Analysing the American National Election Study of 1948 using the codebooks package

Martin Elff
University of Essex

January 9, 2015

1 Introduction

This vignette gives an example for the analysis of a typical social science data set. It is the data file of the *American National Election Study* of 1948[1], available from the American National Election Studies website (http://www.electionstudies.org). The data file contains data from to USA-wide surveys conducted October and November 1948 by the Survey Research Centre, University Michigan (principal investigators: Angus Campbell and Robert L. Kahn). The total number of cases in the data set is 662 and the number of variables is 65 (more details about this data set can be found at http://www.electionstudies.org/studypages/1948prepost/1948prepost.htm).

With 662 cases and 65 variables, the 1948 ANES data set is relatively small as compared to current social science data sets. Such larger data sets can be processed along the same lines as in this vignette. Unlike the 1948 ANES data, their size as well as, in some cases, legal restrictions prohibit the inclusion of such a data set into the package, however.

This vignette starts with a demontration how a data file can be examined before loading it and how a subset of the data can be loaded into memory. After loading this subset into memory, some desciptive analyses are conducted that showcase the construction of contingency tables and of

[1] National Election Studies, 1948: *Post-Election Study [dataset].* Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor], 1999. ANES Dataset ID: 1948.T; ICPSR Study Number: 7218.

These materials are based on work supported by the National Science Foundation under Grant Nos.: SBR-9707741, SBR-9317631, SES-9209410, SES-9009379, SES-8808361, SES-8341310, SES-8207580, and SOC77-08885.

Any opinions, findings and conclusions or recommendations expressed in these materials are those of the author(s) and do not necessarily reflect those of the National Science Foundation.

general tables of desriptive statistics using the genTable function. In addition, a logit analysis is demonstrated and the collection of several logit coefficients into a comprehensive table by the mtable function.

It should be noted that the analyses reported in the following are conducted only for purpose of demonstrating the features of the package and are not to be considered of conclusive scientific evidence of any kind.

2 Reading In a "Portable" SPSS Data File

We start with importing the data into R. The following code extracts the SPSS portable file "NES1948.POR" from zip file "NES1948.ZIP" delivered with the codebooks package.

```
> library(mtable)
> library(codebooks)
> options(digits=3)
> nes1948.por <- UnZip("anes/NES1948.ZIP","NES1948.POR",package="codebooks")
```

Now the portable file is in a temporary directory and the path to the file is contained in the string variable nes1948.por. In the next step, the file is declared as a SPSS/PSPP "portable" file using the function PSPPportable, which as first argument takes the path to the file in question. PSPPportable reads in the information about the variables contained in the data set and counts the number of cases in the file. That is, standard I/O operations are used on the file, but the data read in are just thrown away without allocating core memory for the data. This counting of cases can, of course, be suppressed if it would take to long.

```
> nes1948 <- spss.portable.file(nes1948.por)
> print(nes1948)

SPSS portable file âĂŸ/tmp/RtmprplNk6/NES1948.PORâĂŹ
        with 67 variables and 662 observations
```

At this stage, the data are not loaded into the memory yet. But we can see which variables exist inside the data set:

```
> names(nes1948)

 [1] "vversion" "vdsetno"  "v480001"  "v480002"  "v480003"  "v480004"
 [7] "v480005"  "v480006"  "v480007"  "v480008"  "v480009"  "v480010"
[13] "v480011"  "v480012"  "v480013"  "v480014a" "v480014b" "v480015a"
[19] "v480015b" "v480016a" "v480016b" "v480017a" "v480017b" "v480018"
[25] "v480019"  "v480020"  "v480021a" "v480021b" "v480022a" "v480022b"
[31] "v480023"  "v480024"  "v480025a" "v480025b" "v480026"  "v480027"
[37] "v480028"  "v480029"  "v480030"  "v480031a" "v480031b" "v480031c"
[43] "v480032a" "v480032b" "v480032c" "v480033a" "v480033b" "v480034a"
```

```
[49] "v480034b" "v480035a" "v480035b" "v480036a" "v480036b" "v480037"
[55] "v480038"  "v480039"  "v480040"  "v480041"  "v480042"  "v480043"
[61] "v480044"  "v480045"  "v480046"  "v480047"  "v480048"  "v480049"
[67] "v480050"
```

