APPENDIX: CONSTRUCTED VARIABLES

In order to facilitate and encourage the common use of several key variables, and to help individual users to avoid having to "reinvent the wheel" by creating certain scales and indices, the released dataset also includes the following constructed variables at the end of the variable list:

1) Generalized trust (mean of three questions, all with 0-10 scales)
2) Political action (sum of 14 possible types of political acts)
3) Party identification (Democrat v. Republican, with different gradations)
4) Participation in voluntary organizations (inactive, passive, active, super-active)
5) Citizenship norms (citizen duty v. engaged citizen)
6) The diversity of social networks (close friends, neighborhood, school/work/former work)
7) Racial prejudice/negative stereotypes (against Asians, Blacks, Hispanics, and Whites)
8) National pride (of being an American)
9) Attitudes toward immigrants (mean of three questions, all with 0-10 scales)
10) Demographic factors (race, income, education, age, and gender)

For each, we give a brief explanation of the transformation that was made to create the new variable, and we provide the SPSS syntax that was used.

1) Generalized Trust

(new variable = GENTRUST)

This index, measuring R's level of interpersonal trust, comprises three variables, PPLTRST, PPLFAIR, and PPLHLP. The index is not constructed for the cases with more than two missing values among these items. Higher values denote higher levels of trust.

```
COMPUTE GENTRUST = MEAN.2(ppltrst, pplfair, pplhlp).
VAR LABELS GENTRUST Generalized Trust.
EXECUTE.
```

2) Political Action

(new variable = POLACTUS14)

This index measures the total number of political acts that a respondent has undertaken in the previous 12 months. It includes all 14 types of political action included in the U.S. CID survey (10 of which were included in the ESS; the other four were added to the U.S. survey).[1] The maximum possible score on POLACTUS14 is therefore 14. Higher values denote higher levels of political action.

As a first step, the variables were recoded so that "no" responses were counted as 0, rather than 2:

```
RECODE contplt (1=1) (2=0) INTO contpltd .
RECODE wrkprty (1=1) (2=0) INTO wrkprtyd .
RECODE WRKCAMP (1=1) (2=0) INTO WRKCAMPd .
RECODE wrkorg (1=1) (2=0) INTO wrkorgd .
RECODE badge (1=1) (2=0) INTO badged .
RECODE sgnptit (1=1) (2=0) INTO sgnptitd .
RECODE pbldmn (1=1) (2=0) INTO pbldmnd .
RECODE bcdprd (1=1) (2=0) INTO bctprdd .
RECODE bghtprd (1=1) (2=0) INTO bghtprdd .
RECODE dntmny (1=1) (2=0) INTO dntmnyd .
RECODE ilglpst (1=1) (2=0) INTO ilglpstd .
RECODE WEBPOL (1=1) (2=0) INTO WEBPOLd .
RECODE FWDPOL (1=1) (2=0) INTO FWDPOLd .
RECODE WEBPTC (1=1) (2=0) INTO WEBPTCd .
EXECUTE .
```

Then the full index:

```
COMPUTE POLACTUS14 = contpltd + wrkprtyd + WRKCAMPd + wrkorgd + badged +
sgnptitd + pbldmnd + bctprdd + bghtprdd + dntmnyd + ilglpstd
 + WEBPOLd + FWDPOLd + WEBPTCd .
EXECUTE .
VAR LABELS POLACTUS14 Political Action.
```

[1] Note that these variables names all end in "US14" in order to distinguish them from similar variables (ending in "ESS10") that can be applied to the 10 types of political action that were included in the ESS.

3) Party Identification

(new variable = PARTYID)

The U.S. CID survey includes a question and several follow-ups to determine the respondent's party identification. The syntax below represents the process by which we derived a single PARTYID variable, with the familiar categories of Strong Republican, Republican, Leaning Republican, Independent, Leaning Democrat, Democrat, and Strong Democrat.

```
COMPUTE PARTYID = -1 .
missing values PID () .
missing values PID5 (-1) .
missing values PIDC (-1) .
DO IF (PID EQ 1) .
IF (PID5 EQ 1)PARTYID =1 .
IF (PID5 EQ 2 OR PID5 EQ 8 OR PID5 EQ 9)PARTYID =2 .
ELSE IF (PID EQ 2) .
IF (PID5 EQ 1)PARTYID =7 .
IF (PID5 EQ 2 OR PID5 EQ 8 OR PID5 EQ 9)PARTYID =6 .
ELSE IF (PID EQ 3) .
IF (PIDC EQ 1)PARTYID = 3 .
IF (PIDC EQ 2)PARTYID = 5 .
IF (PIDC EQ 3)PARTYID = 4 .
IF (PIDC EQ 7)PARTYID = 4 .
IF (PIDC EQ 8)PARTYID = 4 .
IF (PIDC EQ 9)PARTYID = 4 .
ELSE IF (PID EQ 4) .
IF (PIDC EQ 1)PARTYID = 3 .
IF (PIDC EQ 2)PARTYID = 5 .
IF (PIDC EQ 3)PARTYID = 4 .
ELSE IF (PID EQ 5) .
IF (PIDC EQ 1)PARTYID = 3 .
IF (PIDC EQ 2)PARTYID = 5 .
IF (PIDC EQ 3)PARTYID = 4 .
IF (PIDC EQ 7)PARTYID = 4 .
IF (PIDC EQ 8)PARTYID = 4 .
IF (PIDC EQ 9)PARTYID = 4 .
END IF .
VALUE LABELS PARTYID
  (1)STRONG REPUBLICAN(2)REPUBLICAN(3)LEANING REPUBLICAN
  (4)INDEPENDENT(5)LEANING DEMOCRAT
  (6)DEMOCRAT (7)STRONG DEMOCRAT .
VAR LABELS PARTYID PARTY IDENTIFICATION, CONSTRUCTED .
missing values PARTYID (-1) .
EXECUTE .
```

4) Participation in Voluntary Organizations

(new variables = MEMBINDEXUS17, PARTICINDEXUS17, DONATEINDEXUS17, VOLWORKINDEXUS17, NONEINDEXUS17, ORGINVINDEXUS17)

A core component of the U.S. CID survey is an extensive battery of questions on respondents' involvement with voluntary organizations. The battery includes an opening question that allows respondents to specify whether they are members, participated, donated money, or did voluntary work in each of 17 different types of voluntary organizations (12 of which were included in the ESS; the other five were added to the U.S. CID). Multiple responses are allowed for each of these four options; respondents who perform none of these activities were coded as "none."

