Problem Set #6: Social Capital and Good Government in Italy

Probably one of the most influential ideas in political science in the last decade has been the concept of social capital. Social capital refers to characteristics of social organization among two or more people. Some relationships are more productive than others; conceptually, the more productive a relationship is, the more social capital is embodied in the relationship. Though authors have identified multiple attributes of relationships that can increase their productivity, the most widely cited attributes are those that have to do with trust, norms of reciprocity, and the nature of networks between people. Social capital is an investment good: it takes time and money to establish relationships, imbue them with trust, and build up a reputation for reciprocity. An example of social capital at work is when a friend agrees to take care of a child in an emergency. I’m willing to leave my child with a friend because I trust the friend to do the right thing. Moreover, if my friend is embedded in a larger group of mutual acquaintances, there is an implicit threat that I will “bad-mouth” the friend to other acquaintances in the group if he or she fails to take good care of the child. The friend is willing to take on the burden (at least in part) because there is an expectation that I would do the same later for them. If I fail to reciprocate when asked later, the friend can “bad-mouth” me to our mutual acquaintances, making my investment in relationship with those acquaintances less useful later on.

Robert D. Putnam, a political scientist and former Dean of Harvard’s Kennedy School of Government, has popularized the idea that “good governance” depends on the amount of social capital in a given political unit. When citizens know each other, trust each other, and are willing to work together, they are also more likely to successfully organize for political action, pay more attention to what their government is doing, and work together to demand better governance from their elected and appointed officials. Putnam’s seminal work on this subject is *Making Democracy Work: Civic Traditions in Modern Italy* (MDW). MDW leverages data collected from 1970 to 1985 during a federally-mandated reorganization of Italy’s regional governments. Putnam’s goal was to discover what institutional, political, and social factors tended to lead to greater satisfaction with the reformed governments and greater institutional performance. This study constitutes a “natural experiment.” The structure of the new regional governments was imposed from above; differences in performance are not due to differences in structure; they must be due to differences in community.

While Putnam’s theoretical model is quite elegant, it is also quite abstract. Simply put, there is no way to directly measure “social capital,” “citizen satisfaction,” “institutional performance,” or other concepts that are present in his model. Instead, Putnam leveraged multiple measures of these and other concepts and then attempted to construct indices using factor analysis. In this problem set, we will attempt to build similar indices from his original data.

*Data courtesy of Professor Robert D. Putnam*

*John F. Kennedy School of Government, Harvard University.*
Part I: Data Construction and Familiarization

1. Spend a few minutes becoming familiar with the data and looking for any data issues.

2. Review chapters 3, 4, and 5 of MDW (available through ERes). The basic Putnam argument is:
   a. Citizen satisfaction is correlated with institutional performance.
   b. Institutional performance is correlated with “civic community.”
   c. Civic community is related to other constructs, like “clientelism,” (negatively) “electoral reform,” (positively) “republicanism,” (positively) and “clericalism” (negatively).
   d. Civic traditions are related to civic community and performance.

In short, citizens are more satisfied when their government performs well, and good performance seems to be predicted by strong civic community and a long history of civic traditions. However, all of these are unobservable constructs. We will try to reconstruct some of these and see how the relationships hold up.

Part I: Data Analysis

1. Let’s construct several indices of citizen satisfaction. The dataset includes the mean score for each region. For each part, use v1251, v1252, v1253, v1254, v1255, and v1481 as the variables.
   a. Construct an index using principal components analysis. First run the `factor` command with the `pc` option and then run the `score` command (giving the score some name like `sat_pc`). For example:

```
pca v1251 v1252 v1253 v1254 v1255 v1481
predict sat_pc
```
   b. Create a scatterplot graph of performance in 1970-1976 (`v917`) against `sat_pc`. What relationship does there appear to be between performance and satisfaction? Use the `mlabel(var1RegC)` option to show the points as the two-letter abbreviations for the regions. For instance:

```
twoway (scatter v917 sat_pc, mlabel(var1RegC))
```
   c. Notice carefully the scoring used for the satisfaction variables. Higher numbers represent lower satisfaction. Our first graph is somewhat confusing, so we need to transform the original data into a more appropriate index. One way is to reformat all six variables by indexing them, with 0 being assigned to the lowest value of the variable and 100 being assigned to the highest value. Using the `sum` command, find the min and max for each variable. Then use the following transformation: (max – actual)/(max-min). For instance, `v1251` has a max of 3.313 and a min of 2.441. To recode, use this formula:

```
gen v1251REC = (3.313 - v1251)/(3.313-2.441)
```
Recode the remaining variables (if you know how to use DO files in Stata, this will be faster), re-run the principal components analysis (name the new score sat_pc2), and then re-run the scatterplot. Does this look more reasonable?

