

National Longitudinal Study of Adolescent Health

Wave II Political Context Database

James Fowler
Jaime Settle
University of California, San Diego
Department of Political Science

Tim Monbureau
University of North Carolina at Chapel Hill
Carolina Population Center

April 2010

This research was supported by grants SES-0719404 and SES-0843907 from the National Science Foundation. Further information may be obtained by contacting James Fowler (jhfowler@ucsd.edu) or Jaime Settle (jsettle@ucsd.edu).

TABLE OF CONTENTS

I. INTRODUCTION

1. Documentation Structure
2. Data Form
3. Variable Naming Conventions
4. Missing Data

II. SUBJECT INDEX

III. DATA DICTIONARY

APPENDIX A

APPENDIX B

I. INTRODUCTION

Extant research in political science evinces that the political environment characterizing a community influences individual political and social decisions. These contextual influences have consequence for political socialization in adolescence and continue to exert effect throughout the life course. The Add Health Political Context Database provides an array of measures that describe the political environments in which Add Health respondents reside and, thereby, enables researchers to explore the role of certain contextual influences on adolescent and early adult political behaviors. The Wave II component of the Political Context Database contains 12 contextual variables compiled from a source of political data and then matched to individual Add Health respondent locations. These contextual variables include measures of election results for gubernatorial, presidential, and senatorial races.

The assembled contextual data underwent geocoding in order to attach measures to Add Health respondent locations at the county-level. The merge matched data for the years 1995 to 1996 to respondent locations to coincide with the time of individual Wave II In-Home survey participation. The Wave II In-Home survey was conducted from April to August 1996.¹

To easily and successfully use the data comprising the Political Context Database, please carefully review the following subsections discussing documentation structure, data form, variable naming conventions, and types of missing data. Users are strongly encouraged to read all sections of this document before attempting to utilize any of the contextual data provided.

Documentation Structure

Each section and appendix has an introduction that describes the utility and structure of the information presented. The remainder of this subsection will discuss the relationship between the different sections of the Political Context Database guide.

The INTRODUCTION presents the overarching information necessary to understand the contents and conventions of the data present in the database.

Topics covered in the data are alphabetically listed in the SUBJECT INDEX section. This section is best used to identify general topics, which can guide the identification of specific variables described in the DATA DICTIONARY.

The DATA DICTIONARY lists each variable with its name and, when relevant, the formula used to construct it. This section also provides descriptive subject headings as well as references to source documentation detailed in Appendix A.

APPENDIX A supplies the user source descriptions and source variable documentation. When the data collection or calculation processes require additional explanation, the user is encouraged to refer to any references and notes presented. This appendix is alphabetically ordered by source.

APPENDIX B provides a standard codebook for the Wave II Political Context Database. Variable order reflects that of the data file. The codebook identifies a variable's range, missingness, and frequencies. It also provides value labels intended to aid in the interpretation of both reserve code values that define missing data and categorical variable values.

Data Form

¹ Care must be taken when interpreting the interaction of certain electoral variables with the Add Health survey responses. The Political Context Database provides contextual variables measured at a range of different years, in part, to afford the data user the ability to choose years appropriate for answering the particular questions of interest or specifying particular models. Checking the respondent interview date available in the main In-Home survey will facilitate the proper selection of variable years.

The Wave II Political Context Database contains one observation for each respondent in the corresponding wave of the Add Health In-Home survey. The contextual data were spatially and temporally matched to Add Health respondents based on both the time of the In-Home interview and location of the respondents' residence at Wave II. When using these contextual data consider the possibility that respondents may have moved just prior to or after the date of their collection.

This contextual database contains a total of 13 variables. The first variable, an eight character string, serves as the Respondent Identifier (AID). This identifier enables linkage of the Political Context Database to the Add Health In-Home respondent-level data. Aside from the AID, each variable present is numeric in type. Variable order in the data file mirrors the order in which variables appear in the DATA DICTIONARY.

Contextual data matched to the Wave II respondent location include information regarding gubernatorial election results for the years 1995 and 1996, senatorial election results for 1996, and presidential election results for 1996.