Although it is quite complex, there are several advantages to this detailed battery. One is that it allows us to go beyond simple membership as a measure of involvement with organizations. Another is that researchers can look at different combinations of types of activity.

Moreover, the battery includes a series of follow-up questions that were asked to all respondents who are either members, participate, donate money, or do voluntary work in at least one organization, for each organization in which they are involved. These include whether they have personal friends in the organization, how long have they have been a member, how often they participate in group activities and meetings of the organization, and how racially/ethnically diverse the organization is. Yet another follow-up asks in which organization the respondent is the most active, and then inquires as to whether the respondent has undertaken a series of civic activities within that group.

This battery allows for numerous possible forms of recoding, transformations, and analysis. The released dataset includes a series of five basic indices, corresponding to the total number of organizations that respondents either are members of, participate in, donate money to, do voluntary work for, or "none of the above," for the 17 orgs used in the U.S. survey.[2]

Finally, we include a more complex—but ultimately more accurate—constructed variable, an *organizational involvement index*, which distinguishes between members who are inactive, passive, active, and super-active in voluntary organizations.

This conceptualization assumes that some forms of organizational involvement represent greater "activeness" than others (e.g., being a member or donating money correspond to a lower degree of involvement than participating or volunteering). And several different combinations of these four types of involvement lead to the following overall classifications:

0 Inactive (no involvement of any kind)
1 Passive (member only, donor only, or member and donor only)
2 Active (volunteer only, participant only, any combination of either volunteer or participant with member or donor, or the combination of participant, donor, and member)
3 Super-active (volunteer and participant only, or 3 or more of the above, excluding the combination participant, donor, member)

This measure corresponds to the respondent's *highest* level of involvement in at least one of the organizations. For example, if a respondent participates or volunteers for one organization and is also a member of another group, she will be coded as participating or volunteering overall (i.e., as "active"); if

[2] Note that these variables names all end in "US17" in order to distinguish them from similar variables (ending in "ESS12") that can be applied to the 12 organizations that were included in the ESS.

another respondent participates and volunteers in just one organization, he will be coded as such (i.e., as "super-active").[3] What matters from the perspective of social capital and civil society is that people be interacting in *some* organized capacity, and since computing an aggregate measure that somehow multiplies the type of participation by the number of organizations could be very misleading,[4] this appears to be the best possible measure to distinguish among people who participate in various ways in voluntary organizations.

The following syntax was used to create these variables:

[Note: as a first step, which is not shown in the syntax below, all "Refused," "Don't Know," and "No Answer" responses were recoded as "None" on the basic membership questions.]

The first step creates the basic additive indices for the total number of organizations that people either are members of, participate in, donate money to, do voluntary work for, or none of the above, for the 17 orgs used in the US survey:

```
COMPUTE MEMBINDEXUS17 = sptcmmb + cltommb + trummb + prfommb + cnsommb +
hmnommb + epaommb + rlgommb + prtymmb + setommb + sclcmmb + NEIVMMB +
VETVMMB + NATVMMB + IMPVMMB + SOCVMMB + othvmmb .
EXECUTE .
variable label MEMBINDEXUS17 'number of orgs in which R is a member (all 17 US orgs)'.

COMPUTE PARTICINDEXUS17 = sptcptp + cltoptp + truptp + prfoptp + cnsoptp + hmnoptp +
epaoptp + rlgoptp + prtyptp + setoptp + sclcptp + NEIVPTP + VETVPTP + NATVPTP +
IMPVPTP + SOCVPTP + othvptp .
EXECUTE .
variable label PARTICINDEXUS17 'number of orgs in which R participates (all 17 US orgs)'.

COMPUTE DONATEINDEXUS17 = sptcdm + cltodm + trudm + prfodm + cnsodm + hmnodm
+ epaodm + rlgodm + prtydm + setodm + sclcdm + NEIVDM + VETVDM + NATVDM +
IMPVDM + SOCVDM + othvdm .
EXECUTE .
variable label DONATEINDEXUS17 'number of orgs to which R donates money (all 17 US
orgs)'.

COMPUTE VOLWORKINDEXUS17 = sptcvw + cltovw + truvw + prfovw + cnsovw + hmnovw
+ epaovw + rlgovw + prtyvw + setovw + sclcvw + NEIVVW + VETVVW + NATVVW +
IMPVVW + SOCVVW + othvvw .
EXECUTE .
variable label VOLWORKINDEXUS17 'number of orgs in which R does voluntary work (all 17
US orgs)'.
```

[3] Note that our measure equally weights participation across the twelve different types of organizations. For example according to our measure, super-active participation in sports clubs is equivalent to super-active participation in human rights groups. While in principle it might be preferable to give extra weight to certain types of organizations that are deemed to be more "civic" and less "recreational," this would require potentially erroneous assumptions about the meaning of participation across groups.

[4] For example, it would weigh "down" a respondent who is super-active in one organization, while also being active in several others, thus making it appear that she is less active than someone who is super-active in just one organization.

```
COMPUTE NONEINDEXUS17 = sptcnn + cltonn + trunn + prfonn + cnsonn + hmnonn +
epaonn + rlgonn + prtynn + setonn + sclcnn + NEIVNN + VETVNN + NATVNN + IMPVNN +
SOCVNN + othvnn .
EXECUTE .
variable label NONEINDEXUS17 'number of orgs in which R does nothing (all 17 US orgs)'.
```

The next procedure is an intermediate step that creates dummy variables (not included in the dataset), which then allow us to create the final ORGINVINDEXUS17 index:

```
RECODE
  MEMBINDEXUS17
  (0=0)  (1 thru highest=1)  INTO  memdum .
VARIABLE LABELS memdum 'Member dummy' .
VALUE LABELS memdum 0 'No' 1 'Yes' .
EXECUTE .

RECODE
  PARTICINDEXUS17
  (0=0)  (1 thru highest=1)  INTO  pardum .
VARIABLE LABELS pardum 'Participant dummy' .
VALUE LABELS pardum 0 'No' 1 'Yes'  .
EXECUTE .

RECODE
  DONATEINDEXUS17
(0=0)  (1 thru highest=1)  INTO  dondum .
VARIABLE LABELS dondum 'Donor dummy' .
VALUE LABELS dondum 0 'No' 1 'Yes'  .
EXECUTE .

RECODE
  VOLWORKINDEXUS17
  (0=0)  (1 thru highest=1)  INTO  voldum .
VARIABLE LABELS voldum 'Volunteer dummy' .
VALUE LABELS voldum 0 'No' 1 'Yes' .
```