We also can ask for a description ("variable label") for each variable:

```
> description(nes1948)

 vversion âĂŸNES VERSION NUMBERâĂŹ
 vdsetno  âĂŸNES DATASET NUMBERâĂŹ
 v480001  âĂŸICPSR ARCHIVE NUMBERâĂŹ
 v480002  âĂŸINTERVIEW NUMBERâĂŹ
 v480003  âĂŸPOP CLASSIFICATIONâĂŹ
 v480004  âĂŸCODERâĂŹ
 v480005  âĂŸNUMBER OF CALLS TO RâĂŹ
 v480006  âĂŸR REMEMBER PREVIOUS INTâĂŹ
 v480007  âĂŸINTR INTERVIEW THIS RâĂŹ
 v480008  âĂŸPRVS PRE-ELCTN R REINTâĂŹ
 v480009  âĂŸR INT IN PRE/POSTELCTNâĂŹ
 v480010  âĂŸRENT CNTRL KEPT/DROPPEDâĂŹ
 v480011  âĂŸGOVT CONTROL PRICESâĂŹ
 v480012  âĂŸWHAT TO DO W TFT-HT ACTâĂŹ
 v480013  âĂŸPRESLELCTN OTCM SURPRISEâĂŹ
 v480014a âĂŸWHY PPL VTD FOR TRUMAN 1âĂŹ
 v480014b âĂŸWHY PPL VTD FOR TRUMAN 2âĂŹ
 v480015a âĂŸWHY PPL VTD AGNST TRUMAN 1âĂŹ
 v480015b âĂŸWHY PPL VTD AGNST TRUMAN 2âĂŹ
 v480016a âĂŸWHY PPL VTD FOR DEWEY 1âĂŹ
 v480016b âĂŸWHY PPL VTD FOR DEWEY 2âĂŹ
 v480017a âĂŸWHY PPL VTD AGNST DEWEY 1âĂŹ
 v480017b âĂŸWHY PPL VTD AGNST DEWEY 2âĂŹ
 v480018  âĂŸDID R VOTE/FOR WHOMâĂŹ
 v480019  âĂŸWN DECIDE FOR WHOM TO VTâĂŹ
 v480020  âĂŸCNSD VT FOR SOMEONE ELSEâĂŹ
 v480021a âĂŸXWHY DID NOT VT FOR HIM 1âĂŹ
 v480021b âĂŸXWHY DID NOT VT FOR HIM 2âĂŹ
 v480022a âĂŸWHY VT THE WAY YOU DID 1âĂŹ
 v480022b âĂŸWHY VT THE WAY YOU DID 2âĂŹ
 v480023  âĂŸVOTED STRAIGHT TICKETâĂŹ
 v480024  âĂŸR NOT VT-IF VT,FOR WHOMâĂŹ
 v480025a âĂŸR NOT VT-WHY DID NOT VT 1âĂŹ
 v480025b âĂŸR NOT VT-WHY DID NOT VT 2âĂŹ
 v480026  âĂŸR NOT VT-WAS R REG TO VTâĂŹ
 v480027  âĂŸVTD IN PRVS PRESL ELCTNâĂŹ
 v480028  âĂŸVTD FOR WHOM IN 1944âĂŹ
 v480029  âĂŸOCCUPATION OF HEADâĂŹ
 v480030  âĂŸHEAD BELONG TO LBR UNâĂŹ
 v480031a âĂŸGRPS IDENTIFIED W TRUMAN 1âĂŹ
 v480031b âĂŸGRPS IDENTIFIED W TRUMAN 2âĂŹ
```

```
v480031c âĂŸGRPS IDENTIFIED W TRUMAN 3âĂŹ
v480032a âĂŸGRPS IDENTIFIED W DEWEY 1âĂŹ
v480032b âĂŸGRPS IDENTIFIED W DEWEY 2âĂŹ
v480032c âĂŸGRPS IDENTIFIED W DEWEY 3âĂŹ
v480033a âĂŸISSUES CONNECTED W TRMN 1âĂŹ
v480033b âĂŸISSUES CONNECTED W TRMN 2âĂŹ
v480034a âĂŸISSUES CONNECTED W DEWEY 1âĂŹ
v480034b âĂŸISSUES CONNECTED W DEWEY 2âĂŹ
v480035a âĂŸPERSONAL ATTRIBUTE TRMN 1âĂŹ
v480035b âĂŸPERSONAL ATTRIBUTE TRMN 2âĂŹ
v480036a âĂŸPERSONAL ATTRIBUTE DEWEY 1âĂŹ
v480036b âĂŸPERSONAL ATTRIBUTE DEWEY 2âĂŹ
v480037  âĂŸCMPN INCIDENTS MENTIONEDâĂŹ
v480038  âĂŸ41-PRESLELCTN PLAN TO VTâĂŹ
v480039  âĂŸ41-PLAN TO VT REP/DEMâĂŹ
v480040  âĂŸ41-USA'S CNCRN W OTHERSâĂŹ
v480041  âĂŸ41-SATISD USA TWRD RUSSâĂŹ
v480042  âĂŸ41-INFORMATION LEVELâĂŹ
v480043  âĂŸ41-USA GV IN,AGRT RUSSâĂŹ
v480044  âĂŸ41-USA-RUSS AGRT VIA U.NâĂŹ
v480045  âĂŸSEX OF RESPONDENTâĂŹ
v480046  âĂŸRACE OF RESPONDENTâĂŹ
v480047  âĂŸAGE OF RESPONDENTâĂŹ
v480048  âĂŸEDUCATION OF RESPONDENTâĂŹ
v480049  âĂŸTOTAL 1948 INCOMEâĂŹ
v480050  âĂŸRELIGIOUS PREFERENCEâĂŹ
```