2. Now let’s run a principal factors analysis using the pf option on the factor command.
   a. What are the primary differences between the assumptions made in principal component and principal factor analysis? (Hint: there are two)
   b. Run the following commands:

   ```stata
   factor v1251REC v1252REC v1253REC v1254REC v1255REC v1481REC, pf
   predict sat_pf
   twoway (scatter sat_pf v917, mlabel(var1RegC))
   ```

   Compare your PF graph to the PC graph in 1c. Are there any differences? If so, how are they different?

   c. Now re-run the factor analysis in 2b, but this time do a varimax rotation after running factor but before doing the score by using the command rotate, varimax. Name this score sat_pf2 and create another scatterplot graph. Is this graph different than those found in 1c and 2b? If so, how?

   d. Of the graphs run in 1b, 1c, 2b, and 2c, which is preferable in your opinion? In your answer, be sure to discuss whether PF or PC is more appropriate and whether the rotation was appropriate. (Hint: think about the difference in assumptions between the two and how one might check whether the PC assumption is met.)

3. Now let’s pretend we have no a priori knowledge of the factor structure. Run a principal factors analysis on the variables listed as part of the “Clericalism” (excluding v1654 and v1655) and “Satisfaction” indices (for Satisfaction, use the recoded variables we constructed in question 1).
   a. Use the greigen command to graph the eigenvalues. How many factors does this test suggest should be retained? How can you tell?
   b. Create scores for the first two factors in the analysis of 3a and scatterplot both factors against performance 1970-79 (v917). To get scores for the first two factors, just include two variables names in the predict statement:

   ```stata
   predict clerical sat_pf3
   ```

   c. Now run the principal factors analysis again. This time, run a varimax rotation before scoring. Then create new scatterplots against performance 1970-79 (v917).

   d. Do you think the rotation helped to define the nature of the factors better? Compare the graphs created in 3b and 3c. Are there major differences in scoring? Which is easier to interpret regarding the relationship between performance and clericalism or satisfaction? Is the satisfaction versus performance graph similar to what you found in 1c?
4. The central thrust of Putnam’s argument revolves around the relationship between civic community and institutional performance. Let’s re-create his findings.

   a. Using the Institutional Performance 1978-1985 variables, do a principal components factors analysis and then create an index for each using the `predict` command.

   b. The Civic Community index is a little more complex to create. There are two sets of variables which have to be factor analyzed before we do a factor analysis for the overall index. The first step is to do a principal components analysis of Referendum Turnout 1974-1987 (variables v1491, v1492, v1493, v1494, and v1495) and Preference Voting 1953-1979 (variables v791, v792, v793, v802, v1489, and v1488). Create scores for both Referendum Turnout and Preference Voting. Then run a principal components analysis of these new variables, plus v712, v1661, v1235, and v1660. This is the Civic Community index. You must now do one additional step. Because all of the variables have an inverse relationship to the index we wish to create (i.e., smaller values mean more “civic community”), we need to multiply the Civic Community index by negative 1 (i.e., -1) in order to have a bigger score mean more Civic Community (either generate a new variable or use the `replace` command).

   c. Scatterplot the Institutional Performance and Civic Community indices. Does the data seem to support Putnam’s hypothesis? Run a regression with the Institutional Performance index as the dependent variable. How well does the regression line fit this data?

   d. The data set includes Putnam’s calculated indices for Institutional Performance (v1691) and Civic Community (v1662). Use the `pwcorr` command with the “sig” option to find the pair-wise correlation and level of statistical significance in those correlations between Putnam’s measures and yours.

5. Does your re-analysis of Putnam’s data tend to confirm or disconfirm his findings?
Data Description

soc_capital.dta

The dataset that forms the backbone of Making Democracy Work was collected over a 20-year period using multiple methods. The project was inspired by a unique opportunity to observe the birth of a new level of government. In 1970, the government of Italy created fifteen new regional governments simultaneously. Though Italy’s constitution of 1948 mandated regional governments, none were created until political pressure began to build for government reform in the late 1960s. Each regional government was mandated to have the same structure and organization. Thus, the quality and nature of the governance that resulted in each region would not be confounded with differences in structure. Only differences in the social and political communities themselves could account for differences in institutional performance.

The data came from six sources: a) four waves of personal interviews with regional councilors in six selected regions; b) three waves of personal interviews with community leaders in six selected regions; c) six specially commissioned national surveys, as well as several dozen other regional surveys between 1968 and 1988; d) an experiment in 1983 to test government responsiveness to street-level citizen inquiries in all regions; e) case studies of institutional politics and regional planning in the six selected regions between 1976 and 1989; and f) detailed analysis of legislation produced by all regions between 1970 and 1984. This dataset contains region averages and aggregates that are drawn from these and other archival data sources.