The final step creates the organizational involvement index:

```
IF (memdum = 0 AND pardum = 0 AND dondum = 0 AND voldum = 0) ORGINVINDEXUS17 = 0 .
IF (memdum = 1 AND pardum = 0 AND dondum = 0 AND voldum = 0) ORGINVINDEXUS17 = 1 .
IF (memdum = 0 AND pardum = 1 AND dondum = 0 AND voldum = 0) ORGINVINDEXUS17 = 1 .
IF (memdum = 0 AND pardum = 0 AND dondum = 1 AND voldum = 0) ORGINVINDEXUS17 = 2 .
IF (memdum = 0 AND pardum = 0 AND dondum = 0 AND voldum = 1) ORGINVINDEXUS17 = 2 .
IF (memdum = 1 AND pardum = 1 AND dondum = 0 AND voldum = 0) ORGINVINDEXUS17 = 1 .
IF (memdum = 0 AND pardum = 1 AND dondum = 1 AND voldum = 0) ORGINVINDEXUS17 = 2 .
IF (memdum = 0 AND pardum = 0 AND dondum = 1 AND voldum = 1) ORGINVINDEXUS17 = 3 .
IF (memdum = 0 AND pardum = 1 AND dondum = 0 AND voldum = 1) ORGINVINDEXUS17 = 2 .
IF (memdum = 1 AND pardum = 0 AND dondum = 1 AND voldum = 0) ORGINVINDEXUS17 = 2 .
IF (memdum = 1 AND pardum = 0 AND dondum = 0 AND voldum = 1) ORGINVINDEXUS17 = 2 .
IF (memdum = 1 AND pardum = 1 AND dondum = 1 AND voldum = 0) ORGINVINDEXUS17 = 2 .
IF (memdum = 1 AND pardum = 1 AND dondum = 0 AND voldum = 1) ORGINVINDEXUS17 = 3 .
```

```
IF (memdum = 1 AND pardum = 0 AND dondum = 1 AND voldum = 1) ORGINVINDEXUS17 = 3 .
IF (memdum = 0 AND pardum = 1 AND dondum = 1 AND voldum = 1) ORGINVINDEXUS17 = 3 .
IF (memdum = 1 AND pardum = 1 AND dondum = 1 AND voldum = 1) ORGINVINDEXUS17 = 3 .
VALUE LABELS ORGINVINDEXUS17 0 'inactive' 1 'passive' 2 'active' 3 'super-active' .
VARIABLE LABEL ORGINVINDEXUS17 'Organizational Involvement Index' .
EXECUTE .
```

5) Citizenship Norms

The citizenship battery measures the public's adherence to different potential citizenship norms. Reflecting the logic of Almond and Verba's classic *Civic Culture* study, respondents are asked how they think a 'good' citizen should behave–the perceived citizenship norms of America–rather than personal adherence to each behavior:[5]

> To be a good citizen, how important is it for a person to be . . . [list items]. 0 is extremely unimportant and 10 is extremely important.

The items in CID were selected to tap four broad dimensions of citizenship: participation, autonomy, social order and solidarity. We factor analyzed the interrelationship between items, and these analyses identified two broad dimensions of citizenship.[6] The two dimensions are measured by factor scores calculated using the regression method. There was no substitution for missing data. The N for each factor score is 978.

Variable	**Citizen Duty**	**Engaged Citizen**
Report a crime	.84	.12
Always obey the law	.77	.09
Serve in the military	.64	.15
Serve on a jury	.63	.32
Vote in elections	.56	.43
Form own opinions	.29	.47
Support worse off	.16	.65
Be active in politics	.15	.80
Active in voluntary groups	.10	.84
Eigenvalue Percent variance	2.56 28.5	2.37 25.8

Source: 2005 U.S. CID Survey
Note: the factor scores in the table above were created using the weight NATWT

[5] The intellectual roots of this battery can be traced to the 1984 General Social Survey and the 1987 Swedish Citizenship Survey that included some initial questions on the duties of citizenship. The 1998 Swedish Democracy Audit systematically studied these norms. The European "Citizenship, Involvement and Democracy" (CID) project replicated several of these items across a set of European nations in the late 1990s; and the European Social Survey (ESS) asked a subset of these items for 22 European nations in 2002 (the ESS did not include serving in the military, serving on a jury, or reporting a crime). The International Social Survey Program built upon this research tradition with a module on citizenship in 2004, and this battery is included in the 2004 General Social Survey.

[6] An unrotated analysis has all items loading positively on the first dimension, which normally occurs when a battery of items are rated on a single scale such as importance. The eigenvalues indicate two dimensions to these items (3.56, 1.30). To distinguish between different aspects of citizenship, we used a varimax rotated factor analysis. However, in an oblique solution the two dimensions are correlated at .42.

Syntax for creating these factor scores:

```
FACTOR
  /VARIABLES impsppl impvote impoblw impopin impavo impapol IMPSRVJ IMPRPCRM
IMPSRVML  /MISSING PAIRWISE /ANALYSIS impsppl
  impvote impoblw impopin impavo impapol IMPSRVJ IMPRPCRM IMPSRVML
  /PRINT INITIAL EXTRACTION ROTATION
  /CRITERIA MINEIGEN(1) ITERATE(25)
  /EXTRACTION PC
  /CRITERIA ITERATE(25)
  /ROTATION VARIMAX
  /SAVE REG(ALL)
  /METHOD=CORRELATION .
RENAME VARIABLES FAC1_1=CITDUTY .
RENAME VARIABLES FAC2_1=ENGAGCIT .
VAR LABELS CITDUTY Citizen Duty factor score .
VAR LABELS ENGAGCIT Engaged Citizen factor score .
EXECUTE .
```

6) The Diversity of Social Networks

Diversity of Close Friends **(new variable = CLOSEDIV)**

The CID survey includes various items that measure the composition of friendship networks. Although they can vary in character, friendship networks are the equivalent of strong ties. Instead of opting for questions that measure characteristics of selected friends, this survey includes items that tap into the characteristics of the entire friendship network. The most important aspect here was to make a distinction between whether respondents interact mostly in friendship networks of a *bridging* (with people who are mostly different on a variety of social dimensions) or *bonding* (people are mostly alike on a variety of social dimensions) character. When using these items below, it is perhaps important to control for the actual size of the friendship network CLOSEN; as well as for the frequency of interaction within the friendship network: CLOMEET, CLOTALK, CLOEMAIL.