or even a code book using

```
> codebook(nes1948)
```

(this is not shown here because the output would have taken more then thirty pages).

2.1 Reading In a Subset of the Data

After we have decided which variables to use we can read in a subset of the data:

```
vote.48 <- subset(nes1948,
              select=c(
                  v480018,
                  v480029,
                  v480030,
                  v480045,
                  v480046,
                  v480047,
                  v480048,
                  v480049,
```

```
                 v480050
                 ))
```

```
> vote.48 <- subset(nes1948,
+              select=c(
+                  v480018,
+                  v480029,
+                  v480030,
+                  v480045,
+                  v480046,
+                  v480047,
+                  v480048,
+                  v480049,
+                  v480050
+                  ))
```

The subset of the ANES 1948 we read in is now contained in the variable `vote.48`, which contains an object of class `data.set`. A `data.set` is an "embellished" version of a `data.frame`, a data structure intended to contained `labelled` vectors. `labelled` vectors contain the all the special information attached to the variables in the original data set, such as variable labels, value labels, and general missing values. A short summary of this special information shows up after a call to `str`.

```
> str(vote.48)

Data set with 662 obs. of 9 variables:
 $ v480018: Nmnl. item w/ 7 labels for 1,2,3,... + ms.v. 1 2 1 2 1 2 2 1 2 1 ...
 $ v480029: Nmnl. item w/ 12 labels for 10,20,30,... + ms.v. 70 30 40 10 10 20 80 80 40 40
 $ v480030: Nmnl. item w/ 4 labels for 1,2,8,... + ms.v. 1 2 2 2 2 2 2 2 1 1 ...
 $ v480045: Nmnl. item w/ 3 labels for 1,2,9 + ms.v. 1 2 2 2 1 2 1 2 1 1 ...
 $ v480046: Nmnl. item w/ 4 labels for 1,2,3,... + ms.v. 1 1 1 1 1 1 1 1 1 1 ...
 $ v480047: Nmnl. item w/ 7 labels for 1,2,3,... + ms.v. 3 3 2 3 2 3 4 5 2 2 ...
 $ v480048: Nmnl. item w/ 4 labels for 1,2,3,... + ms.v. 1 2 2 3 3 2 1 1 2 2 ...
 $ v480049: Nmnl. item w/ 8 labels for 1,2,3,... + ms.v. 4 7 5 7 5 7 5 2 5 6 ...
 $ v480050: Nmnl. item w/ 6 labels for 1,2,3,... + ms.v. 1 1 2 1 2 1 1 1 1 2 ...
```

