AUSTRALIAN SURVEY OF SOCIAL ATTITUDES 2011

USER GUIDE

This purpose of this user guide is to give a brief explanation of the Australian Social Survey (AuSSA) 2011 dataset.

Contents

Abstract

The 2011 Australian Survey of Social Attitudes (AuSSA) is the fifth in a biennial series that studies social attitudes and behaviour of Australian citizens for the Australian and international research community. AuSSA provides cross-sectional data on the social attitudes and behaviour of Australians, repeating a core questionnaire for each cross-section and fielding specific modules relevant to the changing needs of the social research community. AuSSA is Australia's official survey in the International Social Survey Program and regularly includes ISSP modules. AuSSA 2011 includes both the ISSP's Environment III (2010) and Health (2011) modules.

The 2011 Survey includes attitudes and behaviours that are organised into seven categories: Issues facing Australia today; Environment; Health; Government Policy and services; Kindness; Australians in History; and The Law and Authority.

AuSSA 2011 also includes demographic and behavioural categories (Personal Background) that survey: sex, year born, income, education, employment, union membership, country of birth, household composition and religion. There are also questions about the partner of the respondent: employment, highest-level of education and income.

Sample information

Over 6,000 individuals were sampled from the 2011 Electoral Roll. Within each state, individuals were chosen using systematic random sampling. The table below displays some information regarding the response rates and the dates of fieldwork.

To minimise sample losses several actions were taken. One week before the questionnaires were mailed out a pre-letter introducing the survey was sent. Around one and a half weeks after the questionnaires were mailed, a reminder postcard was sent. Replacement questionnaires were sent to respondents who had misplaced or lost their original questionnaires but were willing to participate in the survey. A second reminder postcard was sent approximately one month after the first mail-out of the questionnaires.

TABLE 1. SAMPLE INFORMATION AND FIELDWORK DATES

	AuSSA 2011
Sample size	6,250
Final number of respondents	1,946
AAPOR Response Rate 1	31 %
Cooperation Rate 1	68 %
Contact Rate 1	46 %
Less than half of questionnaire filled in	2
Cannot complete due to language	5
Blank questionnaire returned	794
Known respondent refusal	37
Respondent unavailable during field period	15
Death	16
Physically or mentally unable/incompetent	34
Addressee Not Known at Place of Address	515
Nothing ever returned	2881
Duplicate listings	5
Dates of fieldwork	
Pre-letter sent	1-May-12
Questionnaire mailed out	8-May-12
Reminder postcard	15-May-12
Replacement questionnaires mailed out	30-May-12
Second reminder postcard sent	6-Jun-12
Close	8-Aug-12

Missing codes

All missing codes in the dataset are assigned to negative numbers. The following missing codes are used:

-1 Missing Used when the respondent was did not provide an answer to a question.

-2 Skip: Used when the respondent was not asked a particular question because it did not apply to them and they had therefore been filtered away from it.

-4 No partner: Used in questions relating to partner's characteristics, in cases where the respondent did not have a partner. A respondent was classified as not having a partner if they were not living with a partner or if the question on living with a partner was missing (*h20*).

-5 Multiple response: Used when the respondent was meant to give one response only, but instead gave multiple responses. For example in variable *a1* respondents were asked to choose who they thought was the most important issue for Australia today. People who gave multiple responses chose more than one option e.g. both health care and education.

-6 Illegible or impossible value: Used when the respondent did provide some information but it was too faint to read, or it was an impossible value. For example a response of 300 for height (*c28*). The purpose of this variable is to differentiate cases were something was written from those with unit non-response (-1 missing).

-7 Don't know: This was used in a few instances where the respondent had written a note on the questionnaire indicating they did not know. Also used for the height and weight variables (c28 & c29).

Data quality

Variables *h42* and *h43* (income) should be used with caution. A large number of respondents had difficulty interpreting and answering these questions leading to poor quality of data.

Australian Bureau of Statistics (ABS) classifications

The following ABS classifications were used for selected variables.

