ and (β, U), respectively, the inner expectation of $\hat{\pi}$ **(G.4)** is difficult to compute exactly. Following McGibbon and Tomerlin (1989), we used a first-order Taylor linearization to simplify the calculation. Focusing on the predicted propensity for a specific domain and block group, and treating it as a function of $(\hat{\beta}, \hat{U})$, we have that

$$\hat{\pi}_{ib}^{d}(\hat{\beta}, \hat{U}) \doteq \hat{\pi}_{ib}^{d}(\beta, U) + \begin{bmatrix} \dfrac{\partial \hat{\pi}_{ib}^{d}}{\partial \hat{\beta}}(\beta, U) \\ \dfrac{\partial \hat{\pi}_{ib}^{d}}{\partial \hat{U}}(\beta, U) \end{bmatrix}^t \begin{bmatrix} \hat{\beta}-\beta \\ \hat{U}-U \end{bmatrix} \tag{G.5}$$

The first partial derivative of $\hat{\pi}_{ib}^{d}$ with respect to $\hat{\beta}$ is

$$\partial \frac{\hat{\pi}_{ib}^{d}}{\partial \hat{\beta}} = \frac{\partial}{\partial \hat{\beta}} \left[\frac{1}{1+\exp[-(X_{ib}^{d}\hat{\beta} + Z_{ib}\hat{U})]} \right]$$

$$= \frac{-\dfrac{\partial}{\partial \hat{\beta}} \left[1+\exp[-(X_{ib}^{d}\hat{\beta} + Z_{ib}\hat{U})]\right]}{\{1+\exp[-(X_{ib}^{d}\hat{\beta} + Z_{ib}\hat{U})]\}^2}$$

$$= \left(X_{ib}^{d}\right)^t \frac{\exp[-(X_{ib}^{d}\hat{\beta} + Z_{ib}\hat{U})]}{\{1+\exp[-(X_{ib}^{d}\hat{\beta} + Z_{ib}\hat{U})]\}^2} \quad ,$$

where X_{ib}^d and Z_{ib} are defined as in Sections 2.5 and 2.6. Evaluating at (β, U), we have

$$\frac{\partial \hat{\pi}_{ib}^d}{\partial \hat{\beta}}(\beta,U) = \left(X_{ib}^d\right)^t \frac{\exp[-(X_{ib}^d\beta + Z_{ib}U)]}{\{1+\exp[-(X_{ib}^d\beta + Z_{ib}U)]\}^2} = \left(X_{ib}^d\right)^t \pi_{ib}^d(1-\pi_{ib}^d)$$

Similarly,

$$\frac{\partial \hat{\pi}_{ib}^d}{\partial \hat{U}}(\beta,U) = Z_{ib}^t \; \pi_{ib}^d(1-\pi_{ib}^d) \quad .$$

Substituting into **(G.5)** and noting that $\hat{\pi}_{ib}^d(\beta,U) = \pi_{ib}^d$, we have that

$$\hat{\pi}_{ib}^d(\hat{\beta},\hat{U}) \doteq \pi_{ib}^d + \pi_{ib}^d(1-\pi_{ib}^d)\; [X_{ib}^d \; Z_{ib}] \begin{bmatrix} \hat{\beta}-\beta \\ \hat{U}-U \end{bmatrix}$$

$$= \pi_{ib}^d + \pi_{ib}^d(1-\pi_{ib}^d)[X_{ib}^d(\hat{\beta}-\beta) + Z_{ib}(\hat{U}-U)] \quad .$$

Thus $$\hat{\pi}_{ib}^d - \pi_{ib}^d \doteq \pi_{ib}^d(1-\pi_{ib}^d)[X_{ib}^d(\hat{\beta}-\beta) + Z_{ib}(\hat{U}-U)] \qquad \textbf{(G.6)}$$

Substituting into **(G.4)**, we have

$$MSE_i^a = \tilde{N}^t \, E\left\{[\overset{*}{C}\overset{*}{X}(\hat{\beta}-\beta)+\overset{*}{C}\overset{*}{Z}(\hat{U}-U)][\overset{*}{C}\overset{*}{X}(\hat{\beta}-\beta)+\overset{*}{C}\overset{*}{Z}(\hat{U}-U)]^t\right\}\tilde{N} \qquad \textbf{(G.7)}$$

where an asterisk over a C, X, or Z indicates that the matrix has one row for every domain within every block group in the area as opposed to having one row per sample person as was the case for the model fitting.

We now make an additional simplifying assumption, treating C as fixed even through it must be estimated. With that additional assumption, we have that

$$
\begin{aligned}
MSE_i^a &\doteq \tilde{N}^t \overset{*}{C} E\left\{[\overset{*}{X}(\hat{\beta}-\beta)+\overset{*}{Z}(\hat{U}-U)][\overset{*}{X}(\hat{\beta}-\beta)+\overset{*}{Z}(\hat{U}-U)]^t\right\} \overset{*}{C}\tilde{N} \\
&= \tilde{N}^t \overset{*}{C} E[\overset{*}{X}(\hat{\beta}-\beta)(\hat{\beta}-\beta)^t] \overset{*}{C}\tilde{N} \ + \\
&\quad + 2\tilde{N}^t \overset{*}{C} E[\overset{*}{X}(\hat{\beta}-\beta)(\hat{U}-U)^t \overset{*}{Z}{}^t] \overset{*}{C}\tilde{N} \ + \\
&\quad + \tilde{N}^t \overset{*}{C} E[\overset{*}{Z}(\hat{U}-U)(\hat{U}-U)^t \overset{*}{Z}{}^t] \overset{*}{C}\tilde{N} \ .
\end{aligned} \tag{G.8}
$$

Since $\overset{*}{X}$ and $\overset{*}{Z}$ are fixed, we may further simplify to

$$
\begin{aligned}
MSE_i^a &\doteq \tilde{N}^t \overset{*}{C}\overset{*}{X} \mathrm{E}[(\hat{\beta}-\beta)(\hat{\beta}-\beta)^t] \overset{*}{X}{}^t \overset{*}{C}\tilde{N} \\
&\quad + 2\tilde{N}^t \overset{*}{C}\overset{*}{X} \mathrm{E}[(\hat{\beta}-\beta)(\hat{U}-U)^t] \overset{*}{Z}{}^t \overset{*}{C}\tilde{N} \ + \ \tilde{N}^t \overset{*}{C}\overset{*}{Z} \mathrm{E}[(\hat{U}-U)(\hat{U}-U)^t] \overset{*}{Z}{}^t \overset{*}{C}\tilde{N}
\end{aligned} \tag{G.9}
$$

At this point, it is important to note that we estimated $\hat{U}_{ij}=0$ for every PSU j which was not in sample. We could have made random draws for these variables from the normal distribution with mean zero and variance $\hat{\sigma}_2^2$, but we saw no reason to introduce extra noise into the system. Random draws might have been an appropriate step if we had been interested in the distribution of propensities across all PSUs, but this was not a primary goal of our work. Rather, our primary goal was to minimize the mean square error of the predictions for the states and large MSAs. For this goal, we think that estimating a zero random effect for all nonsample counties was the best course of action. Our initial MSE estimates ignored the fact that these zero predictions have a nonzero error associated with them. We show both the original formula for estimating mean square error and the revised formula.

Let U_S denote the vector of random effects for the states and the sample PSUs. Let $U_{\$}$ denote the corresponding vector for the nonsample PSUs in states of interest. Let $Z_{\$}$ be a matrix with as many rows as $\overset{*}{Z}$ and as many columns as there are nonsample PSUs in the states of interest, where each row of $Z_{\$}$ is mostly zeroes. A one in the j-th column of the i-th row of $Z_{\$}$ indicates that the i-th domain and block group combination is in the j-th nonsample PSU. Let Z_S denote the columns of $\overset{*}{Z}$ corresponding to the random effects for states and sample PSUs. Then we may write

$$\lambda = X\beta + \begin{bmatrix} Z_S & Z_\$ \end{bmatrix} \begin{bmatrix} U_S \\ U_\$ \end{bmatrix} \qquad \textbf{(G.10)}$$

and

$$\hat{\pi} - \pi \doteq C\left[X(\hat{\beta}-\beta) + Z_S(\hat{U}_S - U_S) - Z_\$ U_\$\right] \qquad \textbf{(G.11)}$$

Since all the true random effects (those for states, sample PSUs and nonsample PSUs) are assumed to be mutually independent and since the estimates of the model parameters are independent of the true random effects for areas not in sample, the mean square error can be rewritten as

$$\begin{aligned} MSE_i^a \doteq\ & \tilde{N}^t \overset{*}{C} \overset{*}{X} \mathrm{E}[(\hat{\beta}-\beta)(\hat{\beta}-\beta)^t] \overset{*}{X}{}^t \overset{*}{C} \tilde{N} \\ & + 2\tilde{N}^t \overset{*}{C} \overset{*}{X} \mathrm{E}[(\hat{\beta}-\beta)(\hat{U}_S - U_S)^t] \overset{*}{Z}{}^t_S \overset{*}{C} \tilde{N} + \\ & + \tilde{N}^t \overset{*}{C} \overset{*}{Z}_S \mathrm{E}[(\hat{U}_S - U_S)(\hat{U}_S - U_S)^t] \overset{*}{Z}{}^t_S \overset{*}{C} \tilde{N} + \\ & + \tilde{N}^t \overset{*}{C} Z_\$ \mathrm{E}(U_\$ U_\$^t) Z_\$^t \overset{*}{C} \tilde{N} \end{aligned} \qquad \textbf{(G.12)}$$

The fourth term is the term that was omitted from the original MSE calculations.