This output shows, for example, that variable `v480018` has the description (variable label) "DID R VOTE/FOR WHOM" is considered as having nominal level of measurement, has seven value labels and one defined missing value.

Since the variable names in the ANES data set are not very mnemonic, we rename the variables:

```
vote.48 <- rename(vote.48,
                 v480018 = "vote",
                 v480029 = "occupation.hh",
                 v480030 = "unionized.hh",
```

```
                 v480045 = "gender",
                 v480046 = "race",
                 v480047 = "age",
                 v480048 = "education",
                 v480049 = "total.income",
                 v480050 = "religious.pref"
       )
```

```
> vote.48 <- rename(vote.48,
+                  v480018 = "vote",
+                  v480029 = "occupation.hh",
+                  v480030 = "unionized.hh",
+                  v480045 = "gender",
+                  v480046 = "race",
+                  v480047 = "age",
+                  v480048 = "education",
+                  v480049 = "total.income",
+                  v480050 = "religious.pref"
+        )
```

Before we start with analyses, we take a closer look at the data.

```
> codebook(vote.48)

========================================================================

   vote âĂŸDID R VOTE/FOR WHOMâĂŹ

------------------------------------------------------------------------

   Storage mode: double
   Measurement: nominal
   Missing values: 9

           Values and Labels      N     Percent

       1   âĂŸVOTED - FOR TRUMANâĂŹ  212   32.1 32.0
       2   âĂŸVOTED - FOR DEWEYâĂŹ   178   27.0 26.9
       3   âĂŸVOTED - FOR WALLACEâĂŹ   1    0.2  0.2
       4   âĂŸVOTED - FOR OTHERâĂŹ    11    1.7  1.7
       5   âĂŸVOTED - NA FOR WHOMâĂŹ  20    3.0  3.0
       6   âĂŸDID NOT VOTEâĂŹ        238   36.1 36.0
       9 M âĂŸNA WHETHER VOTEDâĂŹ      2         0.3

========================================================================

   occupation.hh âĂŸOCCUPATION OF HEADâĂŹ

------------------------------------------------------------------------
```

Storage mode: double
Measurement: nominal
Missing values: 99

Values and Labels			N	Percent	
10		âĂŸPROFESSIONAL, SEMI-PROFESSIONALâĂŹ	44	6.9	6.6
20		âĂŸSELF-EMPLOYED, MANAGERIAL, SUPERVISORYâĂŹ	73	11.5	11.0
30		âĂŸOTHER WHITE-COLLAR (CLERICAL, SALES, ETâĂŹ	79	12.5	11.9
40		âĂŸSKILLED AND SEMI-SKILLEDâĂŹ	164	25.9	24.8
60		âĂŸPROTECTIVE SERVICEâĂŹ	6	0.9	0.9
70		âĂŸUNSKILLED, INCLUDING FARM AND SERVICE WâĂŹ	85	13.4	12.8
80		âĂŸFARM OPERATORS AND MANAGERSâĂŹ	105	16.6	15.9
92		âĂŸSTUDENTâĂŹ	7	1.1	1.1
94		âĂŸUNEMPLOYEDâĂŹ	5	0.8	0.8
95		âĂŸRETIRED, TOO OLD OR UNABLE TO WORKâĂŹ	38	6.0	5.7
96		âĂŸHOUSEWIFEâĂŹ	28	4.4	4.2
99	M	âĂŸNAâĂŹ	28		4.2