Topic	Variables		ABS coding standard
Country of birth	h36o Respondent country of birth h37o Mother's country of birth h38o Father's country of birth		*SACC- Standard Australian Classification of Countries, 2011*
Ancestry	h39oa Ancestry – other 1 h39ob Ancestry – other 2		*ASCCEG- Australian Standard Classification of Cultural and Ethnic Groups, 2011*
Occupation	occupation_l1 occupation_l2 occupation_l3 occupation_l4	Respondent's occupation	*ANZSCO - Australian and New Zealand Standard Classification of Occupations, First Edition, Revision 1*
	poccupation_l1 poccupation_l2 poccupation_l3 poccupation_l4	Partner's occupation	
Industry	industry_l1 industry_l2 industry_l3 industry_l4	Industry of respondent's place of work	*ANZSIC- Australian and New Zealand Standard Industrial Classification , 2006, Revision 1.0)*
	pindustry_l1 pindustry_l2 pindustry_l3 pindustry_l4	Industry of partner's place of work	

Weighting

Responses were biased toward women, the more highly educated and older people. To correct for these biases, the AuSSA 2011 sample has been benchmarked against the Census 2011 data on Australian *citizens*, following the procedure and the STATA code used by Trevor Breusch (2004) for the 2003 AuSSA data. The weighted sample is designed to have the same proportions as the Census in a three-way cross-tabulation of age, sex, and educational attainment.

	AuSSA 2011	Census
	Sample %	Expected %
Age		
18-29	8	20
30-39	10	17
40-49	16	19
50-64	34	25
65+	30	18
Missing	2	
Total	100	100
Sex		
Male	47	49
Female	52	51
Missing	1	
Total	100	100
Highest education		
Less than Year 12	23	32
Year 12	8	18
Trade, certificate, diploma	33	30
Bachelor degree	15	15
Postgraduate degree or diploma	16	6
Missing	4	
Total	100	100

The cross tabulation by age, sex and highest education level from the Census was done using the 2011 Census Table Builder. A table was created using sex and age (single years) in the rows, and Australian Citizenship, Highest level of education and highest year of school completed in the columns. The final table was created by excluding those who did not have Australian citizenship (this included those who did not state their citizenship). Individuals with postgraduate degrees and graduate diplomas were combined into one category. Individuals with certificates or diplomas/advanced diplomas were also combined into one category. Those who were recorded as having their highest non-school qualification *as inadequately described, not stated* or *not applicable* were assumed not to have a post-school qualification and were coded according to their highest level of school education. Those who had had completed Year 12 were coded as having completed high-school education. Those who completed up to Year 11 or below (included those who did not state their highest level of school education) were coded as having not completed high school.

Constructing sampling weights

The sampling weights are designed to rebalance the sample so the weighted sample frequencies are equal to the expected frequencies in the three-way tabulation by age group (five groups), sex, and educational attainment (five levels).

Only 1,834 of the 1,946 observations had complete information on all three of these variables. Imputation of missing data was possible in some cases however. Where educational attainment was missing but occupational category was available, for the purposes of constructing the weights, people were given the most common education level for their occupation (irrespective of their age).

When age was missing, in some cases, this was imputed from other information available in the questionnaire. If there was information on the respondents main activity they were given the average age of people with the same activity (e.g. if the respondent was retired, they were given the average age of retired people).

For people who were missing on gender, they had a weight allocated to them that was the average of the weights over gender for their education and age. To avoid extreme weights, the weights were trimmed at the 1 per cent and 99 per cent level, before being rescaled so that they averaged to one across all cases in each subsample.

The average calculated weights are shown in the following table:

Highest education

Sex and age	Not complete high school	Complete high school	Certificate or diploma	Bachelor degree	Postgraduate degree or diploma
18-29					
Male	5.73	3.92	1.94	1.56	0.84
Female	5.73	5.73	1.20	1.21	0.91
30-39					
Male	5.73	2.75	2.36	1.41	0.67
Female	5.57	3.52	1.35	1.53	0.47
40-49					
Male	1.80	2.68	1.06	0.96	0.35
Female	2.69	1.86	0.84	0.81	0.53
50-64					
Male	1.08	1.01	0.73	0.87	0.33
Female	1.24	1.43	0.48	0.58	0.21
65 and over					
Male	0.53	0.68	0.61	0.48	0.16
Female	0.89	1.08	0.29	0.53	0.15

The effects

For most variables the effect of weighting is only moderate.