We now examine each of the terms in **(G.12)** in turn. Starting with the β-term, we have that

$$\beta = (X^t V^{-1} X)^{-1} (X^t V^{-1} X) \beta$$

and

$$\hat{\beta} = (X^t V^{-1} X)^{-1} X^t V^{-1} \zeta$$

So

$$\hat{\beta} - \beta = (X^t V^{-1} X)^{-1} X^t V^{-1} (\zeta - X\beta) \qquad \textbf{(G.13)}$$

We demonstrated in Section C.1, that the covariance matrix for ζ is V. Thus, if we treated V as fixed rather than as a parameter to be estimated, we have that

$$
\begin{aligned}
\mathrm{E}\left[(\hat{\beta}-\beta)(\hat{\beta}-\beta)^t\right] &\doteq (X^tV^{-1}X)^{-1}X^tV^{-1}\ \mathrm{E}\left[(\zeta-X\beta)(\zeta-X\beta)^t\right]\ V^{-1}X(X^tV^{-1}X)^{-1} \\
&= (X^tV^{-1}X)^{-1}X^tV^{-1}\ VV^{-1}X(X^tV^{-1}X)^{-1} \\
&= (X^tV^{-1}X)^{-1}(X^tV^{-1}X)(X^tV^{-1}X)^{-1} \\
&= (X^tV^{-1}X)^{-1}
\end{aligned}
\qquad \textbf{(G.14)}
$$

Temporarily skipping the second term of **(G.12)** and looking instead at the third term of **(G.12)**, we have from Section C.1 that

$$\hat{U} = GZ^tP\zeta$$

where we have dropped the S subscript since it is clear that we are only talking about the random effects for states and sample PSUs in this equation. Using Lemma C.1 that $PX=0$, we can rewrite this as

$$\hat{U} = GZ^tP(\zeta-X\beta)$$

By the definition of ζ and ε from Section C.1, we have then that

$$\hat{U} = GZ^tP(ZU+W^{-1}\varepsilon)$$

Thus,

$$
\begin{aligned}
\mathrm{E}(\hat{U}-U)(\hat{U}-U)^t &= \mathrm{E}\left\{[GZ^tP(ZU+W^{-1}\varepsilon)-U][U^tZ^t+\varepsilon^tW^{-1}PZG-U^t]\right\} \\
&= \mathrm{E}(GZ^tPZUU^tZ^tPZG) + \mathrm{E}(GZ^tPW^{-1}\varepsilon U^tZ^tPZG) + \\
&\quad + \mathrm{E}(GZ^tPZU\varepsilon^tW^{-1}PZG) + \mathrm{E}(GZ^tPW^{-1}\varepsilon\varepsilon^tW^{-1}PZG) + \\
&\quad - \mathrm{E}(GZ^tPZUU^t) - \mathrm{E}(UU^tZ^tZ^tPZG) \\
&\quad -\mathrm{E}(U\varepsilon^tW^{-1}PZG) + \mathrm{E}(UU^t)
\end{aligned}
\qquad \textbf{(G.15)}
$$

We now ignore the fact that G and P are both estimated from the data and instead treat them as fixed. We can then bring the expected value operator inside each term to focus just on the variability in U and in ε. Here we note that in Section C.1, we showed that U and ε are uncorrelated, the $\mathrm{Cov}(\varepsilon)=C^{-1}$, and that $PVP=P$.

Using this simplifying assumption and the results from Section C.1, we have that

$$\begin{aligned}
E(\hat{U}-U)(\hat{U}-U)^t &\doteq GZ^tPZGZ^tPZG + 0 + 0 + GZ^tPW^{-1}C^{-1}W^{-1}PZG + \\
&\quad - GZ^tPZG - GZ^tPZG - 0 + G \\
&= GZ^tP(V-R)PZG + GZ^tPRPZG - 2GZ^TPZG + G \\
&= GZ^tPVPZG - 2GZ^tPZG + G \\
&= GZ^tPZG - 2GZ^tPZG + G \\
&= G - GZ^tPZG
\end{aligned} \qquad \textbf{(G.16)}$$

The fourth term is quite simple. By definition, $E(U_\$U_\$^t)=\sigma_2^2 I$, where I is the identity matrix of order equal to the number of nonsample PSUs in the states of interest.

Returning now to the second term of **(G.12)** and using the representations developed for the first and third terms, we have that

$$\begin{aligned}
E(\hat{\beta}-\beta)(\hat{U}-U)^t &= E\{[(X^tV^{-1}X)^{-1}X^tV^{-1}(ZU+W^{-1}\varepsilon)][U^tZ^t+\varepsilon^tW^{-1})PZG-U^t]\} \\
&= E[(X^tV^{-1}X)^{-1}X^tV^{-1}ZUU^tZ^tPZG] + E[(X^tV^{-1}X)^{-1}X^tV^{-1}ZU\varepsilon^tW^{-1}PZG] + \\
&\quad + E[(X^tV^{-1}X)^{-1}X^tV^{-1}W^{-1}\varepsilon U^tZ^tPZG] + E[(X^tV^{-1}X)^{-1}X^tV^{-1}W^{-1}\varepsilon\varepsilon^tW^{-1}PZG] + \\
&\quad - E[(X^tV^{-1}X)^{-1}X^tV^{-1}ZUU^t] - E[(X^tV^{-1}X)^{-1}X^tV^{-1}W^{-1}\varepsilon U^t]
\end{aligned} \qquad \textbf{(G.17)}$$

Again, we ignore the stochastic nature of V, P and G, and write

$$\begin{aligned}
E(\hat{\beta}-\beta)(\hat{U}-U)^t &\doteq (X^tV^{-1}X)^{-1}X^tV^{-1}ZGZ^tPZG + 0 + 0\,(X^tV^{-1}X)^{-1}X^tV^{-1}W^{-1}C^{-1}W^{-1}PZG + \\
&\quad - (X^tV^{-1}X)^{-1}X^tV^{-1}\,ZG - 0 \\
&= (X^tV^{-1}X)^{-1}X^tV^{-1}(V-R)PZG + (X^tV^{-1}X)^{-1}X^tV^{-1}RPZG - (X^tV^{-1}X)^{-1}X^tV^{-1}ZG \\
&= (X^tV^{-1}X)^{-1}X^tPZG - (X^tV^{-1}X)^{-1}X^tV^{-1}ZG \\
&= (X^tV^{-1}X)^{-1}X^t(P-V^{-1})ZG \\
&= -(X^tV^{-1}X)^{-1}X^tV^{-1}X(X^tV^{-1}X)^{-1}X^tV^{-1}ZG \\
&= -(X^tV^{-1}X)^{-1}X^tV^{-1}ZG
\end{aligned} \qquad \textbf{(G.18)}$$

Substituting the results from **(G.14)**, **(G.15)**, and **(G.16)** back into **(G.12)**, we have that

$$MSE_i^a \doteq \tilde{N}^t \overset{*}{C} \overset{*}{X} (X^t V^{-1} X)^{-1} \overset{*}{X}^t \overset{*}{C} \tilde{N} - 2\tilde{N}^t \overset{*}{C} \overset{*}{X} (X^t V^{-1} X)^{-1} X^t V^{-1} Z G \overset{*}{Z}^t \overset{*}{C} \tilde{N} + \tilde{N}^t \overset{*}{C} \overset{*}{Z} (G - G Z^t P Z G) \overset{*}{Z}^t \overset{*}{C} \tilde{N} + \sigma_2^2 \tilde{N}^t \overset{*}{C} Z_\$ Z_\$^t \overset{*}{C} \tilde{N} \tag{G.19}$$

We now make a final approximation which is to substitute estimated parameters for unknown parameters. This substitution should be accurate asymptotically.

$$MSE_i^a \doteq \tilde{N}^t \hat{\overset{*}{C}} \overset{*}{X} (X^t \hat{V}^{-1} X)^{-1} \overset{*}{X}^t \hat{\overset{*}{C}} \tilde{N} - 2\tilde{N}^t \hat{\overset{*}{C}} \overset{*}{X} (X^t \hat{V}^{-1} X)^{-1} X^t \hat{V}^{-1} Z \hat{G} \overset{*}{Z}^t \hat{\overset{*}{C}} \tilde{N} + + \tilde{N}^t \hat{\overset{*}{C}} \overset{*}{Z} (\hat{G} - \hat{G} Z^t \hat{P} Z \hat{G}) \overset{*}{Z}^t \hat{\overset{*}{C}} \tilde{N} + \hat{\sigma}_2^2 \tilde{N}^t \hat{\overset{*}{C}} Z_\$ Z_\$^t \hat{\overset{*}{C}} \tilde{N} \tag{G.20}$$

The original report on this project also gave a computationally fast form of this mean square error estimator. The problems with calculating the forms given here are that V is too large to practically invert and that Z is mostly full of zeroes. By using lemma C.13 and by writing out the summations implicit in the matrix equations given here, the calculations become much more practical. The derivation of that form is, however, extremely tedious and doesn't add very much to the intuitive understanding of the mean square error estimation and was thus dropped from this report. If someone, however, is trying to replicate the results, it would be useful to have that original version of the report, keeping in mind that the fourth term was inadvertently omitted in that derivation.

G.2 Asymptotic Theory

As $n \to \infty$, we expect that most of the approximations employed to estimate $E(\hat{\pi}_i^W - \pi_i)^2$ will tend to become more and more accurate. Recall that there were five major approximations:

(1) Assume that $E\left[(\hat{\pi}_{ib}^{ad} - \pi_{ib}^{ad})(\hat{\pi}_{ib}^{a'd} - \pi_{ib}^{a'd})\right] =$ for $a \neq a'$

(2) Assume that $E\left[(\hat{\pi}-\pi)(\hat{\pi}-\pi)'\right] \doteq E\{[CX(\hat{\beta}-\beta)+CZ(\hat{U}-U)][CZ(\hat{\beta}-\beta)+[CX(\hat{\beta}-\beta)+CZ(\hat{U}-U)]'\}$

(3) Assume that $\mathrm{Cov}(\hat{C}) = 0$

(4) Assume that $\mathrm{Cov}(\hat{G}) = 0$

(5) Assume that C, G, P and V may be replaced by $\hat{C}, \hat{G}, \hat{P}$ and $\hat{V}$ in the final expression.

All of these assumptions should be asymptotically correct with the exception of the first. The third assumption is perhaps the least troubling since there is a fairly large set of block groups represented in the sample. The fourth assumption, on the other hand, is the most troubling since the numbers of states and PSUs represented in the sample are such smaller.