Here are the items that measure the diversity of friendship networks:

```
FREQUENCIES
  VARIABLES=CLORACE CLORELIG CLOVIEWS CLOEDUC
  /ORDER=  ANALYSIS .
```

Recoding of these items can be done this way:

```
recode CLORACE CLORELIG CLOVIEWS CLOEDUC (96 thru 99=sysmis) (-1=sysmis) (else=copy)
INTO RCLORACE RCLORELIG RCLOVIEWS RCLOEDUC.
recode RCLOEDUC (1=9) (2=8) (3=7) (4=6) (5=5) (6=4) (7=3) (8=2) (9=1).
value labels RCLOEDUC 1 'None have different' 2 'Few have different' 9 ' All have different'.

variable labels RCLORACE 'close friends, about how many of them: are of a different race from yours?'.
variable labels RCLORELIG 'close friends, about how many of them: have different religious views from
yours?'.
variable labels RCLOVIEWS 'close friends, about how many of them: have different political views from
yours?'.
variable labels RCLOEDUC 'close friends, about how many of them: Have roughly a different level of
education as you'.
```

[Note: Don't Knows are coded as missing.]
[Note: The education variable had to be reversed because it asks about "sameness" not difference.]

The following procedure creates a summated rating scale to measure the diversity of friendship networks:

```
COMPUTE CLOSEDIV = (RCLORACE + RCLORELIG + RCLOVIEWS)/3 .
EXECUTE .
variable labels CLOSEDIV "Diversity of Close Friends" .
```

[Note: This scale omits RCLOEDUC, as it does not load as strongly as the other diversity measures.]
[Note: Cronbach's alpha for this scale: .56]
[Note: Low values of the SCALE indicate low diversity in friendship networks, whereas high values indicate high diversity. The scale ranges from 1-9.]

Diversity of Neighborhood Ties **(new variable = NEIGHDIV)**

The CID survey includes various items that measure the composition of neighborhood networks. Neighborhood networks are an example of weak ties, although they can vary in character. This survey includes items that tap into the characteristics of the entire neighborhood network. Once again, the most important aspect here was to make a distinction between whether respondents interact mostly in neighborhood networks of a *bridging* (with people who are mostly different on a variety of social dimensions) or *bonding* (people are mostly alike on a variety of social dimensions) character. When using these items below, it is perhaps important to control for how many neighbors the respondents knows and talks to NEIKNOW; as well as for the frequency of interaction with other neighbors: NEITALK.

Here are the items that measure the diversity of the neighborhood networks:

```
FREQUENCIES
  VARIABLES=NEICHAR1 NEICHAR2 NEICHAR3 NEICHAR4 NEICHAR5
  /ORDER=  ANALYSIS
```

Recoding of these items can be done this way:

```
recode NEICHAR1 NEICHAR2 NEICHAR3 NEICHAR4 NEICHAR5 (96 thru 99=sysmis) (-1=sysmis)
(else=copy) INTO NEIRACE NEIRELIG NEIPOLI NEIEDU NEIMMI.
recode NEIEDU (1=9) (2=8) (3=7) (4=6) (5=5) (6=4) (7=3) (8=2) (9=1).
value labels NEIEDU 1 'None have different' 2 'Few have different' 9 ' All have different'.

variable labels NEIRACE 'neighbors R talks to, about how many of them: are of a different race from
yours?'.
variable labels NEIRELIG 'neighbors R talks to, about how many of them: have different religious views
from yours?'.
variable labels NEIPOLI 'neighbors R talks to, about how many of them: have different political views
from yours?'.
variable labels NEIEDU 'neighbors R talks to, about how many of them: Have roughly a different level of
education as you'.
variable labels NEIMMI 'neighbors R talks to, about how many of them: Are recent immigrants to the
United States'.
```

[Note: Don't Knows are coded as missing.]
[Note: The education variable had to be reversed because it asks about "sameness" not difference.]

The following procedure creates a summated rating scale to measure the diversity of neighborhood networks:

```
COMPUTE NEIGHDIV = (NEIRACE + NEIRELIG + NEIPOLI + NEIMMI)/4 .
EXECUTE .
variable labels NEIGHDIV "Diversity of Neighbors with whom R talks" .
```

[Note: Cronbach's alpha for this scale: .84]
[Note: Low values of the scale indicate low diversity in neighborhood networks, whereas high values indicate high diversity. The scale ranges from 1-9.]
[Note: This scale is ideal for non-immigrant respondents (because of the item that measures the number of immigrants in R's neighborhood).]

Diversity of Ties at Work, School, or Former Workplace (new variable = WORKDIV)

The CID survey also includes various items that measure the composition of co-worker networks. Co-worker networks are another example of weak ties, although again they can vary in character. This survey includes items that tap into the characteristics of the group of co-workers with whom the respondent interacts at work. Once again, the most important aspect here was to make a distinction between whether respondents interact mostly with co-workers who are of different backgrounds than the respondent (*bridging ties)* or with people who are like the respondent on various social dimensions (*bonding ties)*. When using these items below, it is perhaps important to control for how the respondent socializes with his/her co-workers (COWSOC), though this would preclude the analysis to those respondents who are currently working.