==

unionized.hh âĂŸHEAD BELONG TO LBR UNâĂŹ

Storage mode: double
Measurement: nominal
Missing values: 8-Inf

Values and Labels			N	Percent	
1		âĂŸYESâĂŹ	150	23.3	22.7
2		âĂŸNOâĂŹ	493	76.7	74.5
8	M	âĂŸDKâĂŹ	5		0.8
9	M	âĂŸNAâĂŹ	14		2.1

==

gender âĂŸSEX OF RESPONDENTâĂŹ

Storage mode: double
Measurement: nominal
Missing values: 9

Values and Labels N Percent

```
      1   âĂŸMALEâĂŹ      302   45.8 45.6
      2   âĂŸFEMALEâĂŹ    357   54.2 53.9
      9 M âĂŸNAâĂŹ          3         0.5

========================================================================

   race âĂŸRACE OF RESPONDENTâĂŹ

------------------------------------------------------------------------

   Storage mode: double
   Measurement: nominal
   Missing values: 9

       Values and Labels  N     Percent

         1   âĂŸWHITEâĂŹ    585   90.7 88.4
         2   âĂŸNEGROâĂŹ     60    9.3  9.1
         3   âĂŸOTHERâĂŹ      0    0.0  0.0
         9 M âĂŸNAâĂŹ        17         2.6

========================================================================

   age âĂŸAGE OF RESPONDENTâĂŹ

------------------------------------------------------------------------

   Storage mode: double
   Measurement: nominal
   Missing values: 9

       Values and Labels  N     Percent

       1   âĂŸ18-24âĂŹ        57    8.7  8.6
       2   âĂŸ25-34âĂŹ       142   21.7 21.5
       3   âĂŸ35-44âĂŹ       174   26.6 26.3
       4   âĂŸ45-54âĂŹ       125   19.1 18.9
       5   âĂŸ55-64âĂŹ        86   13.1 13.0
       6   âĂŸ65 AND OVERâĂŹ  70   10.7 10.6
       9 M âĂŸNAâĂŹ           8         1.2

========================================================================

   education âĂŸEDUCATION OF RESPONDENTâĂŹ

------------------------------------------------------------------------

   Storage mode: double
   Measurement: nominal
```

```
   Missing values: 9

       Values and Labels   N     Percent

       1   âĂŸGRADE SCHOOLâĂŹ 292   44.4 44.1
       2   âĂŸHIGH SCHOOLâĂŹ  266   40.4 40.2
       3   âĂŸCOLLEGEâĂŹ      100   15.2 15.1
       9 M âĂŸNAâĂŹ             4         0.6

==========================================================================

   total.income âĂŸTOTAL 1948 INCOMEâĂŹ

--------------------------------------------------------------------------

   Storage mode: double
   Measurement: nominal
   Missing values: 9

        Values and Labels    N     Percent

       1   âĂŸUNDER $500âĂŹ      25    3.8  3.8
       2   âĂŸ$500-$999âĂŹ       43    6.6  6.5
       3   âĂŸ$1000-1999âĂŹ     110   16.8 16.6
       4   âĂŸ$2000-2999âĂŹ     185   28.2 27.9
       5   âĂŸ$3000-3999âĂŹ     142   21.7 21.5
       6   âĂŸ$4000-4999âĂŹ      66   10.1 10.0
       7   âĂŸ$5000 AND OVERâĂŹ  84   12.8 12.7
       9 M âĂŸNAâĂŹ               7         1.1

==========================================================================

   religious.pref âĂŸRELIGIOUS PREFERENCEâĂŹ

--------------------------------------------------------------------------

   Storage mode: double
   Measurement: nominal
   Missing values: 9

       Values and Labels  N     Percent

       1   âĂŸPROTESTANTâĂŹ  460   70.0 69.5
       2   âĂŸCATHOLICâĂŹ    140   21.3 21.1
       3   âĂŸJEWISHâĂŹ       25    3.8  3.8
       4   âĂŸOTHERâĂŹ        14    2.1  2.1
       5   âĂŸNONEâĂŹ         18    2.7  2.7
       9 M âĂŸNAâĂŹ            5         0.8
```

We now have obtained a *codebook*, which contains information of the class and type of the variables in the data set, the value labels and defined missing values, and counts of the distinct values of the variables.