G.3 Unrepresented Sources of Error

As mentioned in Sections 2.6 and **G.2**, the unrepresented component of most concern is the component due to estimation of G, the matrix of the between-state and between-PSU residual variances. This is a limitation of the empirical Bayes method that had been of concern since the earliest work using the method (e.g., Harville, 1977). Prasad and Rao (1990) developed an approximate method for taking the uncertainty about G into effect when estimating $\mathrm{E}\left[(\hat{U}-U)(\hat{U}-U)^t\right]$, but their method has only been demonstrated to work well for linear mixed models where G does not depend on π and where a strictly unbiased estimator for G is available. We worked out a rough application of their correction to MSE estimates when there was only one level of random effects. This rough correction had almost no impact on the MSEs for the oversampled city estimates, but did inflate the square root of the MSEs on the smaller states and on the balances of the states with oversampled cites by 2 to 20 percent. Since these inflations appeared to be unstable, we decided not to incorporate them into the official MSE estimates.

Leaving out the correlations between age groups does not have any effect on the MSEs for age group specific estimates, but does have an effect on area totals and on estimates for domains that cut across ages such as estimates by sex or race, not crossed by age. We could have incorporated these correlations with larger and faster computers.

The final source of error that was neglected was that due to the clustering of the sample by segment and by household. Other aspects of the sample design such as clustering at the county/MSA level and differential weighting were fully reflected in the estimated MSEs. We had originally intended on fitting models with additional random effects for segments and households, but the data became too thin. With a rare characteristic, the segment and household level residuals tend to be small and negative but are occasionally large and positive. This causes the empirical Bayesian procedure to become biased and unstable. This problem might be partially rectified by going to a full Bayesian approach, but even with such an approach, estimates of variance components at the segment and household levels are likely to remain unstable for rare characteristics. Another possible solution might be to use a leaner X matrix.

In the course of evaluating the effect of initially omitting the fourth term of the MSE, we noticed another possible source of underestimation of the MSE. We noted that on theoretical grounds, one would expect σ_2^2 to be larger (perhaps considerably larger) than σ_1^2. This is because one expects the heterogeneity in the logit propensity to be greater among PSUs than among states. We, in fact, often observed the opposite pattern in our estimated components of variance. Furthermore, we often estimated σ_2^2 to be zero. We theorize that our estimated values of σ_2^2 are too small due to overfitting of the fixed part of the model. If we had not allowed so many fixed predictors into the model, our estimates of σ_2^2 would certainly have been larger and that would have increased the importance of the fourth component of the MSE.

G.4 Derivation of Goodness of Fit Statistic

As discussed in Section 3.1.3, our approach was to form a Wald statistic along the lines of

$$T=(\underset{\sim}{\hat{\pi}}^W-\underset{\sim}{\hat{\pi}}^D)^t\,\hat{\Psi}^{-1}\,(\underset{\sim}{\hat{\pi}}^W-\underset{\sim}{\hat{\pi}}^D)$$

where the squiggly line under a symbol indicates the vector of estimated propensities for the L homogenous groups described in that section and

$$\hat{\Psi}=\mathrm{Cov}(\underset{\sim}{\hat{\pi}}^W-\underset{\sim}{\hat{\pi}}^D)=\mathrm{E}(\underset{\sim}{\hat{\pi}}^W-\underset{\sim}{\hat{\pi}}^D)^t(\underset{\sim}{\hat{\pi}}^W-\underset{\sim}{\hat{\pi}}^D).$$

Under the null hypothesis that $\mathrm{E}(\underset{\sim}{\hat{\pi}}^W-\underset{\sim}{\hat{\pi}}^D)=0$ and certain regularity assumptions, T is approximately asymptotically distributed as a chi-square variable with L degrees of freedom. The model is rejected if T is too large relative to what is expected from a chi-square variable. The problem that we faced was how to define the meaning of the expected value in null hypothesis and in the definition of the variance-covariance matrix $\hat{\Psi}$. We would have found it most satisfying to have defined the expectation as over all possible samples, over the distribution of the state and PSU random effects, and over the distribution of person level random outcomes. However, we were not able to derive a method of doing this within the time and budget available. Accordingly, we did something simpler. We conditioned on the state and PSU random effects, on the estimated β vector, on the estimated set of random effects, and on the partition induced by the fitted model. This means that the expectation we used was just over all possible samples and over the conditional distribution of the person level random events given that state and PSU random effects are fixed **and** equal to our estimated state and PSU random effects and that the true β vector is equal to the estimated β vector. A less conditional definition of the expectation in the equation for $\hat{\Psi}$ would have almost certainly resulted in larger variances which would, in turn, have led to less significant test results since T would have tended to be smaller. Although it would probably have been possible with more work to derive the variance not conditioning on the random effects or on the estimated parameters, it is doubtful that we could have found the variance without conditioning on the partition induced by the model. Since the partition is random, the distribution of $\underset{\sim}{\hat{\pi}}^{W}-\underset{\sim}{\hat{\pi}}^{D}$ depends on the order statistics of the entire set of person-level predicted propensities.

Conditioning on fixed and random effects, $\mathrm{E}(\underset{\sim}{\hat{\pi}}^{D}|S,\hat{\beta},\hat{U})=\underset{\sim}{\hat{\pi}}^{W}$ is a constant, so the conditional variance of $\underset{\sim}{\hat{\pi}}^{W}-\underset{\sim}{\hat{\pi}}^{D}$ given the NHSDA sample (denoted by S) and the estimated model parameters is

$$\Psi_1=\mathrm{Cov}(\underset{\sim}{\hat{\pi}}^{W}-\underset{\sim}{\hat{\pi}}^{D}|S,\hat{\beta},\hat{U})=\begin{bmatrix} \mathrm{Var}\left(\dfrac{\sum\limits_{i,j,k\in G_1} w_{ijk}y_{ijk}}{\sum\limits_{i,j,k\in G_1} w_{ijk}}|S,\hat{\beta},\hat{U}\right) & & 0 \\ & \ddots & \\ 0 & & \mathrm{Var}\left(\dfrac{\sum\limits_{i,j,k\in G_L} w_{ijk}y_{ijk}}{\sum\limits_{i,j,k\in G_L} w_{ijk}}|S,\hat{\beta},\hat{U}\right) \end{bmatrix}$$

$$=\begin{bmatrix} \dfrac{\hat{\pi}_1^W(1-\hat{\pi}_1^W)Deff_1}{n_1} & & 0 \\ & \ddots & \\ 0 & & \dfrac{\hat{\pi}_L^W(1-\hat{\pi}_L^W)Deff_L}{n_L} \end{bmatrix},$$

G_g is the set of sample cases in the g-th homogenous group, n_g is the sample size for the group, and $Deff_g$ is the design effect for the group due to unequal weighting. Standard sampling theory informs us that this design effect is equal to one plus the relative variation in the sampling weights within the group. That is

$$Deff_g = 1 + \frac{\left(\sum_{i,j,k \in G_g} w_{ijk} - \bar{w}_g \right)^2}{\bar{w}_g^2}.$$

We could have used Ψ_1 to form a Wald statistic $T_1 = (\hat{\underset{\sim}{\pi}}^W - \hat{\underset{\sim}{\pi}}^D)^t \, \Psi_1^{-1} (\hat{\underset{\sim}{\pi}}^W - \hat{\underset{\sim}{\pi}}^D)$, but we were concerned that the variance was too small due to the strongly conditional nature of the definition of Ψ_1. To remove the conditioning on the particular sample that was selected for the 1991-1993 NHSDA, we used SUDAAN. We instructed SUDAAN to get the variance covariance matrix for the L group means, where the dependent variable was defined as $W^{-1}\varepsilon$, as defined in Section 1 of Appendix C and the weight vector was defined as $WCW\underset{\sim}{1}_n$, where $\underset{\sim}{1}_n$ is just a column vector of n ones. Since $\varepsilon = C^{-1}(y - \pi)$, the weighted dependent variable as defined here is equal to $WCWW^{-1}C^{-1}(y - \pi) = W(y - \pi)$, as desired. If the NHSDA were a simple random sample and we instructed SUDAAN accordingly, SUDAAN would then estimate the variance covariance matrix of the L group means to be Ψ_1. However, we instructed SUDAAN to treat the sample as the stratified two stage design that the NHSDA really is. For first stage strata, we used pairs of noncertainty PSUs and collections of 25 area segments from each certainty PSU, including the six oversampled MSAs. For the first stage clusters, we used noncertainty PSUs and half samples of 12 to 13 area segments from each of the strata defined in the certainty PSUs. We also instructed SUDAAN to treat the PSUs as drawn with unequal probability and with replacement. As a result, the variance covariance matrix estimated by SUDAAN is Ψ_2, the variance of $\hat{\underset{\sim}{\pi}}^W - \hat{\underset{\sim}{\pi}}^D$ conditioned on the estimated parameters but not conditioned on the sample or on the person level outcomes.

We could then have calculated the Wald statistic as $T_2 = (\hat{\underset{\sim}{\pi}}^W - \hat{\underset{\sim}{\pi}}^D)^t \, \Psi_2^{-1} (\hat{\underset{\sim}{\pi}}^W - \hat{\underset{\sim}{\pi}}^D)$, as indeed, SUDAAN has an option to support. However, it is fairly well known that when the dimension of the $\hat{\underset{\sim}{\pi}}^W - \hat{\underset{\sim}{\pi}}^D$ vector is large, then the stability of the variance-covariance matrix becomes more critical for the chi-square approximation to the distribution of T to be reasonably good. Since we knew that in this case the vector did have high dimension (40) and that the stability of our estimated variance covariance matrix was limited by perhaps fewer than 40 degrees of freedom, we did not feel comfortable using this test. Instead, we used the Satterwaithe adjustment to T_1 suggested by Rao and Scott (1981).