There are three sets of variables that measure the diversity of the co-workers with whom the respondent interacts (for R at work; and for R at school; and for R's past work in case R does not work anymore). Below please find the syntax for collapsing the three into one variable:

```
FREQUENCIES
  VARIABLES= SCHCHAR1 SCHCHAR2 SCHCHAR3 SCHCHAR4 SCHCHAR5 COWCHAR1
COWCHAR2 COWCHAR3 COWCHAR4 COWCHAR5 JOBCHAR1 JOBCHAR2
JOBCHAR3 JOBCHAR4 JOBCHAR5 /ORDER=  ANALYSIS .
```

Recoding of these items can be done as follows:

```
recode SCHCHAR1 SCHCHAR2 SCHCHAR3 SCHCHAR4 SCHCHAR5 COWCHAR1 COWCHAR2
COWCHAR3 COWCHAR4 COWCHAR5 JOBCHAR1 JOBCHAR2 JOBCHAR3 JOBCHAR4
JOBCHAR5 (96 thru 99=sysmis) (-1=sysmis) (else=copy) INTO Schoolrace Schoolrelig Schoolpoli
Schooledu
Schoolimmi Workrace Workrelig Workpoli Workedu Workimmi Oldrace Oldrelig Oldpoli Oldedu
Oldimmi .
recode Schooledu (1=9) (2=8) (3=7) (4=6) (5=5) (6=4) (7=3) (8=2) (9=1).
recode Workedu (1=9) (2=8) (3=7) (4=6) (5=5) (6=4) (7=3) (8=2) (9=1).
recode Oldedu (1=9) (2=8) (3=7) (4=6) (5=5) (6=4) (7=3) (8=2) (9=1).
```

[Note: Don't Knows are coded as missing.]
[Note: The education variable had to be reversed because it asks about "sameness" not difference.]

The following procedure collapses the current school, current and old workplace variables:

```
IF (workrace >= 1) fworkrace = workrace .
EXECUTE .
IF (schoolrace>=1) fworkrace = schoolrace.
EXECUTE .
IF (oldrace>=1) fworkrace = oldrace.
EXECUTE .
IF (workrelig >= 1) fworkrelig = workrelig .
EXECUTE .
IF (schoolrelig>=1) fworkrelig = schoolrelig .
EXECUTE .
IF (oldrelig>=1) fworkrelig = oldrelig .
EXECUTE .
IF (workpoli >= 1) fworkpoli = workpoli .
```

```
EXECUTE .
IF (schoolpoli>=1) fworkpoli = schoolpoli.
EXECUTE .
IF (oldpoli>=1) fworkpoli = oldpoli.
EXECUTE .
IF (workedu >= 1) fworkedu = workedu .
EXECUTE .
IF (schooledu>=1) fworkedu = schooledu.
EXECUTE .
IF (oldedu>=1) fworkedu = oldedu.
EXECUTE .
IF (workimmi >= 1) fworkimmi = workimmi .
EXECUTE .
IF (schoolimmi>=1) fworkimmi = schoolimmi.
EXECUTE .
IF (oldimmi>=1) fworkimmi = oldimmi.
EXECUTE .
```

Variable Labels:

```
variable labels FWORKRACE ‘co-workers R talks to, about how many of them: are of a different race from yours?’.
variable labels FWORKRELIG ‘co-workers R talks to, about how many of them: have different religious views from yours?’.
variable labels FWORKPOLI ‘co-workers R talks to, about how many of them: have different political views from yours?’.
variable labels FWORKEDU ‘co-workers R talks to, about how many of them: Have roughly a different level of education than you’.
variable labels FWORKIMMI ‘co-workers R talks to, about how many of them: Are recent immigrants to the United States’.
```

The following procedure creates a summated rating scale to measure the diversity of co-workers:

```
COMPUTE WORKDIV = (fworkrace + fworkrelig + fworkpoli + fworkimmi)/4 .
EXECUTE .
variable labels WORKDIV “Diversity of co-workers with whom R interacts” .
```

[Note: Cronbach’s alpha for this scale: .70]
[Note: Low values of the scale indicate low diversity among co-workers, whereas high values indicate high diversity. The scale ranges from 1-9.]
[Note: This scale is ideal for non-immigrant respondents (because of the item that measures the number of immigrants at R’s work).]

Summary Diversity Scales **(see new variable names below)**

Strong Tie scale: CLOSEDIV

Weak Tie scale: WEAKDIV

COMPUTE WEAKDIV= (NEIGHDIV+WORKDIV)/2
variable labels WEAKDIV “Diversity of weak ties (neighbors and co-workers)” .

Combined Tie Scale: Friendscale+Neighborscale+ Workscale

New Neighborscale:

COMPUTE NEIGHDIV2 = (NEIRACE + NEIRELIG + NEIPOLI)/3.
EXECUTE .

New Workscale:

COMPUTE WORKDIV2 = (Fworkrace + Fworkrelig + Fworkpoli)/3.
EXECUTE .

[Note: this scale only combines items that have been asked for all arenas.]

Overall Diversity Scale: DIVERSITY

Compute DIVERSITY=(CLOSEDIV+NEIGHDIV2+WORKDIV2)/3
variable labels DIVERSITY “Diversity of ALL ties (close friends, neighbors, and co-workers)” .

7) Racial Prejudice/Negative Stereotypes

(new variables = PREJASN2, PREJBLK2, PREJHIS2, PREJWHT2)