Variable Naming Conventions

Each variable name has a length of 8 characters. The first two of which designate the Add Health wave, such that W2 denotes Wave II. The third and fourth characters, PC, identify the specific data file as belonging to the Political Context Database. Subjects or sub-subjects are represented by the fifth character. In the Wave II Political Context Database, character and subject assignments are as follows:

G	Gubernatorial Election Results
P	Presidential Election Results
S	Senatorial Election Results

The sequence of numbers occupying the sixth position in the variable name, that increase in increments of 1 by alphabetically ordered subject or sub-subject, make variable names unique and, thereby, distinguish variables within subjects or sub-subjects. The final two characters denote the year in which the variable was measured. For example, "95" would indicate data from the year 1995. When measures fail to fall on a specific year but rather fall within a span of several years, the variable name indicates the most recent year on which the variable may have been measured; that is, the time span's upper bound.

Missing Data

Add Health data generally distinguishes types of missing data with special reserve codes. Wave II of the Political Context Database has two forms of missing data: those legitimately missing due to a lack of applicability and those missing as a result of a failure to successfully assign geocodes to respondents.

Certain missing responses result from a lack of applicable data. For instance, states hold senatorial elections in different years; consequently, respondents inhabiting states without senate races in 1996 lack non-missing senatorial election results. These respondents, therefore, have the Add Health reserve code denoting legitimate skip assigned to their senatorial election variables. This particular reserve code equals a 7 preceded by a series of 9s, the number of which varies depending upon the variables' maximum, non-missing values.

The second missing data code identifies respondent data missing as a result of a failure to successfully attach useable geocodes to respondent locations. Respondents not associated with a county or state due to this failure receive the Add Health reserve code for geocode missing, which equals 8 preceded by a series of 9s; the number of which varies depending upon the variables' maximum values. A total of 83 respondents were ultimately assigned missing geocode values.

II. SUBJECT INDEX

The SUBJECT INDEX lists the various subjects measured by the variables contained in the Wave II Political Context Database, variables nominally listed in the DATA DICTIONARY. Subject headings are alphabetized and listings are grouped into subheadings when appropriate.

Election Results

Gubernatorial

Proportion of votes cast for the Democratic gubernatorial candidate. (1995-1996)

Proportion of votes cast for the Republican gubernatorial candidate. (1995-1996)

The difference between the proportion of votes cast for the Democratic candidate and the Republican candidate. Positive numbers indicate more votes for the Democrat. (1995-1996)

Presidential

Proportion of votes cast for the Democratic presidential candidate. (1996)

Proportion of votes cast for the Republican presidential candidate. (1996)

The difference between the proportion of votes cast for the Democratic candidate and the Republican candidate. Positive numbers indicate more votes for the Democrat. (1996)

Senatorial

Proportion of votes cast for the Democratic senatorial candidate. (1996)

Proportion of votes cast for the Republican senatorial candidate. (1996)

The difference between the proportion of votes cast for the Democratic candidate and the Republican candidate. Positive numbers indicate more votes for the Democrat. (1996)

III. DATA DICTIONARY

Organized by subject, the DATA DICTIONARY lists all the variables contained in the Wave II Political Context Database.

The following information is provided in the DATA DICTIONARY for each variable: Name, Description, and Formula.

Name

Variables obtained from source files are renamed to adhere to established Add Health nomenclature, which is designed to provide information about the variable. This renaming process also serves to emphasize that source variables typically undergo some degree of transformation before their inclusion in Add Health and, therefore, no longer necessarily possess values equal to those found in their source. See the INTRODUCTION section entitled Variable Naming Conventions for an explanation of Add Health's nomenclature

Description

The description column provides an extended variable label. The data collection year or span of years for each variable is also included in the description.

Formula

Formulae for each contextual variable provide information about the component source variables used in its construction. Source information and original variable descriptions can be found in APPENDIX A, which is alphabetically ordered by source. Each original variable in the appendix is listed with a description along with the original source variable name or reference name. The formulae used for converting percentages to proportions and rounding do not receive explication.