Let $\Lambda = \Psi_1^{-1}\Psi_2$. This is often referred to as a design effect matrix. If both Ψ_1 and Ψ_2 were diagonal, then the matrix Λ would have the design effects for the L group means on its main diagonal. From linear algebra, we know that the eigenvalues of a diagonal matrix are equal to the elements on the main diagonal. Rao and Scott used these facts to suggest using the eigenvalues of Λ to adjust T_1. Let $\bar{\mu}$ be the average eigenvalue of Λ and let V_μ^2 be the relative variance of the eigenvalues. Then Rao and Scott's Satterwaithe-adjusted Wald statistic is

$$T_3 = \frac{T_1}{\bar{\mu}(1 + V_\mu^2)}$$

with adjusted degrees of freedom $L' = \dfrac{L}{1 + V_\mu^2}$.

Appendix H: Inadequacy of Traditional Measure of Design-Based Mean Square Error

Appendix H. Inadequacy of Traditional Measure of Design-Based Mean Square Error

The first formal approach to evaluation of small area estimates that was proposed in the literature on the subject was to examine the mean-squared difference between the model-based estimates and the design-consistent estimates across the small areas (Gonzalez, 1973). This approach has some appeal for estimators based on fixed-effect models that do not have area-specific effects, but it does not apply to composite estimators or estimators based on mixed effect models. The approach is based on the following error decomposition, where E_D indicates expectation with respect to the sample design:

$$
\begin{aligned}
E_D\left[\frac{1}{L}\sum_i\left(\pi_i^F-\pi_i^D\right)^2\right] &= E_D\left[\frac{1}{L}\sum_i\left(\pi_i^F-E_D\pi_i^F+E_D\pi_i^F-E_D\pi_i^D+E_D\pi_i^D-\pi_i^D\right)^2\right] \\
&= E_D\left[\frac{1}{L}\sum_i\left(\pi_i^F-E_D\pi_i^F\right)^2\right] + E_D\left[\frac{1}{L}\sum_i\left(E_D\pi_i^F-E_D\pi_i^D\right)^2\right] + \\
&\quad + E_D\left[\frac{1}{L}\sum_i\left(\pi_i^D-E_D\pi_i^D\right)^2\right] - E_D\left[\frac{1}{L}\sum_i\left(\pi_i^F-E_D\pi_i^F\right)\left(\pi_i^D-E_D\pi_i^D\right)\right] \\
&= \frac{1}{L}\sum_i E_D\left(\pi_i^F-E_D\pi_i^F\right)^2 + \frac{1}{L}\sum_i E_D\left(E_D\pi_i^F-E_D\pi_i^D\right)^2 + \\
&\quad + \frac{1}{L}\sum_i E_D\left(\pi_i^D-E_D\pi_i^D\right)^2 - \frac{1}{L}\sum_i E_D\left(\pi_i^F-E_D\pi_i^F\right)\left(\pi_i^D-E_D\pi_i^D\right) \\
&= \frac{1}{L}\sum_i \mathrm{Var}_D\left(\pi_i^F\right) + \frac{1}{L}\sum_i \mathrm{Bias}_D\left(\pi_i^F\right) + \frac{1}{L}\sum_i \mathrm{Var}_D\left(\pi_i^D\right) - \frac{1}{L}\sum_i \mathrm{Cov}_D\left(\pi_i^F,\pi_i^D\right)
\end{aligned}
$$

By definition, the design-based mean square error of π_i^F is

$$\mathrm{MSE}_D\left(\pi_i^F\right) = \mathrm{Var}_D\left(\pi_i^F\right)+\mathrm{Bias}_D\left(\pi_i^F\right).$$

Often when unbiased estimators are impractical, one tries to obtain an estimator with minimum mean square error (MSE). To estimate the MSE separately for each small area is too unstable. By averaging across the small areas, the stability of this evaluation measure is improved. The sum of the first two terms on the right is the average of the design-based mean square error of π_i^F across all the small areas. It is straightforward to estimate the two variance terms using survey methods. If the covariance term can be assumed to be zero, then the design-based variance of π_i^D can be subtracted off the observed mean squared deviation between the two estimators across the areas so as to leave an unbiased estimate of the average mean square error of π_i^F. As mentioned above, when a fixed effect model is used that does not contain any area-specific effects, it is reasonable to assume that π_i^F and π_i^D are independent and thus that their covariance is zero. However, when random effect models are used and the estimator is π_i^M, the covariance between π_i^M and π_i^D can become quite strong, thereby invalidating this approach. This is most easily seen by recalling that small area estimators based upon mixed effect models are closely related to composite estimators. Since $\pi_i^C = \Gamma_i \pi_i^F + (1-\Gamma_i)\pi_i^D$, it is obvious that by choosing $\Gamma_i = 0$, we would achieve

$$E_D\left[\frac{1}{L}\sum_i \left(\pi_i^C - \pi_i^D\right)^2\right] = 0.$$

If one were to naively assume that the covariance between π_i^C and π_i^D were zero and then used this approach to estimate the mean square error of π_i^C, one would estimate negative mean square error. This approach is thus not appropriate for evaluating small area estimates based on compositing or upon mixed effect models. Accordingly, the approach was not used. The various estimates were compared across the small domains of interest only to see whether the expected shrinkage patterns were realized.

Appendix I: Estimates and Confidence Intervals for Estimates

Appendix I. Estimated Percentage of Population with Substance Abuse Behaviors for 26 States. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals

State	Licit Substance Use In Past Month				Illicit Substance Use In Past Month					
	Cigarette Use		Alcohol Use		Any Illicit Drug Use		Any Illicit But Mrj Use		Cocaine Use	
Total United States	**25.46**	(25.29-25.64)	**49.92**	(49.32-50.51)	**5.83**	(5.69-5.98)	**2.44**	(2.37-2.52)	**0.80**	(0.75-0.85)
North East Region										
New Jersey	**24.54**	(23.72-25.37)	**56.36**	(54.43-58.26)	**5.40**	(4.85-6.02)	**1.89**	(1.64-2.19)	**1.01**	(0.80-1.26)
New York	**23.73**	(23.27-24.19)	**53.37**	(51.71-55.02)	**6.44**	(6.01-6.89)	**2.37**	(2.19-2.56)	**0.82**	(0.74-0.91)
Pennsylvania	**26.87**	(26.26-27.49)	**52.32**	(50.11-54.53)	**5.07**	(4.60-5.59)	**2.04**	(1.81-2.30)	**0.59**	(0.48-0.72)
South Region										
Florida	**24.36**	(23.82-24.90)	**45.84**	(43.89-47.81)	**4.95**	(4.52-5.41)	**1.99**	(1.78-2.23)	**0.69**	(0.58-0.83)
Georgia	**26.34**	(25.43-27.27)	**45.10**	(42.39-47.83)	**5.82**	(5.02-6.73)	**2.04**	(1.72-2.41)	**0.59**	(0.42-0.83)
Kentucky	**31.44**	(30.48-32.41)	**38.45**	(35.44-41.55)	**4.53**	(3.93-5.21)	**1.66**	(1.38-2.00)	**0.44**	(0.32-0.61)
Louisiana	**25.73**	(24.67-26.83)	**45.79**	(43.19-48.42)	**4.38**	(3.78-5.08)	**2.01**	(1.68-2.40)	**0.81**	(0.59-1.12)
North Carolina	**26.60**	(25.91-27.31)	**43.76**	(41.14-46.42)	**5.85**	(5.27-6.50)	**1.79**	(1.55-2.07)	**0.59**	(0.46-0.75)
Oklahoma	**26.93**	(25.92-27.96)	**37.52**	(34.32-40.84)	**6.96**	(5.99-8.06)	**4.15**	(3.69-4.66)	**0.56**	(0.42-0.76)
South Carolina	**28.95**	(27.70-30.24)	**43.56**	(40.04-47.15)	**4.99**	(4.15-5.99)	**1.66**	(1.35-2.05)	**0.66**	(0.44-0.98)
Tennessee	**29.51**	(28.35-30.70)	**38.20**	(34.90-41.62)	**4.53**	(3.80-5.39)	**1.91**	(1.60-2.28)	**0.96**	(0.71-1.30)
Texas	**26.19**	(25.47-26.92)	**48.90**	(47.06-50.74)	**5.57**	(5.05-6.15)	**2.24**	(1.97-2.54)	**0.79**	(0.63-0.98)
Virginia	**25.01**	(24.29-25.73)	**48.09**	(45.74-50.45)	**5.55**	(4.95-6.23)	**2.71**	(2.46-2.98)	**1.25**	(1.00-1.57)
West Virginia	**30.63**	(29.48-31.79)	**36.40**	(32.84-40.12)	**4.23**	(3.61-4.95)	**1.41**	(1.15-1.73)	**0.56**	(0.41-0.75)
North Central Region										
Illinois	**25.96**	(25.40-26.52)	**50.64**	(48.99-52.28)	**4.64**	(4.36-4.94)	**1.94**	(1.78-2.10)	**0.69**	(0.61-0.78)
Indiana	**24.52**	(23.71-25.34)	**44.50**	(41.37-47.68)	**4.49**	(3.95-5.11)	**1.85**	(1.56-2.19)	**0.50**	(0.38-0.67)
Kansas	**24.15**	(23.07-25.26)	**52.56**	(49.35-55.76)	**5.01**	(4.26-5.89)	**3.19**	(2.79-3.63)	**0.77**	(0.58-1.02)
Michigan	**26.95**	(26.12-27.79)	**52.23**	(49.38-55.07)	**5.52**	(4.89-6.23)	**1.55**	(1.29-1.85)	**0.61**	(0.48-0.79)
Minnesota	**22.57**	(21.70-23.47)	**58.68**	(55.77-61.54)	**4.62**	(4.02-5.31)	**1.38**	(1.14-1.68)	**0.32**	(0.23-0.45)
Missouri	**25.22**	(24.44-26.01)	**50.52**	(47.85-53.19)	**4.98**	(4.37-5.67)	**2.82**	(2.50-3.19)	**0.89**	(0.71-1.11)
Ohio	**29.19**	(28.53-29.86)	**48.87**	(46.49-51.25)	**5.37**	(4.87-5.92)	**2.52**	(2.27-2.80)	**0.78**	(0.64-0.94)
Wisconsin	**23.51**	(22.50-24.54)	**55.07**	(51.48-58.61)	**4.07**	(3.49-4.73)	**1.93**	(1.64-2.27)	**0.84**	(0.64-1.11)
West Region										
California	**22.84**	(22.46-23.22)	**52.90**	(52.15-53.66)	**8.23**	(7.86-8.62)	**3.91**	(3.69-4.14)	**0.97**	(0.86-1.10)
New Mexico	**28.42**	(26.73-30.16)	**49.82**	(46.71-52.92)	**7.76**	(6.44-9.33)	**3.54**	(2.90-4.32)	**0.85**	(0.57-1.27)
Oregon	**25.53**	(24.61-26.47)	**52.29**	(48.87-55.69)	**7.08**	(6.28-7.99)	**3.31**	(2.89-3.78)	**0.38**	(0.28-0.51)
Washington	**23.73**	(22.80-24.67)	**54.48**	(51.64-57.29)	**6.07**	(5.32-6.91)	**3.20**	(2.74-3.74)	**0.46**	(0.34-0.63)