3 Analysis

3.1 Some Descriptive Analyses

We start our analyses with a contingency table, but first we make some preparations: We recode the variables of interest into a smaller number of categories in order to get results that are easier to read and interpret.

```
vote.48 <- within(vote.48,{
  vote3 <- recode(vote,
    1   -> "Truman",
    2   -> "Dewey",
    3:4 -> "Other"
    )
  occup4 <- recode(occupation.hh,
    10:20 -> "Upper white collar",
    30    -> "Other white collar",
    40:70 -> "Blue collar",
    80    -> "Farmer"
    )
  relig3 <- recode(religious.pref,
    1   -> "Protestant",
    2   -> "Catholic",
    3:5 -> "Other/none"
    )
   race2 <- recode(race,
    1 -> "White",
    2 -> "Black"
    )
  })
```

```
> vote.48 <- within(vote.48,{
+   vote3 <- recode(vote,
+     1 -> "Truman",
+     2 -> "Dewey",
+     3:4 -> "Other"
+     )
+   occup4 <- recode(occupation.hh,
+     10:20 -> "Upper white collar",
+     30 -> "Other white collar",
+     40:70 -> "Blue collar",
+     80 -> "Farmer"
+     )
+   relig3 <- recode(religious.pref,
```

```
+     1 -> "Protestant",
+     2 -> "Catholic",
+     3:5 -> "Other,none"
+     )
+   race2 <- recode(race,
+     1 -> "White",
+     2 -> "Black"
+     )
+   })
```

Having constructed the unordered factors `vote3`, `occup4`, `relig3`, and `race2` we can proceed examining the association the vote, occupational class, relgious denomination, and race. First, we look upon a simple contingency table. We use the `toLatex` method defined for tables to get a nicely formatted output. (The generic function `toLatex` is defined in package `utils`.)

```
> toLatex(xtabs(~vote3+occup4,data=vote.48))
```

	Upper white collar	Other white collar	Blue collar	Farmer
Truman	17	30	114	26
Dewey	67	31	36	14
Other	2	0	4	3

Tables of percentages may seem more informative about the impact of various factors on the vote. So we use the function `genTable` to obtain such tables of percentages:

```
> toLatex(t(genTable(percent(vote3)~occup4,data=vote.48)),
+   digits=c(1,1,1,0))
```

	Truman	Dewey	Other	N
Upper white collar	19.8	77.9	2.3	86
Other white collar	49.2	50.8	0.0	61
Blue collar	74.0	23.4	2.6	154
Farmer	60.5	32.6	7.0	43

Obviously, voters from farmer and blue collar worker households were especially supportive of President Truman, while voters of upper white collar background largely supported the Republican Candidate Dewey.

```
> toLatex(t(genTable(percent(vote3)~relig3,data=vote.48)),
+   digits=c(1,1,1,0))
```

	Truman	Dewey	Other	N
Protestant	44.7	51.0	4.3	255
Catholic	66.0	34.0	0.0	103
Other,none	68.2	29.5	2.3	44

This table shows that Catholics and adherents of other denominations were more supportive of Truman than of Dewey.

```
> toLatex(t(genTable(percent(vote3)~race2,data=vote.48)),
+   digits=c(1,1,1,0))
```

	Truman	Dewey	Other	N
White	51.3	45.5	3.2	376
Black	64.7	35.3	0.0	17

African Americans apparently supported Truman by a large majority. The number of members of this group in the sample is very small, however, so that such an inference would be very shaky.

```
> inc.tab <- t(genTable(percent(vote3)~total.income,data=vote.48))
> rownames(inc.tab) <- gsub("$","\\$",rownames(inc.tab),fixed=TRUE)
> toLatex(inc.tab,digits=c(1,1,1,0))
```