[H31] Frequency of attending religious services

	Unweighted (%)	Weighted (%)
Several times a week or more often	2	2
Once a week	11	9
2 or 3 times a month	3	3
Once a month	3	2
Several times a year	12	12
Once a year	7	7
Less frequently than once a year	17	16
Never	42	44
Missing	3	6
Total	100	100

*Total may add up to more than 100, due to rounding

For other variables more closely related to age, sex or highest education the effect of weighting is more noticeable, as with the example of home ownership.

[H46] Own or rent dwelling

	Unweighted (%)	Weighted (%)
Own outright	46	34
Own, paying off mortgage	31	34
Rent from private landlord	11	12
Rent from public housing authority	3	3
Other (boarding, living at home, etc)	6	12
Missing	3	6
Total	100	100

*Total may add up to more than 100, due to rounding

Citing the survey

To cite the survey, please use the following:

Evans, A. *Australian Survey of Social Attitudes 2011*, The Australian National University: Australian Demographic & Social Research Institute, 2011

For further information regarding the survey please visit the website:

http://aussa.anu.edu.au/index.php

APPENDIX A: *STATA* code used to generate weights

```
************************************************************
* Name: weight_final.do
* Purpose: Sampling Weights for AuSSA 2011
*
* Author: Anna Reimondos - adapted from Trevor Breusch
*              weights used in 2003 AuSSA
* Date: November 2012
* Stata version: 12.1
************************************************************

* Weight for age group, gender and education attainment

preserve

capture drop gender
clonevar gender=h1

*Gen tempage group
capture drop tempage
gen tempage=age
replace tempage=. if tempage==-1
*If age is missing estimate based on average for that sex*retired
capture drop retired
gen retired=.
replace retired=1 if h19==6
replace retired=0 if h19!=6 & h19!=.
mean(tempage), over(gender retired)
replace tempage=48 if gender==1 & retired==0 & tempage==.
replace tempage=73 if gender==1 & retired==1 & tempage==.

replace tempage=47 if gender==2 & retired==0 & tempage==.
replace tempage=71 if gender==2 & retired==1 & tempage==.

*If retirement missing use partner's retirement?
gen pretired=.
replace pretired=1 if h28==6
mean(tempage) if pretired==1, over(gender)

replace tempage=72 if gender==1 & pretired==1 & tempage==.
replace tempage=69 if gender==2 & pretired==1 & tempage==.

*Otherwise just use average
mean(tempage) ,over(gender)
replace tempage=56 if gender==1 & tempage==.
replace tempage=54 if gender==2 & tempage==.

mean(tempage), over(retired)
replace tempage=48 if gender==. & tempage==. & retired==0
replace tempage=71 if gender==. & tempage==. & retired==1

mean(tempage)
replace tempage=55 if gender==. & tempage==. & retired==.


gen age_gp=.
replace age_gp=1 if tempage>=18 & tempage<=29
replace age_gp=2 if tempage>=30 & tempage<=39
replace age_gp=3 if tempage>=40 & tempage<=49
replace age_gp=4 if tempage>=50 & tempage<=64
replace age_gp=5 if tempage>=65 & tempage<=105
tab age_gp,m

label define age_gp ///
```

```
         1 "18-29" ///
         2 "30-39" ///
         3 "40-49" ///
         4 "50-64" ///
         5 "65+"

label values age_gp age_gp

*
* EDUCATION ATTAINMENT.
* We separate bachelor and postgraduate qualifications (because we can,
* although that means combining ABS categories) but we combine
* trade/certificate/diploma (because we are uncertain that AuSSA
* categories match ABS ones)


capture drop edimp
gen edimp=.
replace edimp=1 if highed==1
replace edimp=2 if highed==2
replace edimp=3 if inlist(highed,3,4)
replace edimp=4 if inlist(highed,5) & h6==6