This set of variables measures R's level of prejudice against other racial groups. Each index is based on the questions on negative stereotypes (i.e., hard to understand, cannot be friend, prone to crime, cannot trust, selfish, and lazy). The new variables are constructed for the cases with two and more valid values among these questions. Higher values indicate higher prejudice.

```
COMPUTE rustanda = -1.
DO IF (rustanda EQ -1).
IF (gq129a EQ 1) rustanda = rustand1.
IF (gq129b EQ 1) rustanda = rustand2.
IF (gq129c EQ 1) rustanda = rustand3.
END IF.
VAR LABELS rustanda hard to understand ASIANS.
EXECUTE.
RECODE rustanda (-1 = 10).
missing values rustanda (7 thru 10).
EXECUTE.

COMPUTE rustandb = -1.
DO IF (rustandb EQ -1).
IF (gq129a EQ 2) rustandb = rustand1.
IF (gq129b EQ 2) rustandb = rustand2.
IF (gq129c EQ 2) rustandb = rustand3.
END IF.
VAR LABELS rustandb hard to understand BLACKS.
EXECUTE.
RECODE rustandb (-1 = 10).
missing values rustandb (7 thru 10).
EXECUTE.

COMPUTE rustandh = -1.
DO IF (rustandh EQ -1).
IF (gq129a EQ 3) rustandh = rustand1.
IF (gq129b EQ 3) rustandh = rustand2.
IF (gq129c EQ 3) rustandh = rustand3.
END IF.
VAR LABELS rustandh hard to understand HISPANICS.
EXECUTE.
RECODE rustandh (-1 = 10).
missing values rustandh (7 thru 10).
EXECUTE.

COMPUTE rustandw = -1.
DO IF (rustandw EQ -1).
IF (gq129a EQ 4) rustandw = rustand1.
IF (gq129b EQ 4) rustandw = rustand2.
IF (gq129c EQ 4) rustandw = rustand3.
END IF.
```

```
VAR LABELS rustandw hard to understand WHITES.
EXECUTE.
RECODE rustandw (-1 = 10).
missing values rustandw (7 thru 10).
EXECUTE.

COMPUTE rfrienda = -1.
DO IF (rfrienda EQ -1).
IF (gq129a EQ 1) rfrienda = rfriend1.
IF (gq129b EQ 1) rfrienda = rfriend2.
IF (gq129c EQ 1) rfrienda = rfriend3.
END IF.
VAR LABELS rfrienda cannot be friend with ASIANS.
EXECUTE.
RECODE rfrienda (-1 = 10).
missing values rfrienda (7 thru 10).
EXECUTE.

COMPUTE rfriendb = -1.
DO IF (rfriendb EQ -1).
IF (gq129a EQ 2) rfriendb = rfriend1.
IF (gq129b EQ 2) rfriendb = rfriend2.
IF (gq129c EQ 2) rfriendb = rfriend3.
END IF.
VAR LABELS rfriendb cannot be friend with BLACKS.
EXECUTE.
RECODE rfriendb (-1 = 10).
missing values rfriendb (7 thru 10).
EXECUTE.

COMPUTE rfriendh = -1.
DO IF (rfriendh EQ -1).
IF (gq129a EQ 3) rfriendh = rfriend1.
IF (gq129b EQ 3) rfriendh = rfriend2.
IF (gq129c EQ 3) rfriendh = rfriend3.
END IF.
VAR LABELS rfriendh cannot be friend with HISPANICS.
EXECUTE.
RECODE rfriendh (-1 = 10).
missing values rfriendh (7 thru 10).
EXECUTE.

COMPUTE rfriendw = -1.
DO IF (rfriendw EQ -1).
IF (gq129a EQ 4) rfriendw = rfriend1.
IF (gq129b EQ 4) rfriendw = rfriend2.
IF (gq129c EQ 4) rfriendw = rfriend3.
END IF.
VAR LABELS rfriendw cannot be friend with WHITES.
EXECUTE.
RECODE rfriendw (-1 = 10).
```

```
missing values rfriendw (7 thru 10).
EXECUTE.

COMPUTE rcrimea = -1.
DO IF (rcrimea EQ -1).
IF (gq129a EQ 1) rcrimea = rcrime1.
IF (gq129b EQ 1) rcrimea = rcrime2.
IF (gq129c EQ 1) rcrimea = rcrime3.
END IF.
VAR LABELS rcrimea prone to crime ASIANS.
EXECUTE.
RECODE rcrimea (-1 = 10).
missing values rcrimea (7 thru 10).
EXECUTE.

COMPUTE rcrimeb = -1.
DO IF (rcrimeb EQ -1).
IF (gq129a EQ 2) rcrimeb = rcrime1.
IF (gq129b EQ 2) rcrimeb = rcrime2.
IF (gq129c EQ 2) rcrimeb = rcrime3.
END IF.
VAR LABELS rcrimeb prone to crime BLACKS.
EXECUTE.
RECODE rcrimeb (-1 = 10).
missing values rcrimeb (7 thru 10).
EXECUTE.

COMPUTE rcrimeh = -1.
DO IF (rcrimeh EQ -1).
IF (gq129a EQ 3) rcrimeh = rcrime1.
IF (gq129b EQ 3) rcrimeh = rcrime2.
IF (gq129c EQ 3) rcrimeh = rcrime3.
END IF.
VAR LABELS rcrimeh prone to crime HISPANICS.
EXECUTE.
RECODE rcrimeh (-1 = 10).
missing values rcrimeh (7 thru 10).
EXECUTE.

COMPUTE rcrimew = -1.
DO IF (rcrimew EQ -1).
IF (gq129a EQ 4) rcrimew = rcrime1.
IF (gq129b EQ 4) rcrimew = rcrime2.
IF (gq129c EQ 4) rcrimew = rcrime3.
END IF.
VAR LABELS rcrimew prone to crime WHITES.
EXECUTE.
RECODE rcrimew (-1 = 10).
missing values rcrimew (7 thru 10).
EXECUTE.
```

```
COMPUTE rutrusta = -1.
DO IF (rutrusta EQ -1).
IF (gq129a EQ 1) rutrusta = rutrust1.
IF (gq129b EQ 1) rutrusta = rutrust2.
IF (gq129c EQ 1) rutrusta = rutrust3.
END IF.
VAR LABELS rutrusta cannot trust ASIANS.
EXECUTE.
RECODE rutrusta (-1 = 10).
missing values rutrusta (7 thru 10).
EXECUTE.

COMPUTE rutrustb = -1.
DO IF (rutrustb EQ -1).
IF (gq129a EQ 2) rutrustb = rutrust1.
IF (gq129b EQ 2) rutrustb = rutrust2.
IF (gq129c EQ 2) rutrustb = rutrust3.
END IF.
VAR LABELS rutrustb cannot trust BLACKS.
EXECUTE.
RECODE rutrustb (-1 = 10).
missing values rutrustb (7 thru 10).
EXECUTE.

COMPUTE rutrusth = -1.
DO IF (rutrusth EQ -1).
IF (gq129a EQ 3) rutrusth = rutrust1.
IF (gq129b EQ 3) rutrusth = rutrust2.
IF (gq129c EQ 3) rutrusth = rutrust3.
END IF.
VAR LABELS rutrusth cannot trust HISPANICS.
EXECUTE.
RECODE rutrusth (-1 = 10).
missing values rutrusth (7 thru 10).
EXECUTE.

COMPUTE rutrustw = -1.
DO IF (rutrustw EQ -1).
IF (gq129a EQ 4) rutrustw = rutrust1.
IF (gq129b EQ 4) rutrustw = rutrust2.
IF (gq129c EQ 4) rutrustw = rutrust3.
END IF.
VAR LABELS rutrustw cannot trust WHITES.
EXECUTE.
RECODE rutrustw (-1 = 10).
missing values rutrustw (7 thru 10).
EXECUTE.

COMPUTE rselfa = -1.
DO IF (rselfa EQ -1).
IF (gq129a EQ 1) rselfa = rself1.
```

```
IF (gq129b EQ 1) rselfa = rself2.
IF (gq129c EQ 1) rselfa = rself3.
END IF.
VAR LABELS rselfa be selfish ASIANS.
EXECUTE.
RECODE rselfa (-1 = 10).
missing values rselfa (7 thru 10).
EXECUTE.

COMPUTE rselfb = -1.
DO IF (rselfb EQ -1).
IF (gq129a EQ 2) rselfb = rself1.
IF (gq129b EQ 2) rselfb = rself2.
IF (gq129c EQ 2) rselfb = rself3.
END IF.
VAR LABELS rselfb be selfish BLACKS.
EXECUTE.
RECODE rselfb (-1 = 10).
missing values rselfb (7 thru 10).
EXECUTE.

COMPUTE rselfh = -1.
DO IF (rselfh EQ -1).
IF (gq129a EQ 3) rselfh = rself1.
IF (gq129b EQ 3) rselfh = rself2.
IF (gq129c EQ 3) rselfh = rself3.
END IF.
VAR LABELS rselfh be selfish HISPANICS.
EXECUTE.
RECODE rselfh (-1 = 10).
missing values rselfh (7 thru 10).
EXECUTE.

COMPUTE rselfw = -1.
DO IF (rselfw EQ -1).
IF (gq129a EQ 4) rselfw = rself1.
IF (gq129b EQ 4) rselfw = rself2.
IF (gq129c EQ 4) rselfw = rself3.
END IF.
VAR LABELS rselfw be selfish WHITES.
EXECUTE.
RECODE rselfw (-1 = 10).
missing values rselfw (7 thru 10).
EXECUTE.

COMPUTE rlazya = -1.
DO IF (rlazya EQ -1).
IF (gq129a EQ 1) rlazya = rlazy1.
IF (gq129b EQ 1) rlazya = rlazy2.
IF (gq129c EQ 1) rlazya = rlazy3.
END IF.
```

```
VAR LABELS rlazya be lazy ASIANS.
EXECUTE.
RECODE rlazya (-1 = 10).
missing values rlazya (7 thru 10).
EXECUTE.

COMPUTE rlazyb = -1.
DO IF (rlazyb EQ -1).
IF (gq129a EQ 2) rlazyb = rlazy1.
IF (gq129b EQ 2) rlazyb = rlazy2.
IF (gq129c EQ 2) rlazyb = rlazy3.
END IF.
VAR LABELS rlazyb be lazy BLACKS.
EXECUTE.
RECODE rlazyb (-1 = 10).
missing values rlazyb (7 thru 10).
EXECUTE.

COMPUTE rlazyh = -1.
DO IF (rlazyh EQ -1).
IF (gq129a EQ 3) rlazyh = rlazy1.
IF (gq129b EQ 3) rlazyh = rlazy2.
IF (gq129c EQ 3) rlazyh = rlazy3.
END IF.
VAR LABELS rlazyh be lazy HISPANICS.
EXECUTE.
RECODE rlazyh (-1 = 10).
missing values rlazyh (7 thru 10).
EXECUTE.

COMPUTE rlazyw = -1.
DO IF (rlazyw EQ -1).
IF (gq129a EQ 4) rlazyw = rlazy1.
IF (gq129b EQ 4) rlazyw = rlazy2.
IF (gq129c EQ 4) rlazyw = rlazy3.
END IF.
VAR LABELS rlazyw be lazy WHITES.
EXECUTE.
RECODE rlazyw (-1 = 10).
missing values rlazyw (7 thru 10).
EXECUTE.

COMPUTE prejasn = MEAN.2(rustanda, rfrienda, rcrimea, rutrusta, rselfa, rlazya).
COMPUTE prejblk = MEAN.2(rustandb, rfriendb, rcrimeb, rutrustb, rselfb, rlazyb).
COMPUTE prejhis = MEAN.2(rustandh, rfriendh, rcrimeh, rutrusth, rselfh, rlazyh).
COMPUTE prejwht = MEAN.2(rustandw, rfriendw, rcrimew, rutrustw, rselfw, rlazyw).

COMPUTE PREJASN2 = 6 - prejasn.
COMPUTE PREJBLK2 = 6 - prejblk.
COMPUTE PREJHIS2 = 6 - prejhis.
COMPUTE PREJWHT2 = 6 - prejwht.
```

```
VAR LABELS PREJASN2 Prejudice against ASIANS.
VAR LABELS PREJBLK2 Prejudice against BLACKS.
VAR LABELS PREJHIS2 Prejudice against HISPANICS.
VAR LABELS PREJWHT2 Prejudice against WHITES.
EXECUTE.
```