Appendix I. Estimated Percentage of Population with Substance Abuse Behaviors for 26 States. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals (Continued)

State	Past Year Dependency				Past Year Treatment						Past Year Arrest	
	Dependent On Illicit Drugs		Dependent On Alcohol		Received Treatment For Illicit Drug Use		Received Treatment For Alcohol Use		Needed Treatment For Illicit Drug Use			
Total United States	**1.24**	(1.19-1.29)	**3.08**	(2.97-3.20)	**0.70**	(0.66-0.74)	**0.69**	(0.65-0.74)	**2.85**	(2.76-2.95)	**1.79**	(1.72-1.86)
North East Region												
New Jersey	**1.26**	(1.11-1.43)	**2.48**	(2.14-2.88)	**0.68**	(0.56-0.83)	**0.64**	(0.49-0.83)	**2.03**	(1.77-2.34)	**1.35**	(1.16-1.56)
New York	**1.09**	(1.00-1.18)	**2.02**	(1.82-2.24)	**0.64**	(0.55-0.75)	**0.41**	(0.33-0.51)	**2.46**	(2.28-2.66)	**1.32**	(1.21-1.43)
Pennsylvania	**0.88**	(0.78-0.99)	**2.76**	(2.41-3.16)	**0.56**	(0.48-0.65)	**0.69**	(0.56-0.85)	**2.19**	(1.95-2.45)	**1.28**	(1.16-1.42)
South Region												
Florida	**0.93**	(0.84-1.04)	**2.36**	(2.07-2.69)	**0.69**	(0.60-0.80)	**0.58**	(0.47-0.70)	**2.46**	(2.21-2.74)	**1.61**	(1.47-1.75)
Georgia	**1.21**	(1.06-1.38)	**3.61**	(3.01-4.32)	**0.71**	(0.58-0.87)	**0.45**	(0.33-0.62)	**3.78**	(3.23-4.42)	**2.49**	(2.25-2.76)
Kentucky	**0.90**	(0.79-1.02)	**2.06**	(1.72-2.45)	**0.57**	(0.47-0.67)	**0.48**	(0.35-0.65)	**2.27**	(1.93-2.66)	**1.67**	(1.48-1.88)
Louisiana	**0.99**	(0.86-1.14)	**3.93**	(3.31-4.67)	**0.64**	(0.53-0.77)	**0.59**	(0.42-0.82)	**2.77**	(2.29-3.36)	**2.32**	(2.06-2.62)
North Carolina	**1.14**	(1.02-1.27)	**2.55**	(2.21-2.95)	**0.64**	(0.53-0.76)	**0.50**	(0.39-0.64)	**2.40**	(2.11-2.74)	**1.68**	(1.53-1.85)
Oklahoma	**1.50**	(1.34-1.69)	**4.05**	(3.32-4.93)	**0.86**	(0.71-1.04)	**1.49**	(0.97-2.26)	**3.77**	(3.07-4.63)	**1.37**	(1.19-1.58)
South Carolina	**1.18**	(1.02-1.36)	**3.01**	(2.41-3.77)	**0.53**	(0.42-0.66)	**0.55**	(0.38-0.81)	**2.14**	(1.72-2.66)	**1.80**	(1.55-2.08)
Tennessee	**0.91**	(0.80-1.04)	**2.06**	(1.69-2.51)	**0.60**	(0.50-0.70)	**0.53**	(0.38-0.76)	**2.08**	(1.72-2.52)	**2.10**	(1.83-2.41)
Texas	**1.47**	(1.33-1.63)	**3.37**	(2.95-3.84)	**0.61**	(0.51-0.73)	**0.69**	(0.55-0.87)	**3.06**	(2.73-3.43)	**1.84**	(1.66-2.05)
Virginia	**1.11**	(1.01-1.23)	**3.03**	(2.63-3.48)	**0.65**	(0.55-0.76)	**0.61**	(0.50-0.75)	**2.91**	(2.54-3.33)	**1.54**	(1.39-1.69)
West Virginia	**0.84**	(0.73-0.97)	**2.12**	(1.73-2.59)	**0.49**	(0.39-0.61)	**0.73**	(0.52-1.02)	**2.15**	(1.77-2.62)	**1.29**	(1.10-1.50)
North Central Region												
Illinois	**0.88**	(0.81-0.94)	**3.08**	(2.83-3.36)	**0.54**	(0.48-0.59)	**0.66**	(0.56-0.78)	**2.32**	(2.15-2.51)	**1.72**	(1.57-1.90)
Indiana	**0.97**	(0.87-1.09)	**2.15**	(1.80-2.55)	**0.52**	(0.43-0.62)	**0.66**	(0.49-0.89)	**1.98**	(1.70-2.30)	**2.30**	(2.08-2.55)
Kansas	**1.13**	(1.00-1.27)	**3.51**	(2.88-4.26)	**0.66**	(0.55-0.79)	**0.86**	(0.59-1.24)	**2.73**	(2.25-3.31)	**2.40**	(2.10-2.74)
Michigan	**1.08**	(0.96-1.22)	**3.36**	(2.82-3.99)	**0.76**	(0.65-0.89)	**0.70**	(0.52-0.93)	**3.05**	(2.63-3.53)	**1.99**	(1.79-2.21)
Minnesota	**1.00**	(0.87-1.14)	**3.12**	(2.58-3.77)	**0.84**	(0.70-1.02)	**0.80**	(0.59-1.10)	**2.19**	(1.84-2.60)	**1.67**	(1.45-1.91)
Missouri	**1.03**	(0.92-1.15)	**2.16**	(1.83-2.55)	**0.70**	(0.59-0.84)	**0.68**	(0.50-0.91)	**3.26**	(2.82-3.76)	**2.10**	(1.85-2.38)
Ohio	**0.96**	(0.87-1.07)	**2.73**	(2.36-3.15)	**0.70**	(0.59-0.83)	**0.81**	(0.64-1.01)	**2.56**	(2.28-2.87)	**2.25**	(2.06-2.47)
Wisconsin	**0.98**	(0.88-1.11)	**2.95**	(2.41-3.60)	**0.61**	(0.51-0.73)	**0.59**	(0.44-0.79)	**2.41**	(1.99-2.92)	**1.53**	(1.30-1.80)
West Region												
California	**1.91**	(1.78-2.04)	**4.87**	(4.62-5.14)	**0.97**	(0.87-1.08)	**0.81**	(0.71-0.94)	**4.23**	(3.99-4.48)	**2.15**	(1.99-2.33)
New Mexico	**1.62**	(1.41-1.87)	**3.66**	(2.93-4.57)	**0.77**	(0.61-0.96)	**0.92**	(0.55-1.56)	**3.30**	(2.61-4.17)	**2.63**	(2.07-3.32)
Oregon	**1.99**	(1.79-2.22)	**2.75**	(2.32-3.25)	**0.90**	(0.75-1.08)	**0.61**	(0.43-0.85)	**2.22**	(1.90-2.58)	**1.88**	(1.64-2.15)
Washington	**1.96**	(1.74-2.19)	**3.51**	(2.94-4.19)	**0.85**	(0.72-1.00)	**0.51**	(0.36-0.71)	**3.55**	(3.01-4.18)	**1.91**	(1.61-2.26)

Appendix I. Estimated Percentage of Population with Substance Abuse Behaviors for 25 MSAs. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals

MSA	Licit Substance Use In Past Month				Illicit Substance Use In Past Month					
	Cigarette Use		Alcohol Use		Any Illicit Drug Use		Any Illicit But Mrj Use		Cocaine Use	
Anaheim-Santa Ana, CA	**21.01**	(20.19-21.84)	**52.25**	(50.58-53.91)	**8.82**	(7.78-9.98)	**4.85**	(4.25-5.53)	**1.23**	(0.93-1.61)
Atlanta, GA	**23.70**	(22.80-24.63)	**50.98**	(49.52-52.45)	**5.86**	(5.15-6.66)	**1.95**	(1.62-2.36)	**0.56**	(0.40-0.79)
Baltimore, MD	**24.96**	(24.00-25.95)	**44.84**	(43.48-46.20)	**5.03**	(4.37-5.80)	**1.80**	(1.48-2.19)	**0.70**	(0.53-0.93)
Boston, MA	**27.51**	(26.70-28.33)	**61.51**	(60.37-62.63)	**6.73**	(5.98-7.55)	**2.57**	(2.17-3.03)	**0.94**	(0.74-1.20)
Chicago, IL	**24.73**	(24.42-25.04)	**53.82**	(53.49-54.16)	**5.53**	(5.34-5.72)	**2.34**	(2.23-2.45)	**0.94**	(0.87-1.02)
Dallas, TX	**25.95**	(25.06-26.85)	**50.23**	(48.39-52.08)	**5.69**	(4.95-6.53)	**2.38**	(2.04-2.78)	**1.05**	(0.80-1.37)
Denver, CO	**27.13**	(26.83-27.44)	**58.42**	(58.09-58.75)	**8.30**	(8.07-8.54)	**2.94**	(2.80-3.07)	**1.01**	(0.93-1.08)
Detroit, MI	**26.52**	(25.61-27.46)	**54.17**	(52.79-55.54)	**5.50**	(4.83-6.25)	**1.38**	(1.12-1.71)	**0.60**	(0.45-0.81)
El Paso, TX	**20.41**	(18.86-22.05)	**45.85**	(42.18-49.56)	**3.58**	(2.62-4.89)	**1.85**	(1.35-2.51)	**0.53**	(0.33-0.85)
Houston, TX	**25.07**	(24.03-26.13)	**50.94**	(49.37-52.50)	**4.05**	(3.47-4.71)	**2.43**	(2.07-2.86)	**0.60**	(0.44-0.82)
Los Angeles, CA	**21.82**	(21.54-22.11)	**49.34**	(49.01-49.68)	**6.68**	(6.48-6.89)	**2.72**	(2.60-2.84)	**0.90**	(0.84-0.97)
Miami-Hialeah, FL	**20.88**	(20.60-21.17)	**44.40**	(44.05-44.75)	**3.75**	(3.59-3.92)	**2.26**	(2.15-2.37)	**0.89**	(0.82-0.96)
Minneapolis-St. Paul, MN	**22.64**	(21.73-23.57)	**65.87**	(64.62-67.10)	**5.19**	(4.54-5.92)	**1.51**	(1.22-1.87)	**0.36**	(0.26-0.50)
Nassau-Suffolk, NY	**20.22**	(19.38-21.08)	**59.60**	(58.36-60.84)	**6.57**	(5.84-7.38)	**1.99**	(1.65-2.41)	**0.37**	(0.28-0.49)
New York, NY	**23.15**	(22.85-23.45)	**48.84**	(48.50-49.19)	**5.99**	(5.80-6.19)	**2.40**	(2.29-2.51)	**1.06**	(0.99-1.15)
Newark, NJ	**24.10**	(23.10-25.13)	**62.17**	(60.45-63.85)	**6.20**	(5.39-7.11)	**1.91**	(1.60-2.28)	**1.37**	(0.98-1.92)
Oakland, CA	**23.33**	(22.61-24.07)	**65.02**	(63.55-66.46)	**11.39**	(10.31-12.57)	**5.11**	(4.59-5.69)	**1.47**	(1.17-1.85)
Philadelphia, PA-NJ	**25.89**	(25.20-26.59)	**59.13**	(58.06-60.19)	**5.72**	(5.11-6.40)	**2.22**	(1.94-2.55)	**0.64**	(0.51-0.81)
Phoenix, AZ	**24.83**	(23.85-25.84)	**53.69**	(52.12-55.25)	**6.82**	(5.95-7.81)	**3.99**	(3.45-4.62)	**1.01**	(0.77-1.31)
San Antonio, TX	**26.09**	(24.71-27.52)	**54.43**	(51.55-57.27)	**4.35**	(3.39-5.55)	**2.17**	(1.67-2.81)	**0.78**	(0.51-1.17)
San Bernardino, CA	**23.32**	(22.39-24.27)	**49.12**	(47.64-50.60)	**7.42**	(6.62-8.32)	**4.60**	(4.07-5.20)	**0.64**	(0.49-0.84)
San Diego, CA	**22.25**	(21.36-23.16)	**49.08**	(47.60-50.55)	**7.07**	(6.22-8.03)	**3.59**	(3.09-4.17)	**0.93**	(0.70-1.23)
St. Louis, MO-IL	**25.27**	(24.52-26.03)	**55.02**	(53.73-56.30)	**5.19**	(4.56-5.90)	**2.53**	(2.21-2.88)	**0.95**	(0.75-1.21)
Tampa-St. Petersburg, FL	**25.55**	(24.76-26.36)	**41.22**	(39.75-42.70)	**5.27**	(4.56-6.08)	**2.03**	(1.66-2.47)	**0.57**	(0.43-0.74)
Washington, DC	**22.61**	(22.40-22.81)	**53.82**	(53.59-54.05)	**5.38**	(5.26-5.51)	**2.58**	(2.50-2.67)	**1.03**	(0.98-1.08)

Appendix I. Estimated Percentage of Population with Substance Abuse Behaviors for 25 MSAs. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals (Continued)

MSA	**Past Year Dependency**				**Past Year Treatment**						Past Year Arrest	
	Dependent On Illicit Drugs		Dependent On Alcohol		Received Treatment For Illicit Drug Use		Received Treatment For Alcohol Use		Needed Treatment For Illicit Drug Use			
Anaheim-Santa Ana, CA	**2.17**	(1.88-2.51)	**6.01**	(5.37-6.72)	**1.01**	(0.77-1.31)	**0.55**	(0.39-0.78)	**4.45**	(3.91-5.06)	**1.53**	(1.20-1.94)
Atlanta, GA	**1.24**	(1.06-1.45)	**4.47**	(3.84-5.20)	**0.74**	(0.60-0.91)	**0.38**	(0.27-0.55)	**4.83**	(4.21-5.53)	**1.74**	(1.52-1.98)
Baltimore, MD	**0.89**	(0.75-1.05)	**2.15**	(1.74-2.65)	**0.58**	(0.45-0.75)	**0.45**	(0.32-0.65)	**2.28**	(1.90-2.73)	**1.30**	(1.11-1.52)
Boston, MA	**1.68**	(1.46-1.92)	**4.38**	(3.83-5.01)	**0.97**	(0.77-1.22)	**2.18**	(1.69-2.81)	**3.88**	(3.41-4.41)	**1.66**	(1.39-1.98)
Chicago, IL	**0.89**	(0.82-0.97)	**3.46**	(3.33-3.60)	**0.53**	(0.49-0.59)	**0.71**	(0.66-0.77)	**2.75**	(2.63-2.88)	**1.51**	(1.41-1.62)
Dallas, TX	**1.63**	(1.45-1.84)	**3.31**	(2.80-3.91)	**0.65**	(0.53-0.80)	**0.45**	(0.31-0.66)	**2.47**	(2.10-2.90)	**1.55**	(1.31-1.82)
Denver, CO	**1.26**	(1.16-1.37)	**4.57**	(4.41-4.73)	**0.63**	(0.57-0.70)	**0.96**	(0.89-1.04)	**3.83**	(3.68-3.99)	**2.26**	(2.12-2.40)
Detroit, MI	**1.18**	(1.02-1.37)	**3.59**	(3.11-4.14)	**0.78**	(0.66-0.92)	**0.51**	(0.38-0.69)	**3.58**	(3.15-4.06)	**1.77**	(1.53-2.04)
El Paso, TX	**0.79**	(0.57-1.08)	**3.66**	(2.59-5.13)	**0.55**	(0.36-0.85)	**0.53**	(0.28-1.03)	**2.50**	(1.92-3.25)	**1.62**	(1.20-2.19)
Houston, TX	**2.10**	(1.87-2.37)	**1.96**	(1.53-2.50)	**0.81**	(0.68-0.98)	**0.71**	(0.53-0.94)	**3.86**	(3.40-4.37)	**1.56**	(1.32-1.84)
Los Angeles, CA	**1.50**	(1.39-1.61)	**7.13**	(6.96-7.31)	**0.75**	(0.65-0.88)	**0.50**	(0.46-0.55)	**4.04**	(3.89-4.20)	**1.85**	(1.74-1.97)
Miami-Hialeah, FL	**0.74**	(0.67-0.82)	**2.04**	(1.94-2.14)	**0.56**	(0.44-0.72)	**0.48**	(0.44-0.52)	**2.16**	(2.04-2.29)	**1.22**	(1.13-1.32)
Minneapolis-St. Paul, MN	**1.15**	(0.99-1.34)	**3.45**	(2.93-4.06)	**0.93**	(0.75-1.16)	**0.93**	(0.65-1.34)	**2.38**	(2.04-2.78)	**1.77**	(1.50-2.08)
Nassau-Suffolk, NY	**1.00**	(0.85-1.17)	**1.10**	(0.83-1.44)	**0.64**	(0.49-0.83)	**0.50**	(0.36-0.68)	**1.87**	(1.56-2.24)	**1.14**	(0.93-1.40)
New York, NY	**0.94**	(0.85-1.02)	**1.77**	(1.68-1.87)	**0.66**	(0.52-0.84)	**0.26**	(0.23-0.29)	**2.50**	(2.38-2.63)	**0.79**	(0.72-0.87)
Newark, NJ	**1.43**	(1.24-1.65)	**2.50**	(2.00-3.12)	**0.89**	(0.74-1.07)	**0.77**	(0.59-1.01)	**1.90**	(1.56-2.31)	**1.40**	(1.10-1.77)
Oakland, CA	**2.98**	(2.68-3.31)	**3.88**	(3.24-4.65)	**1.37**	(1.19-1.57)	**2.07**	(1.69-2.53)	**5.22**	(4.63-5.87)	**1.94**	(1.64-2.28)
Philadelphia, PA-NJ	**0.98**	(0.85-1.13)	**2.82**	(2.45-3.25)	**0.82**	(0.71-0.95)	**0.61**	(0.47-0.78)	**2.99**	(2.67-3.33)	**1.37**	(1.20-1.57)
Phoenix, AZ	**1.55**	(1.37-1.75)	**3.61**	(3.04-4.28)	**0.65**	(0.53-0.80)	**0.67**	(0.46-0.98)	**3.14**	(2.68-3.68)	**1.48**	(1.12-1.96)
San Antonio, TX	**1.38**	(1.19-1.61)	**2.53**	(1.69-3.76)	**0.53**	(0.37-0.74)	**0.42**	(0.22-0.79)	**2.90**	(2.20-3.81)	**1.60**	(1.30-1.96)
San Bernardino, CA	**2.09**	(1.82-2.39)	**3.92**	(3.42-4.50)	**1.15**	(0.96-1.37)	**1.19**	(0.89-1.59)	**3.84**	(3.35-4.39)	**2.41**	(2.01-2.89)
San Diego, CA	**1.38**	(1.15-1.65)	**2.99**	(2.51-3.56)	**0.74**	(0.56-0.98)	**0.52**	(0.34-0.79)	**3.19**	(2.75-3.71)	**2.18**	(1.82-2.60)
St. Louis, MO-IL	**0.96**	(0.85-1.09)	**2.46**	(2.09-2.90)	**0.74**	(0.62-0.87)	**0.62**	(0.45-0.85)	**2.75**	(2.37-3.19)	**2.09**	(1.77-2.45)
Tampa-St. Petersburg, FL	**0.98**	(0.84-1.13)	**2.10**	(1.67-2.63)	**0.66**	(0.55-0.80)	**0.41**	(0.29-0.60)	**2.08**	(1.76-2.45)	**1.44**	(1.22-1.70)
Washington, DC	**1.07**	(1.02-1.13)	**3.59**	(3.49-3.69)	**0.56**	(0.51-0.60)	**0.87**	(0.83-0.92)	**2.66**	(2.58-2.75)	**1.33**	(1.27-1.40)