	Truman	Dewey	Other	N
UNDER $500	50.0	50.0	0.0	8
$500-$999	61.5	38.5	0.0	13
$1000-1999	64.4	32.2	3.4	59
$2000-2999	67.0	30.1	2.9	103
$3000-3999	47.5	48.5	4.0	101
$4000-4999	45.8	50.0	4.2	48
$5000 AND OVER	31.8	68.2	0.0	66

3.2 Logit Modelling of Candidate Choice

In the following we conduct a logit analysis of the vote for Truman. First, we assign non-standard contrasts the categorical predictors. Here, the function `contr` is used to assign treatment (dummy) contrasts to `occup4` and `total.income` with baseline category 3 and 4, respectively.

```
> vote.48 <- within(vote.48,{
+   contrasts(occup4) <- contr("treatment",base = 3)
+   contrasts(total.income) <- contr("treatment",base = 4)
+   })
```

We now fit some logistic regression models of the impact occupational class, income, and religious denomination on the vote choice supporting Truman. The contrasts of the occupational class and income factors are such that they compare the choices of the members of the blue-collar class with all other classes and the middle income group ($ 2000-2999) with the other income groups. The religious denomination factor compares Protestants with Catholics and those with other or no denominations.

```
> model1 <- glm((vote3=="Truman")~occup4,data=vote.48,
+               family="binomial")
> model2 <- glm((vote3=="Truman")~total.income,data=vote.48,
+               family="binomial")
> model3 <- glm((vote3=="Truman")~occup4+total.income,data=vote.48,
+               family="binomial")
> model4 <- glm((vote3=="Truman")~relig3,data=vote.48,
+               family="binomial")
> model5 <- glm((vote3=="Truman")~occup4+relig3,data=vote.48,
+               family="binomial")
```

First, we use `mtable` to construct a comparative table of the estimates of `model1`, `model2`, and `model3`. We thus can compare the impact of occupational class and income on the choice of candidate Truman.

```
> mtable(model1,model2,model3,summary.stats=c("Nagelkerke R-sq.","Deviance","AIC","N"))

Calls:
model1: glm(formula = (vote3 == "Truman") ~ occup4, family = "binomial",
    data = vote.48)
model2: glm(formula = (vote3 == "Truman") ~ total.income, family = "binomial",
    data = vote.48)
model3: glm(formula = (vote3 == "Truman") ~ occup4 + total.income, family = "binomial",
    data = vote.48)

======================================================================
                                          model1    model2    model3
----------------------------------------------------------------------
(Intercept)                               1.047***  0.708***  1.316***
                                         (0.184)   (0.210)   (0.268)
occup4: Upper white collar/Blue collar   -2.448***           -2.328***
                                         (0.327)             (0.357)
occup4: Other white collar/Blue collar   -1.080***           -1.015**
                                         (0.315)             (0.323)
occup4: Farmer/Blue collar               -0.622              -0.792*
                                         (0.362)             (0.383)
total.income: UNDER $500/$2000-2999                -0.708    -0.662
                                                   (0.737)   (1.056)
total.income: $500-$999/$2000-2999                 -0.238     0.912
                                                   (0.607)   (1.143)
total.income: $1000-1999/$2000-2999                -0.115     0.144
```

```
                                                    (0.343)   (0.440)
total.income: $3000-3999/$2000-2999                 -0.807**  -0.527
                                                    (0.289)   (0.338)
total.income: $4000-4999/$2000-2999                 -0.875*   -0.509
                                                    (0.358)   (0.411)
total.income: $5000 AND OVER/$2000-2999             -1.470*** -0.535
                                                    (0.337)   (0.405)
----------------------------------------------------------------------
Nagelkerke R-sq.                           0.246     0.085     0.274
Deviance                                 404.190   524.433   390.551
AIC                                      412.190   538.433   410.551
N                                        344       398       340
======================================================================
```