The final prejudice variables (PREJASN2, PREJBLK2, PREJHIS2, and PREJWHT2) correspond to the average level of agreement with racist stereotypes against each particular group (with higher levels indicating greater levels of prejudice).

8) National Pride

(new variable = NATPRIDE)

This index is based on two questions, NATIO1 and NATIO2. Higher values indicate higher level of national pride.

```
RECODE natio1 (5=1) (4=2) (2=4) (1=5).
RECODE natio2 (5=1) (4=2) (2=4) (1=5).
COMPUTE NATPRIDE = MEAN.1(natio1, natio2).
VAR LABELS NATPRIDE National Pride.
EXECUTE.
```

9) Attitudes toward Immigrants

(new variable = IMMATT)

This index, measuring R's subjective perceptions on the impact of immigrants, is based on three variables, IMBLECO, IMUECLT, and IMWBCRM. The index is constructed with the cases that have just one valid value among these items. Higher values denote positive attitudes toward immigrants.

```
COMPUTE IMMATT = MEAN.1(imbleco, imueclt, imwbcrm).
VAR LABELS IMMATT Attitudes toward Immigrants.
EXECUTE.
```

10) Respondent Demographics

The U.S. CID data includes numerous demographic questions. For ease and consistency of analysis, we recommend using the following constructed variables on the main demographic themes of race, income, education, age, and gender:

Race (new variable = RRACE)

The questionnaire included several questions on respondents' racial/ethnic identification. While recognizing that some people identify with more than one race, our goal was to yield a single race category for analysis, and to avoid the "Other" category if possible.