replace edimp=5 if inlist(highed,5) & inlist(h6,1,2,3,4,5)

label define edimp 1 "< Year 12", modify
label define edimp 2 "Year 12", modify
label define edimp 3 "Certificate or diploma", modify
label define edimp 4 "Bachelors", modify
label define edimp 5 "Postgraduate", modify
label values edimp edimp


*(1)Give each person with missing education the most common for their
* occupational group (irrespective of tempage)
tab edimp  occupation_l1, col nofreq

* These are the managers (5 changes)
replace edimp=3 if occupation_l1==1 & edimp==.
* These are the professionals (4 changes)
replace edimp=5 if occupation_l1==2 &edimp==.
* These are the technicians until sales workers(22 changes)
replace edimp=3 if inlist(occupation_l1,3,4,5,6) &edimp==.
* These are the lower occupations (8 changes)
replace edimp=1 if inlist(occupation_l1,7,8)    &edimp==.


*(2)For people with unstated education and occupation, give them the
* lowest level of education
replace edimp=2 if  edimp==. & h5==1
replace edimp=1 if  edimp==.

*(3)More treatment of missing data below, after weights are calculated
tab edimp,m



* Set up matrices for population and sample counts
* Rows `i' age_gp=1/5, then gender=1/2
* Cols `j' edimp=1/5
gen rindex=2*(age_gp-1)+gender
gen cindex=edimp

mat input ABS=(277721     479757 419457 168877 24134\ ///
205638 486886 381275 252856 39508\ ///
207646 198392 441204 213950 81514\ ///
200797 227990 359292 307909 105931\ ///
325065 154443 486846 179355 92138\ ///
```

```
406376 200985 376797 228709 102908\ ///
511216 203430 626392 212093 118190\ ///
766073 216740 414506 230665 117222\ ///
506622 122333 372363 87074  43686\ ///
878947 155307 162661 83535  27974)
matrix list ABS


* Construct an equivalent matrix of sample values
tab gender edimp if age_gp==1, matcell(N1)
tab gender edimp if age_gp==2, matcell(N2)
tab gender edimp if age_gp==3, matcell(N3)
tab gender edimp if age_gp==4, matcell(N4)
tab gender edimp if age_gp==5, matcell(N5)

mat R=N1\N2\N3\N4\N5
matrix list R

*Calculate totals of all cells in matrices
sca sumABS=0
sca sumR=0
forval i=1/10 {
forval j=1/5 {
sca sumABS=sumABS+ABS[`i',`j']
sca sumR=sumR+R[`i',`j']
}
}

* Weight is ratio of population relative cell count to sample
* relative cell count
gen weight=.
forval i=1/10 {
forval j=1/5 {
qui        replace        weight=(ABS[`i',`j']/sumABS)/(R[`i',`j']/sumR)        if
rindex==`i'&cindex==`j'&R[`i',`j']~=0
}
}

* More treatment of missing data
* (3)For those with missing gender, give them the average weight for
* their other characteristics
replace rindex=2*(age_gp-1)+1 if gender==.
forval i=1(2)10 {
forval j=1/5 {
replace
weight=((ABS[`i',`j']/sumABS)/(R[`i',`j']/sumR)+(ABS[`i'+1,`j']/sumABS)/(R[`i'+1,`j
']/sumR))/2 if rindex==`i'&cindex==`j'&weight==.
}
}

* Weights should average=1.0000

summ weight
summ weight, detail

*Trim at 1 and 99% percentile
replace weight=5.9 if weight>5.9
replace weight=0.15 if weight<0.15

sum weight
sum weight, detail

egen sumw=sum(weight)
replace weight=weight*1946/sumw
summ weight
label var weight "Sampling weight"
* Display the weights by gender, education and tempage group
```

```
by gender,s: tab weight edimp if age_gp==1
by gender,s: tab weight edimp if age_gp==2
by gender,s: tab weight edimp if age_gp==3
by gender,s: tab weight edimp if age_gp==4
by gender,s: tab weight edimp if age_gp==5

tab age_gp
tab age_gp [aw=weight]
tab gender,m
tab gender [aw=weight],m
tab edimp,m
tab edimp [aw=weight],m


table age_gp edimp, contents(mean weight ) by(gender)
capture drop gender tempage age_gp educ edimp rindex cindex sumw

keep formid weight

save weight.dta, replace
```