`mtable` returns an object of class `"mtable"`. The `print` method for this class produces output as seen above. This output has a format close to the requirements of social science publications. With the `toLatex` method for objects of this class one can produce almost publication-ready output:

```
> toLatex(relabel(mtable(
+             "Model 1"=model1,
+             "Model 2"=model2,
+             "Model 3"=model3,
+             summary.stats=c("Nagelkerke R-sq.","Deviance","AIC","N")),
+           "[(]Intercept[)]"="\\\\emph{Intercept}",
+           "[$]"="\\\\$",
+           UNDER="under",
+           "AND OVER"="and over",
+           occup4="Occup. class",
+           total.income="Income",
+           gsub=TRUE
+           ),
+           ddigits=5
+         )
```

	Model 1	Model 2	Model 3
Intercept	1.047^{***}	0.708^{***}	1.316^{***}
	(0.184)	(0.210)	(0.268)
Occup. class: Upper white collar/Blue collar	-2.448^{***}		-2.328^{***}
	(0.327)		(0.357)
Occup. class: Other white collar/Blue collar	-1.080^{***}		-1.015^{**}
	(0.315)		(0.323)
Occup. class: Farmer/Blue collar	-0.622		-0.792^{*}
	(0.362)		(0.383)
Income: under \$500/\$2000-2999		-0.708	-0.662
		(0.737)	(1.056)
Income: \$500-\$999/\$2000-2999		-0.238	0.912
		(0.607)	(1.143)
Income: \$1000-1999/\$2000-2999		-0.115	0.144
		(0.343)	(0.440)
Income: \$3000-3999/\$2000-2999		-0.807^{**}	-0.527
		(0.289)	(0.338)
Income: \$4000-4999/\$2000-2999		-0.875^{*}	-0.509
		(0.358)	(0.411)
Income: \$5000 and over/\$2000-2999		-1.470^{***}	-0.535
		(0.337)	(0.405)
Nagelkerke R-sq.	0.246	0.085	0.274
Deviance	404.190	524.433	390.551
AIC	412.190	538.433	410.551
N	344	398	340

The comparison of the pseudo-R-Square values of model 1 and 2 suggests that occupational class has a stronger influence on a preference for Truman than household income. Indeed, if occupational class is taken into account, the effect of income is no longer statistically significant as the column corresponding to model 3 indicates.

Second, we compare the effect of occupational class and religious denomination on the preference for Truman along the same lines as above. We use `mtable` to collect the estimates of `model1`, `model4`, and `model5` into a common table.

```
> toLatex(relabel(mtable(
+              "Model 1"=model1,
+              "Model 4"=model4,
+              "Model 5"=model5,
+              summary.stats=c("Nagelkerke R-sq.","Deviance","AIC","N")),
+          "[(]Intercept[)]"="\\\\emph{Intercept}",
+            occup4="Occup. class",
+            relig3="Religion",
+            gsub=TRUE
+            ),
+          ddigits=5)
```

	Model 1	Model 4	Model 5
Intercept	1.047^{***}	-0.213	0.698^{**}
	(0.184)	(0.126)	(0.216)
Occup. class: Upper white collar/Blue collar	-2.448^{***}		-2.385^{***}
	(0.327)		(0.337)
Occup. class: Other white collar/Blue collar	-1.080^{***}		-1.098^{***}
	(0.315)		(0.326)
Occup. class: Farmer/Blue collar	-0.622		-0.346
	(0.362)		(0.374)
Religion: Catholic/Protestant		0.877^{***}	0.685^{*}
		(0.243)	(0.292)
Religion: Other,none/Protestant		0.975^{**}	1.191^{**}
		(0.347)	(0.441)
Nagelkerke R-sq.	0.246	0.060	0.281
Deviance	404.190	537.711	393.105
AIC	412.190	543.711	405.105
N	344	402	344

A comparison of the pseudo-R-squared values suggests that also the effect of religious denomination is weaker than that of occupational class. However, as the third column in the above table indicates the effect of religious denomination remains statistically significant.