The first race question simply asked the respondent to select his/her race (variables RACEME1-8). Multiple responses were allowed if the respondent chose two or more. In that case, the respondent was asked to choose one of the options that *best* describes his/her race (RACEMEB). Finally, if the respondent still did not choose a race, the interviewer then chose the "best fitting respondent race" (RACEMEI).

The syntax below creates the variable RRACE, which includes 45 Asians, 111 Blacks, 124 Hispanics, 706 Whites, and 14 "Other."

```
COUNT RACEMEN=RACEME1 TO RACEME5(1) .
VAR LABELS RACEMEN NUMBER OF RACES THE RESPONDENT CLAIMS .
COMPUTE RRACE = -1 .
VAR LABELS RRACE RESPONDENT'S RACE, SELF CLASSIFICATION .
DO IF (RACEMEN EQ 1) .
IF (RACEME1 EQ 1)RRACE =1 .
IF (RACEME2 EQ 1)RRACE =2 .
IF (RACEME3 EQ 1)RRACE =3 .
IF (RACEME4 EQ 1)RRACE =4 .
IF (RACEME5 EQ 1)RRACE =5 .
END IF .
VALUE LABELS RRACE (1)Asian(2)Black(3)Hispanic(4)White(5)Other .
FREQUENCIES VARI = RRACE .
DO IF (RRACE EQ -1) .
IF (RACEMEB EQ 2)RRACE = 2  .
IF (RACEMEB EQ 3)RRACE = 3  .
IF (RACEMEB EQ 4)RRACE = 4  .
IF (RACEMEB EQ 6)RRACE = 5  .
END IF .
FREQUENCIES VARI = RRACE .
DO IF (RRACE EQ -1) .
COMPUTE RRACE = RACEOBS .
END IF .
FREQUENCIES VARI=RACEMEN RRACE .
RECODE RACEOBS (9=4) .
FREQUENCIES VARI = RACEOBS .
EXECUTE .
```

Income **(new variable = INCOMEC)**

[Note: The following procedure was derived using the weight NATWT. Users who use dweight or NATWTB should be cautious in applying the INCOMEC variable.]

Since the main income variable (hinctnt) has a great deal of missing data (95 Refused, 29 Don't Know, 10 No Answer), we have created a new variable (INCOMEC, Corrected, Adjusted Income), which is based on imputations from a combination of home ownership (OWNHOME), the interviewer's estimate of the respondent's social class (RCLASS), the respondent's response to a question about "how you feel about your household's income nowadays" (HINCFEL), and a follow-up question about whether the respondent's income is "under $50,000 or $50,000 or more" (INCOME50). The equation is strong, and its use will reduce the missing cases to 18.

We used the following syntax to create the INCOMEC variable:

```
MISSING VALUES HINCTNT () .
MISSING VALUES HINCFEL (7,8,9) .
missing values OWNHOME (7 thru 9) .
missing values RCLASS (9) .
COMPUTE INCOME = HINCTNT .
VAR LABELS INCOME HOUSEHOLD INCOME .
VALUE LABELS INCOME(1)LESS THAN $15,000
  (11)$200,000 OR MORE
  (97)REFUSED(98)DON'T KNOW (99)NO ANSWER .
COMPUTE INCOMEC1 = 10.932 - (1.115 * RCLASS)
   - (1.063 * OWNHOME) - (.895 * HINCFEL ) .
IF (INCOME LE 11)INCOMEC1 = 0 .
MISSING VALUES INCOMEC1 (-1) .
COMPUTE INCOMEC2 = RND(INCOMEC1) .
COMPUTE INCOME4 = INCOME .
*MISSING VALUES INCOME(97,98,99) .
RECODE INCOME4 (97,98,99=0) .
COMPUTE INCOMEC3 = INCOME4 + INCOMEC2 .
IF (INCOME50 EQ 1) AND (INCOMEC3 EQ 7)INCOMEC3 = 6 .
IF (INCOME50 EQ 2) AND (INCOMEC3 EQ 5)INCOMEC3 = 6 .
COMPUTE INCOMEC = INCOMEC3 .
VAR LABELS INCOMEC CORRECTED, ADJUSTED INCOME .
VALUE LABELS INCOMEC(1)LESS THAN $15,000
  (11)$200,000 OR MORE .
EXECUTE .
```

Education **(new variables = EDUC2 and SPEDUC2)**

We recommend using a modified version of the education variables for both the respondent and spouse (educ and edulvlp), which we have called EDUC2 and SPEDUC2. The new variables simply merge together the categories "high school graduate" and "GED," and they recode the scale to begin at zero. The following syntax was used to create these variables:

```
COMPUTE EDUC2 = EDUC .
COMPUTE SPEDUC2 = EDULVLP .
```

```
VAR LABELS EDUC2 RESPONDENT HIGHEST GRADE COMPLETED, RECODED .
VAR LABELS SPEDUC2 SPOUSE HIGHEST GRADE COMPLETED, RECODED .
missing values EDUC2 SPEDUC2 (97 thru 99) .
RECODE EDUC2 SPEDUC2
  (1=0)(2=1)(3,4=2)(5=3)(6=4)(7=5)(8=6) .
VALUE LABELS EDUC2 SPEDUC2
  (0)None, or grade 1-8
  (1)High school incomplete--grades 9 - 11
  (2)High school graduate or GED
  (3)Business, technical, or vocational school after high school
  (4)Some college, no 4-year degree
  (5)College graduate
  (6)Post-graduate training, professional schooling after college .
EXECUTE .
```

***Age* (new variable = AGE)**

Since the questionnaire asks the respondent to name his/her year of birth, we have constructed a new variable for age that is a simple subtraction of the year or birth from 2005, the year of the survey. The following syntax was used:

```
missing values yrbrn (9997 thru 9999) .
COMPUTE AGE = 2005 - yrbrn .
VARIABLE LABELS AGE 'AGE IN YEARS AS OF 2005' .
EXECUTE .
```

***Gender* (new variable = GENDER2)**

For those who prefer using a gender variable that has male as the high value, we have created the variable GENDER2 with the following syntax:

```
RECODE GNDR (2=0) (1=1) INTO GENDER2 .
ADD VALUE LABELS GENDER2(0)Female (1)Male .
VAR LABELS GENDER2 RESPONDENT GENDER WITH HI = MALE .
EXECUTE .
```