Appendix I. Estimated Number of Persons with Substance Abuse Behaviors for 25 MSAs. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals

Appendix I. Estimated Number of Persons with Substance Abuse Behaviors for 26 States. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals

State	Licit Substance Use In Past Month				Illicit Substance Use In Past Month					
	Cigarette Use		Alcohol Use		Any Illicit Drug Use		Any Illicit But Mrj Use		Cocaine Use	
Total United States	**52,444**	(52,084-52,805)	**102,802**	(101,582-104,022)	**12,015**	(11,712-12,326)	**5,031**	(4,878-5,189)	**1,647**	(1,552-1,747)
North East Region										
New Jersey	**1,581**	(1,529-1,635)	**3,631**	(3,507-3,754)	**348**	(312-388)	**122**	(106-141)	**65**	(51-81)
New York	**3,534**	(3,466-3,602)	**7,948**	(7,701-8,193)	**959**	(896-1,026)	**352**	(326-381)	**122**	(110-136)
Pennsylvania	**2,672**	(2,612-2,733)	**5,203**	(4,983-5,423)	**505**	(458-556)	**203**	(180-229)	**59**	(48-72)
South Region										
Florida	**2,744**	(2,683-2,805)	**5,164**	(4,944-5,386)	**557**	(509-610)	**225**	(201-252)	**78**	(65-93)
Georgia	**1,433**	(1,384-1,484)	**2,454**	(2,307-2,603)	**317**	(273-366)	**111**	(94-131)	**32**	(23-45)
Kentucky	**960**	(930-989)	**1,174**	(1,082-1,268)	**138**	(120-159)	**51**	(42-61)	**14**	(10-19)
Louisiana	**872**	(835-909)	**1,551**	(1,463-1,640)	**148**	(128-172)	**68**	(57-81)	**27**	(20-38)
North Carolina	**1,501**	(1,461-1,541)	**2,468**	(2,321-2,618)	**330**	(297-366)	**101**	(87-117)	**33**	(26-42)
Oklahoma	**690**	(664-717)	**962**	(880-1,047)	**178**	(154-207)	**106**	(95-119)	**14**	(11-19)
South Carolina	**851**	(814-889)	**1,280**	(1,177-1,386)	**147**	(122-176)	**49**	(40-60)	**19**	(13-29)
Tennessee	**1,215**	(1,167-1,264)	**1,573**	(1,437-1,713)	**187**	(157-222)	**79**	(66-94)	**39**	(29-53)
Texas	**3,601**	(3,502-3,702)	**6,724**	(6,472-6,977)	**767**	(695-845)	**308**	(271-349)	**108**	(87-135)
Virginia	**1,307**	(1,270-1,345)	**2,514**	(2,391-2,637)	**290**	(259-326)	**141**	(128-156)	**65**	(52-82)
West Virginia	**458**	(441-476)	**545**	(492-601)	**63**	(54-74)	**21**	(17-26)	**8**	(6-11)
North Central Region										
Illinois	**2,434**	(2,382-2,488)	**4,749**	(4,594-4,903)	**435**	(409-463)	**182**	(167-197)	**65**	(57-73)
Indiana	**1,123**	(1,086-1,161)	**2,039**	(1,895-2,184)	**206**	(181-234)	**85**	(72-100)	**23**	(17-31)
Kansas	**485**	(464-508)	**1,057**	(992-1,121)	**101**	(86-118)	**64**	(56-73)	**15**	(12-20)
Michigan	**2,052**	(1,989-2,116)	**3,978**	(3,760-4,194)	**420**	(372-474)	**118**	(98-141)	**47**	(36-61)
Minnesota	**807**	(776-839)	**2,098**	(1,994-2,200)	**165**	(144-190)	**49**	(41-60)	**12**	(8-16)
Missouri	**1,065**	(1,032-1,098)	**2,134**	(2,021-2,246)	**210**	(185-239)	**119**	(106-135)	**38**	(30-47)
Ohio	**2,612**	(2,552-2,672)	**4,372**	(4,159-4,585)	**481**	(436-530)	**226**	(203-251)	**70**	(58-84)
Wisconsin	**945**	(905-987)	**2,215**	(2,070-2,357)	**164**	(141-190)	**78**	(66-91)	**34**	(26-45)
West Region										
California	**5,559**	(5,466-5,652)	**12,877**	(12,693-13,061)	**2,004**	(1,913-2,099)	**951**	(897-1,007)	**237**	(210-268)
New Mexico	**341**	(321-362)	**597**	(560-635)	**93**	(77-112)	**43**	(35-52)	**10**	(7-15)
Oregon	**612**	(590-634)	**1,253**	(1,171-1,335)	**170**	(150-191)	**79**	(69-91)	**9**	(7-12)
Washington	**971**	(934-1,010)	**2,230**	(2,114-2,345)	**248**	(218-283)	**131**	(112-153)	**19**	(14-26)

Appendix I. Estimated Number of Persons with Substance Abuse Behaviors for 25 MSAs. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals

MSA	**Licit Substance Use In Past Month**				**Illicit Substance Use In Past Month**					
	Cigarette Use		Alcohol Use		Any Illicit Drug Use		Any Illicit But Mrj Use		Cocaine Use	
Anaheim-Santa Ana, CA	**419**	(403-436)	**1,043**	(1,010-1,076)	**176**	(155-199)	**97**	(85-110)	**25**	(19-32)
Atlanta, GA	**575**	(553-597)	**1,236**	(1,201-1,272)	**142**	(125-161)	**47**	(39-57)	**14**	(10-19)
Baltimore, MD	**498**	(479-518)	**895**	(868-922)	**100**	(87-116)	**36**	(29-44)	**14**	(11-18)
Boston, MA	**865**	(840-891)	**1,934**	(1,899-1,970)	**212**	(188-238)	**81**	(68-95)	**30**	(23-38)
Chicago, IL	**1,232**	(1,217-1,247)	**2,681**	(2,664-2,698)	**275**	(266-285)	**117**	(111-122)	**47**	(44-51)
Dallas, TX	**550**	(531-569)	**1,065**	(1,026-1,104)	**121**	(105-139)	**51**	(43-59)	**22**	(17-29)
Denver, CO	**365**	(361-369)	**786**	(782-791)	**112**	(109-115)	**40**	(38-41)	**14**	(13-15)
Detroit, MI	**953**	(920-987)	**1,946**	(1,897-1,995)	**197**	(173-225)	**50**	(40-61)	**22**	(16-29)
El Paso, TX	**93**	(86-101)	**209**	(192-226)	**16**	(12-22)	**8**	(6-11)	**2**	(2-4)
Houston, TX	**667**	(640-695)	**1,355**	(1,314-1,397)	**108**	(92-125)	**65**	(55-76)	**16**	(12-22)
Los Angeles, CA	**1,555**	(1,535-1,576)	**3,517**	(3,493-3,541)	**476**	(462-491)	**194**	(185-203)	**65**	(60-69)
Miami-Hialeah, FL	**334**	(330-339)	**710**	(705-716)	**60**	(57-63)	**36**	(34-38)	**14**	(13-15)
Minneapolis-St. Paul, MN	**461**	(442-480)	**1,341**	(1,315-1,366)	**106**	(92-120)	**31**	(25-38)	**7**	(5-10)
Nassau-Suffolk, NY	**440**	(422-459)	**1,298**	(1,271-1,325)	**143**	(127-161)	**43**	(36-52)	**8**	(6-11)
New York, NY	**1,640**	(1,619-1,662)	**3,461**	(3,437-3,485)	**425**	(411-439)	**170**	(163-178)	**75**	(70-81)
Newark, NJ	**361**	(346-377)	**932**	(907-958)	**93**	(81-107)	**29**	(24-34)	**21**	(15-29)
Oakland, CA	**403**	(391-416)	**1,123**	(1,098-1,148)	**197**	(178-217)	**88**	(79-98)	**25**	(20-32)
Philadelphia, PA-NJ	**1,045**	(1,017-1,073)	**2,387**	(2,344-2,430)	**231**	(206-258)	**90**	(78-103)	**26**	(21-33)
Phoenix, AZ	**440**	(422-457)	**951**	(923-978)	**121**	(105-138)	**71**	(61-82)	**18**	(14-23)
San Antonio, TX	**271**	(257-286)	**566**	(536-596)	**45**	(35-58)	**23**	(17-29)	**8**	(5-12)
San Bernardino, CA	**495**	(475-515)	**1,042**	(1,011-1,074)	**158**	(140-177)	**98**	(86-110)	**14**	(10-18)
San Diego, CA	**465**	(446-484)	**1,025**	(995-1,056)	**148**	(130-168)	**75**	(65-87)	**19**	(15-26)
St. Louis, MO-IL	**507**	(492-522)	**1,104**	(1,078-1,130)	**104**	(91-118)	**51**	(44-58)	**19**	(15-24)
Tampa-St. Petersburg, FL	**466**	(451-480)	**751**	(724-778)	**96**	(83-111)	**37**	(30-45)	**10**	(8-13)
Washington, DC	**756**	(749-763)	**1,800**	(1,793-1,808)	**180**	(176-184)	**86**	(84-89)	**34**	(33-36)