List of Variables, Ordered by Variable Number

| v1     | Region                                              |
| v30    | Church marriages, 1976                             |
| v690   | Anti-divorce vote, 1974                            |
| v710   | Divorce rate, 1973                                 |
| v712   | Newspaper Readership, c. 1975                      |
| v791   | Preference voting (%), 1953                        |
| v792   | Preference voting (%), 1958                        |
| v793   | Preference voting (%), 1963                        |
| v802   | Preference voting (%), 1976                        |
| v864   | National Electoral Turnout (%) 1919                |
| v865   | National Electoral Turnout (%) 1921                |
| v882   | Communal Electoral Turnout (%) 1920                |
| v885   | Provincial Electoral Turnout (%) 1920              |
| v890   | Socialist vote (%) 1919                           |
| v891   | Catholic vote (%) 1919                            |
| v892   | Socialist vote (%) 1921                           |
| v893   | Catholic vote (%) 1921                            |
| v896   | Socialist councils (%) 1920                       |
| v897   | Catholic councils (%) 1920                        |
| v917   | Institutional Performance, 1970-1976 (Factan Index)|
| v925   | Electoral turnout, 1919-1921 (Factan Index)        |
| v1010  | Mutual aid societies (per 100,000 pop), 1873       |
v1011 Mutual aid societies (per 100,000 pop), 1878
v1012 Mutual aid societies (per 100,000 pop), 1885
v1013 Mutual aid societies (per 100,000 pop), 1895
v1033 Mutual aid societies, 1873-1904 (Factan Index)
v1061 Mutual aid societies (per 100,000 pop), 1904
v1105 Strength of mass parties, 1919-1921
  (Factan Index of [v890+v891], [v892+v893], and [v896+v897])
v1235 Inhabitants per cultural/recreational assn.
v1241 Associations founded before 1860 (%)
v1251 Mean public satisfaction with regional govt, 1977
v1252 Mean public satisfaction with regional govt, 1979
v1253 Mean public satisfaction with regional govt, 1981
v1254 Mean public satisfaction with regional govt, 1982
v1255 Mean public satisfaction with regional govt, 1987
v1301 Bureaucratic responsiveness, 1983 (Factan Index)
v1313 Church marriages, 1986
v1314 Divorce rate, 1986
v1347 Industrial Policy Instruments, 1984
v1352 Agricultural spending capacity (%), 1978-1980
v1361 USL (local health unit) spending capacity (Italy=100), 1983
v1364 Reform legislation, 1978-1984
v1401 Budget delay (in days), 1979-1985
v1410 Statistical and Information Services, 1981
v1442 Day Care Centers (per 1000 children), 1983
v1444 Family clinics (per million pop), 1978
v1458 Legislative Innovation: Innovative Laws (1978-83) Mean Score
v1471 Housing and urban development spending capacity, 1979-1987 (Factan Index)
v1481 Mean public satisfaction with regional govt, 1988
v1482 DOXA satisfaction scores (Factan Index 1977-1988)
v1488 Preference voting (%), 1979
v1489 Preference voting (%), 1972
v1490 Preference voting, 1953-1979 (Factan Index)
v1491 Referendum turnout (per 1000 voters), 1974
v1492 Referendum turnout (per 1000 voters), 1978
v1493 Referendum turnout (per 1000 voters), 1981
v1494 Referendum turnout (per 1000 voters), 1985
v1495 Referendum turnout (per 1000 voters), 1987
v1496 Referendum Turnout, 1974-1987 (Factan Index)
v1509 Cooperatives per 1000 inhabitants, 1889
v1511 Cooperatives per 1000 inhabitants, 1901
v1513 Cooperatives per 1000 inhabitants, 1910
v1522 Observers rating of region: Programmatic (1)-Clientelistic (5)
v1535 Cooperatives per 1000 inhabitants, 1915
v1537 Cooperatives, 1889-1915 (Factan Index)
v1538 Civic Traditions, 1860-1920 (Factan Index)
v1653 "How important is religion to you?" (Eurobarometer survey, 1976-1989)
v1654 "Are you a religious person?" (Eurobarometer survey, 1976-1989)
v1655 Church attendance frequency (Eurobarometer survey, 1976-1989)
v1660 Inhabitants per sports assn.
v1661 Sports/Cult Assn. Scarcity, 1981 (Factan Index)
v1662 Civic Community (Factan Index)
v1667 Clericalism Factan Index
v1687 Number of giuntas, 1975-1985
v1691 Institutional performance (Factan Index)
LIST OF VARIABLES, ORDERED BY CONTENT