Name	Description	Formula
Election Results: Gubernatorial		
W2PCG196	Proportion of votes cast for the Democratic gubernatorial candidate in 1995 or 1996.	DEMGUB95 or DEMGUB96

W2PCG296	Proportion of votes cast for the Republican gubernatorial candidate in 1995 or 1996.	REPGUB95 or REPGUB96
W2PCG396	Difference between the proportion of votes cast for the Democratic gubernatorial candidate and Republican gubernatorial candidate in 1995 or 1996. Positive numbers indicate more votes for the Democrat.	W2PCG196 - W2PCG296
Election Results: Presidential		
W2PCP196	Proportion of votes cast for the Democratic presidential candidate in 1996.	DEMPRE96
W2PCP296	Proportion of votes cast for the Republican presidential candidate in 1996.	REPPRE96
W2PCP396	Difference between the proportion of votes cast for the Democratic presidential candidate and the Republican presidential candidate in 1996. Positive numbers indicate more votes for the Democrat.	W2PCP196 - W2PCP296
Election Results: Senatorial		
W2PCS196	Proportion of votes cast for the Democratic senatorial candidate in 1996.	DEMSEN96
W2PCS296	Proportion of votes cast for the Republican senatorial candidate in 1996.	REPSSEN96
W2PCS396	Difference between the proportion of votes cast for the Democratic senatorial candidate and the Republican senatorial candidate in 1996. Positive numbers indicate more votes for the Democrat.	W2PCS196 - W2PCS296
Voter Registration Law²		
W2PCA196	Mail voter registration allowed after the National Voter Registration Act, 1996	ANVRACD
W2PCA296	When absentee ballots are due, 1996	ANVRAAD
W2PCA396	The restrictiveness of the eligibility for absentee voting, 1996	ANVRAAE

² Laws in effect before the act were measured with data from 1994-1995, while laws in effect after were measured with data from 1995-1996, as state compliance was not required until 1995.

APPENDIX A—Source Notes and Source Variable Documentation

SOURCE DESCRIPTION

Congressional Quarterly Press Electoral Database

This appendix contains information about each data source used to construct the variables contained in the Wave II Political Context Database. The appendix is ordered alphabetically by source. Essential notes are presented along with a complete list of the original source variables for users interested in more detailed information about variables' measurement and source.

For each source, data file descriptions and user notes are followed by a table of source variables. This table provides the source variable name or assigned reference name and a variable description. Each of these is addressed below. Users can reference source variable names in Appendix A to better understand the construction of the Political Context Database variables as described by formulae provided in the DATA DICTIONARY.

Source Variable Name

Typically, the source name identifies the original name of the variables used to create those comprising the Political Context Database. Since all of the source data used to produce the Wave II Political Context Database originate from tables lacking variable names, names have been assigned in this codebook to facilitate reference to formulae in the DATA DICTIONARY.

Source Variable Description

This column provides a complete description of each source variable involved in the construction of the Add Health Political Context Database. The description aids in the identification of source variables taken from tables without typical variable nomenclature and clarification of contextual variable meanings when a derived measure's definition lacks sufficient detail. Though descriptions presented in APPENDIX A often directly reflect information provided by the source, users may find it necessary to consult source documentation for additional discussion of the original measures.

Congressional Quarterly Press Voting and Elections Collection

The Congressional Quarterly Press (CQ Press) Voting and Elections Collection has data, analyses, explanations, and historical material relating to voters, major and minor political parties, campaigns and elections, and both historical and contemporary races for Congress, the presidency, and governorships. The data can be exported and downloaded by political race, state, and year. Data at the county level were downloaded for the 1995 and 1996 races of interest into a table format. Thus, no original variable names are provided.

Data used from CQ Press for constructing measures in the Wave II Political Context Database are: gubernatorial election results by county for the years 1995 and 1996, presidential election results by county for 1996, and senatorial election results by county for 1996.

Independent cities with their own 5-digit FIPS codes are treated as county equivalents in accordance with U.S. Census Bureau practices.