Appendix I. Estimated Number of Persons with Substance Abuse Behaviors for 25 MSAs. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals

Appendix I. Estimated Number of Persons with Substance Abuse Behaviors for 26 States. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals

State	Past Year Dependency				Past Year Treatment						Past Year Arrest	
	Dependent On Illicit Drugs		Dependent On Alcohol		Received Treatment For Illicit Drug Use		Received Treatment For Alcohol Use		Needed Treatment For Illicit Drug Use			
Total United States	**2,549**	(2,444-2,658)	**6,350**	(6,126-6,582)	**1,447**	(1,365-1,533)	**1,429**	(1,336-1,529)	**5,877**	(5,690-6,071)	**3,689**	(3,549-3,835)
North East Region												
New Jersey	**81**	(72-92)	**160**	(138-185)	**44**	(36-53)	**41**	(32-53)	**131**	(114-151)	**87**	(75-100)
New York	**162**	(149-175)	**301**	(270-334)	**95**	(82-111)	**62**	(50-76)	**367**	(340-395)	**196**	(180-213)
Pennsylvania	**87**	(78-98)	**275**	(240-315)	**55**	(47-65)	**68**	(55-85)	**217**	(194-243)	**128**	(115-142)
South Region												
Florida	**105**	(94-117)	**266**	(233-303)	**78**	(67-90)	**65**	(53-79)	**277**	(248-309)	**181**	(166-197)
Georgia	**66**	(58-75)	**197**	(164-235)	**39**	(32-47)	**25**	(18-34)	**206**	(176-240)	**136**	(123-150)
Kentucky	**28**	(24-31)	**63**	(53-75)	**17**	(14-21)	**15**	(11-20)	**69**	(59-81)	**51**	(45-57)
Louisiana	**34**	(29-39)	**133**	(112-158)	**22**	(18-26)	**20**	(14-28)	**94**	(77-114)	**79**	(70-89)
North Carolina	**64**	(58-72)	**144**	(125-166)	**36**	(30-43)	**28**	(22-36)	**136**	(119-154)	**95**	(86-105)
Oklahoma	**39**	(34-43)	**104**	(85-126)	**22**	(18-27)	**38**	(25-58)	**97**	(79-119)	**35**	(30-41)
South Carolina	**35**	(30-40)	**89**	(71-111)	**15**	(12-19)	**16**	(11-24)	**63**	(50-78)	**53**	(46-61)
Tennessee	**38**	(33-43)	**85**	(70-103)	**25**	(21-29)	**22**	(15-31)	**86**	(71-104)	**87**	(75-99)
Texas	**202**	(182-225)	**463**	(405-528)	**84**	(71-100)	**95**	(75-119)	**421**	(376-472)	**254**	(228-282)
Virginia	**58**	(53-64)	**158**	(138-182)	**34**	(29-40)	**32**	(26-39)	**152**	(133-174)	**80**	(73-89)
West Virginia	**13**	(11-15)	**32**	(26-39)	**7**	(6-9)	**11**	(8-15)	**32**	(27-39)	**19**	(17-22)
North Central Region												
Illinois	**82**	(76-89)	**289**	(265-315)	**50**	(45-55)	**62**	(52-73)	**218**	(201-235)	**162**	(147-178)
Indiana	**45**	(40-50)	**98**	(83-117)	**24**	(20-28)	**30**	(23-41)	**91**	(78-105)	**105**	(95-117)
Kansas	**23**	(20-25)	**70**	(58-86)	**13**	(11-16)	**17**	(12-25)	**55**	(45-67)	**48**	(42-55)
Michigan	**82**	(73-93)	**256**	(215-304)	**58**	(50-68)	**53**	(40-71)	**232**	(200-269)	**151**	(136-168)
Minnesota	**36**	(31-41)	**112**	(92-135)	**30**	(25-37)	**29**	(21-39)	**78**	(66-93)	**60**	(52-68)
Missouri	**43**	(39-48)	**91**	(77-108)	**30**	(25-35)	**29**	(21-39)	**138**	(119-159)	**89**	(78-100)
Ohio	**86**	(78-95)	**244**	(211-281)	**63**	(53-74)	**72**	(57-91)	**229**	(204-256)	**202**	(184-221)
Wisconsin	**40**	(35-44)	**118**	(97-145)	**25**	(21-29)	**24**	(18-32)	**97**	(80-117)	**62**	(52-72)
West Region												
California	**464**	(433-498)	**1,186**	(1,124-1,251)	**236**	(213-262)	**198**	(172-228)	**1,029**	(971-1,090)	**524**	(483-568)
New Mexico	**19**	(17-22)	**44**	(35-55)	**9**	(7-12)	**11**	(7-19)	**40**	(31-50)	**32**	(25-40)
Oregon	**48**	(43-53)	**66**	(56-78)	**22**	(18-26)	**15**	(10-20)	**53**	(46-62)	**45**	(39-52)

Appendix I. Estimated Number of Persons with Substance Abuse Behaviors for 25 MSAs. Estimates Based on Data From the 1991-1993 NHSDA. 95% Confidence Intervals

Washington	**80**	(71-90)	**144**	(120-172)	**35**	(29-41)	**21**	(15-29)	**145**	(123-171)	**78**	(66-93)

MSAs	Past Year Dependency				Past Year Treatment						Past Year Arrest	
	Dependent On Illicit Drugs		Dependent On Alcohol		Received Treatment For Illicit Drug Use		Received Treatment For Alcohol Use		Needed Treatment For Illicit Drug Use			
Anaheim-Santa Ana, CA	**43**	(38-50)	**120**	(107-134)	**20**	(15-26)	**11**	(8-16)	**89**	(78-101)	**31**	(24-39)
Atlanta, GA	**30**	(26-35)	**108**	(93-126)	**18**	(14-22)	**9**	(7-13)	**117**	(102-134)	**42**	(37-48)
Baltimore, MD	**18**	(15-21)	**43**	(35-53)	**12**	(9-15)	**9**	(6-13)	**45**	(38-54)	**26**	(22-30)
Boston, MA	**53**	(46-60)	**138**	(120-157)	**30**	(24-38)	**69**	(53-88)	**122**	(107-139)	**52**	(44-62)
Chicago, IL	**45**	(41-49)	**172**	(166-179)	**27**	(24-29)	**36**	(33-38)	**137**	(131-144)	**75**	(70-81)
Dallas, TX	**35**	(31-39)	**70**	(59-83)	**14**	(11-17)	**10**	(7-14)	**52**	(44-61)	**33**	(28-39)
Denver, CO	**17**	(16-18)	**61**	(59-64)	**9**	(8-9)	**13**	(12-14)	**52**	(50-54)	**30**	(29-32)
Detroit, MI	**42**	(37-49)	**129**	(112-149)	**28**	(24-33)	**18**	(14-25)	**129**	(113-146)	**63**	(55-73)
El Paso, TX	**4**	(3-5)	**17**	(12-23)	**3**	(2-4)	**2**	(1-5)	**11**	(9-15)	**7**	(5-10)
Houston, TX	**56**	(50-63)	**52**	(41-66)	**22**	(18-26)	**19**	(14-25)	**103**	(90-116)	**41**	(35-49)
Los Angeles, CA	**107**	(99-115)	**508**	(496-521)	**54**	(46-63)	**36**	(33-39)	**288**	(277-300)	**132**	(124-141)
Miami-Hialeah, FL	**12**	(11-13)	**33**	(31-34)	**9**	(7-11)	**8**	(7-8)	**35**	(33-37)	**20**	(18-21)
Minneapolis-St. Paul, MN	**23**	(20-27)	**70**	(60-83)	**19**	(15-24)	**19**	(13-27)	**49**	(42-57)	**36**	(31-42)
Nassau-Suffolk, NY	**22**	(19-25)	**24**	(18-31)	**14**	(11-18)	**11**	(8-15)	**41**	(34-49)	**25**	(20-30)
New York, NY	**66**	(61-73)	**125**	(119-132)	**47**	(37-59)	**18**	(16-21)	**177**	(168-186)	**56**	(51-62)
Newark, NJ	**21**	(19-25)	**38**	(30-47)	**13**	(11-16)	**12**	(9-15)	**28**	(23-35)	**21**	(17-27)
Oakland, CA	**51**	(46-57)	**67**	(56-80)	**24**	(21-27)	**36**	(29-44)	**90**	(80-101)	**33**	(28-39)
Philadelphia, PA-NJ	**40**	(34-46)	**114**	(99-131)	**33**	(29-38)	**25**	(19-32)	**121**	(108-135)	**55**	(48-63)
Phoenix, AZ	**27**	(24-31)	**64**	(54-76)	**12**	(9-14)	**12**	(8-17)	**56**	(48-65)	**26**	(20-35)
San Antonio, TX	**14**	(12-17)	**26**	(18-39)	**5**	(4-8)	**4**	(2-8)	**30**	(23-40)	**17**	(14-20)
San Bernardino, CA	**44**	(39-51)	**83**	(72-95)	**24**	(20-29)	**25**	(19-34)	**81**	(71-93)	**51**	(43-61)
San Diego, CA	**29**	(24-35)	**62**	(52-74)	**15**	(12-20)	**11**	(7-16)	**67**	(57-78)	**45**	(38-54)
St. Louis, MO-IL	**19**	(17-22)	**49**	(42-58)	**15**	(12-18)	**12**	(9-17)	**55**	(47-64)	**42**	(36-49)
Tampa-St. Petersburg, FL	**18**	(15-21)	**38**	(30-48)	**12**	(10-15)	**8**	(5-11)	**38**	(32-45)	**26**	(22-31)
Washington, DC	**36**	(34-38)	**120**	(117-124)	**19**	(17-20)	**29**	(28-31)	**89**	(86-92)	**45**	(42-47)