v917 Institutional Performance, 1970-1976 (Factan Index)

v1691 Institutional performance, 1978-1985 (Factan Index)
  v1364 Reform legislation, 1978-1984
  v1442 Day Care Centers (per 1000 children), 1983
  v1410 Statistical and Information Services, 1981
  v1687 Number of giuntas, 1975-1985
  v1444 Family clinics (per million pop), 1978
  v1347 Industrial Policy Instruments, 1984
  v1401 Budget delay (in days), 1979-1985
  v1361 USL (local health unit) spending capacity (Italy=100), 1983
  v1352 Agricultural spending capacity, 1978-1980
  v1458 Legislative Innovation: Innovative Laws (1978-83) Mean Score
  v1301 Bureaucratic responsiveness, 1983 (Factan Index)
  v1471 Housing and urban devel spending, 1979-1987 (Factan Index)

v1662 Civic Community (Factan Index)
  v712 Newspaper Readership, c. 1975
  v1496 Referendum Turnout, 1974-1987 (Factan Index)
    v1491 Referendum turnout (per 1000 voters), 1974
    v1492 Referendum turnout (per 1000 voters), 1978
    v1493 Referendum turnout (per 1000 voters), 1981
    v1494 Referendum turnout (per 1000 voters), 1985
    v1495 Referendum turnout (per 1000 voters), 1987

v1490 Preference Voting, 1953-1979 (Factan Index)
  v791 Preference voting (%), 1953
  v792 Preference voting (%), 1958
  v793 Preference voting (%), 1963
  v802 Preference voting (%), 1976
  v1489 Preference voting (%), 1972
  v1488 Preference voting (%), 1979

v1661 Sports/Cult Assn. Scarcity, 1981 (Factan Index)
  v1235 Inhabitants per cultural/recreational assn.
  v1660 Inhabitants per sports assn.
v1667 Clericalism (Factan Index)
   v30 Church marriages, 1976
   v1313 Church marriages, 1986
   v710 Divorce rate, 1973
   v1314 Divorce rate, 1986
   v690 Anti-divorce vote, 1974
   v1653 "How important is religion to you?" (Eurobarometer survey, 1976-1989)
   v1654 "Are you a religious person?" (Eurobarometer survey, 1976-1989)
   v1655 Church attendance frequency (Eurobarometer survey, 1976-1989)

v1482 Public Satisfaction Scores (Factan Index 1977-1988)
   v1251 Mean public satisfaction with regional govt, 1977
   v1252 Mean public satisfaction with regional govt, 1979
   v1253 Mean public satisfaction with regional govt, 1981
   v1254 Mean public satisfaction with regional govt, 1982
   v1255 Mean public satisfaction with regional govt, 1987
   v1481 Mean public satisfaction with regional govt, 1988
   All the above scores are based on the following question: "Regarding the activities of the regional government here in [region of interview], do you generally feel_?
   1 very satisfied
   2 Somewhat satisfied
   3 Not very satisfied
   4 Not at all satisfied
v1538 Civic Traditions, 1860-1920 (Factan Index)

v1105 Strength of mass parties (Soc & Cath), 1919-1921 (Factan Index)

v890 Socialist vote (%) 1919
v891 Catholic vote (%) 1919
v892 Socialist vote (%) 1921
v893 Catholic vote (%) 1921
v896 Socialist councils (%) 1920
v897 Catholic councils (%) 1920

v1537 Cooperatives, 1889-1915 (Factan Index)

v1509 Cooperatives per 1000 inhabitants, 1889
v1511 Cooperatives per 1000 inhabitants, 1901
v1513 Cooperatives per 1000 inhabitants, 1910
v1535 Cooperatives per 1000 inhabitants, 1915

v1033 Mutual Aid Societies, 1873-1904 (Factan Index)

v1010 Mutual aid societies (per 100,000 pop), 1873
v1011 Mutual aid societies (per 100,000 pop), 1878
v1012 Mutual aid societies (per 100,000 pop), 1885
v1013 Mutual aid societies (per 100,000 pop), 1895
v1061 Mutual aid societies (per 100,000 pop), 1904

v925 Electoral Turnout, 1919-1921 (Factan Index)

v864 National Electoral Turnout (%) 1919
v865 National Electoral Turnout (%) 1921
v882 Communal Electoral Turnout (%) 1920
v885 Provincial Electoral Turnout (%) 1920

v1241 Associations founded before 1860 (%)