The CQ Press data originally reported the total number of votes cast in an election, as well as the number of votes cast specifically for the Democratic candidate and the number of votes cast for the Republican candidate. The source reports the percentage of votes cast for each party candidate, as well. The Add Health Political Context Database relies upon the percentage of votes cast for the Democratic candidate and the percentage cast for the Republican candidate. Note that the summation of both proportions may not equal 1, especially in the presence of strong third party or unaffiliated candidates. Consequently, these summations will provide an indication of the strength of third party or unaffiliated candidates in a particular election. Finally, a competitiveness score was generated by calculating the difference between the proportion of votes cast for the Democratic candidate minus the proportion of votes cast for the Republican candidate. Positive numbers indicate an advantage for the Democratic candidate.

Gubernatorial Election Results

Gubernatorial races occur every year, but most states hold races in even years when no concurrent presidential election occurs. The Political Contextual Database reports results of the gubernatorial race in an election cycle so as to suppress in which year the race happened. The reported cycle at Wave II encompasses 1995 and 1996.

Reference Name	Description:
DEMGUB95	The percent of votes cast for the Democratic gubernatorial candidate in the county, 1995.
REPGUB95	The percent of votes cast for the Republican gubernatorial candidate in the county, 1995.
DEMGUB96	The percent of votes cast for the Democratic gubernatorial candidate in the county, 1996.
REPGUB96	The percent of votes cast for the Republican gubernatorial candidate in the county, 1996.

Presidential Election Results

Reference Name	Description:
DEMPRE96	The percent of votes cast for the Democratic presidential candidate in the county, 1996.
REPPRE96	The percent of votes cast for the Republican presidential candidate in the county, 1996.

Senatorial Election Results

Reference Name	Description:
DEMSEN96	The percent of votes cast for the Democratic senatorial candidate in the county, 1996.
REPSSEN96	The percent of votes cast for the Republican senatorial candidate in the county, 1996.

The Council of State Governments, Book of States Files

Data regarding changes in state voter registration law after the National Voter Registration Act (NVRA) of 1993 were acquired from The Book of the States, a reference book published since 1935 that provides data and comparisons for all 56 U.S. states and territories. Registration law data originated from each state's election administration offices. Data joined to the Political Context Database were obtained from tables reporting voter registration law before and after NVRA went into effect. Laws in effect before the act were measured with data from 1994-1995, while laws in effect after the act were measured with data from 1995-1996, as state compliance was not required until 1995.

Reference Name:	Description:
ANVRACD	The number of days prior to the election required for a voter to register for an upcoming election, after the National Voter Registration Act's compliance deadline (1995-1996). Measured in days.
ANVRAAD	When absentee ballots are due, after the National Voter Registration Act's compliance deadline (1995-1996). <ol style="list-style-type: none">1. Before election day2. On election day3. After election day
ANVRAAE	The restrictiveness of the eligibility for absentee voting, after the National Voter Registration Act's compliance deadline (1995-1996). <ol style="list-style-type: none">0. Single category (military/overseas or disabled)1. Temporarily out of jurisdiction2. Multiple categories of eligibility3. Everyone

APPENDIX B—Codebook

Listed in the same order as present in the data file, the following variable entries allow users to quickly determine variable frequencies, ranges, missingness, and value labels for both the reserve codes that define missing data and categorical variables.

W2POLCON: Wave II Political Context Database

Wave II Political Context Database

Number of observations: 14,738

AID		Char	Respondent Identifier NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
1	0%	10316654	
1	0%	10316952	
1	0%	10506342	
1	0%	10570810	
1	0%	10606128	
14728	100%	Values omitted	
1	0%	99886995	
1	0%	99886996	

1	0%	99886997	
1	0%	99886998	
1	0%	99886999	

W2PCA196		Num	The number of days prior to the election required for a voter to register for an upcoming election, post NVRA
Frequency	Percent	Value	Label
378	3%	0	
256	2%	10	
370	3%	15	
521	4%	20	
1511	10%	25	
11619	79%	30	
83	1%	98	Geocode missing

W2PCA296		Num	When absentee ballots are due, post NVRA
Frequency	Percent	Value	Label
2589	18%	1	Before election day
11311	77%	2	On election day
755	5%	3	After election day

83	1%	8	Geocode missing
----	----	---	-----------------

W2PCA396		Num	The restrictiveness of the eligibility for absentee voting, post NVRA
Frequency	Percent	Value	Label
1002	7%	0	Single category
1235	8%	1	Temporarily out of jurisdiction
12418	84%	2	Multiple categories of eligibility
83	1%	8	Geocode missing

W2PCG196		Num	Proportion of votes cast for the Democratic gubernatorial candidate, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
1	0%	0.195	
1	0%	0.228	
213	1%	0.231	
3	0%	0.260	
1	0%	0.263	
2332	16%	.29-.656	NOTE: Range of values omitted from display
187	1%	0.684	

1	0%	0.713	
1	0%	0.716	
11915	81%	7	Legitimate skip
83	1%	8	Geocode missing

W2PCG296		Num	Proportion of votes cast for the Republican gubernatorial candidate, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
1	0%	0.204	
1	0%	0.284	
193	1%	0.293	
67	0%	0.301	
1	0%	0.323	
2262	15%	.338-.74	NOTE: Range of values omitted from display
1	0%	0.757	
213	1%	0.769	
1	0%	0.805	
11915	81%	7	Legitimate skip
83	1%	8	Geocode missing

W2PCG396		Num	Difference between the proportion of votes cast for the Democratic and Republican gubernatorial candidates, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
1	0%	-0.610	
213	1%	-0.538	
1	0%	-0.529	
3	0%	-0.480	
1	0%	-0.474	
2332	16%	-.41- .333	NOTE: Range of values omitted from display
187	1%	0.391	
1	0%	0.432	
1	0%	0.509	
11915	81%	7	Legitimate skip
83	1%	8	Geocode missing

W2PCP196		Num	Proportion of votes cast for the Democratic presidential candidate, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
3	0%	0.248	

153	1%	0.255	
130	1%	0.271	
8	0%	0.274	
12	0%	0.281	
13977	95%	.291- .762	NOTE: Range of values omitted from display
178	1%	0.793	
1	0%	0.800	
188	1%	0.801	
5	0%	0.858	
83	1%	8	Geocode missing

W2PCP296		Num	Proportion of votes cast for the Republican presidential candidate, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
5	0%	0.105	
1	0%	0.138	
188	1%	0.151	
178	1%	0.155	
1	0%	0.157	

13995	95%	.208-.623	NOTE: Range of values omitted from display
1	0%	0.632	
130	1%	0.672	
3	0%	0.674	
153	1%	0.680	
83	1%	8	Geocode missing

W2PCP396		Num	Difference between the proportion of votes cast for the Democratic and Republican presidential candidates, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
3	0%	-0.426	
153	1%	-0.425	
130	1%	-0.401	
1	0%	-0.330	
1	0%	-0.324	
13995	95%	-.31-.565	NOTE: Range of values omitted from display
178	1%	0.638	
188	1%	0.650	

1	0%	0.662	
5	0%	0.753	
83	1%	8	Geocode missing

W2PCS196		Num	Proportion of votes cast for the Democratic senatorial candidate, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
3	0%	0.087	
5	0%	0.167	
3	0%	0.171	
150	1%	0.174	
164	1%	0.197	
6963	47%	.199-.77	NOTE: Range of values omitted from display
1	0%	0.777	
144	1%	0.858	
1	0%	0.874	
7221	49%	7	Legitimate skip
83	1%	8	Geocode missing

W2PCS296		Num	Proportion of votes cast for the Republican senatorial candidate, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
1	0%	0.126	
144	1%	0.142	
1	0%	0.223	
1	0%	0.230	
19	0%	0.240	
7257	49%	.281- .812	NOTE: Range of values omitted from display
3	0%	0.818	
5	0%	0.821	
3	0%	0.903	
7221	49%	7	Legitimate skip
83	1%	8	Geocode missing

W2PCS396		Num	Difference between the proportion of votes cast for the Democratic and Republican senatorial candidates, 1996 NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
3	0%	-0.816	

5	0%	-0.654	
3	0%	-0.647	
150	1%	-0.638	
164	1%	-0.590	
6963	47%	-.588-.54	NOTE: Range of values omitted from display
1	0%	0.554	
144	1%	0.716	
1	0%	0.748	
7221	49%	7	Legitimate skip
83	1%	8	Geocode missing