sex-bias-in-CSEW-violence / annotated_syntax.md

BJCrim Update annotated_syntax.md 5aedfa9 · 1 hour ago History

Preview Code Blame Raw

Set the working directory, which will contain the data files for 2019/20. For example:

```
cd "C:\...\Changing violence\data and syntax\19-20".
```

Open the non-victim form dataset. This file has one line of data for each respondent. SPSS refers to this file as DataSet2.

```
GET FILE='csew_apr19mar20_nvf.sav'.
DATASET NAME DataSet2.
DATASET ACTIVATE DataSet2.
COMPUTE SAMPYEAR=2019.
```

Get the number of cases by sex.

```
MEANS TABLES=rowlabel BY sex
  /CELLS=COUNT.
WEIGHT BY C11IndivWgt.
```

Get the weighted number of cases by sex -this will be the population values of those in households aged 16 or more.

```
MEANS TABLES=rowlabel BY sex
  /CELLS=COUNT.
WEIGHT OFF.
```

```
+ Male    15505
+ Female  18229
+ Total   33734
```

Now make sure the data set is in the correct order.

```
SORT CASES BY rowlabel.
```

Now open the victim form dataset. This file has multiple lines of data for each respondent - one for each crime report. If the repondent is a non-victim, then then there are no records.

```
GET  FILE='csew_apr19mar20_vf.sav'.
DATASET NAME DataSet1.
DATASET ACTIVATE DataSet1.
```

And make sure that this file is also sorted in the same order as Dataset2.

```
SORT CASES BY rowlabel.
```

Now merge them together, keeping only useful variables.

```
MATCH FILES /FILE=*
  /TABLE='DataSet2'
  /BY ROWLABEL
  /KEEP= ROWLABEL SEX AGE SAMPYEAR PINCID BEFOR99 SUSPEND SAMPTYPE YRINCIB WHERHAPP tocrfc_vf  VINTRO OFFENCE
NSERIES NUMINC NUMBER NUMBER_UNCAPPED
             VIOLNR_VFNOCAP ALLASSAU_VF  VIOLNR_VF VIOLGRP  VIOLGRPNR  tocrfc_vf  tocrfc_vfnocap OFFREL4 OFFREL4A to
OFFREL4R KNEWOFF1 SEENOFF1
             KNEWOFF SEENOFF NUMOFF C11INDIVWGT C11WEIGHTI  WEIGHTI_UNCAP mthinc2 monthid.
EXECUTE.
DATASET CLOSE Dataset2.
FILTER OFF.
USE ALL.
```

The active file is now Dataset1.

We now want to make a new variable **VIOLTYPE** specifyting the type and severity of the violent crime event. We also renove from the active file records where there is no violent act.

```
SELECT IF (offence =11 or offence =12 or offence =13 or offence =21 or offence =32 or offence =33).
RECODE OFFENCE (11,32=1)(12,33=2)(13=3)(21=4) INTO VIOLTYPE.
VALUE LABELS VIOLTYPE 1 'Serious Wounding' 2 'Other Wounding' 3 'Common Assault' 4 'Attempted Assault'.
```

Now we construct some new variables. **NUMBER_UNCAPPED** is the uncapped crime counts per crime event. **NUMBER** is the ONS variable for violent crime counts capped at the 98th percentile, and **NUMINC** is the variable for crime counts capped at 5. We save this data file for future use.

```
COMPUTE NUMINC=NUMBER.
IF(NUMBER>5) NUMINC=5.
FORMATS  VIOLTYPE (F5.0).
EXECUTE.
SAVE OUTFILE='CSEW_19_20_VIOL.sav'/COMPRESSED.
```

We now produce the estimated number of violent crimes in the population, by sex below (capped at 5, capped at 98th percentile, uncapped) using the individual weights stored in the variable **C11IndivWgt**.

```
WEIGHT BY C11IndivWgt.
MEANS TABLES= numinc number number_uncapped BY sex
  /CELLS=SUM.
```

```
+  Sex    Number of incidents cap at 5      Number of incidents (capped at 98%ile)     Uncapped number of incidents
+   Male       679089                                  740048                                   826826
+  Female      468901                                  498524                                   583192
+   Total     1147989                                 1238572                                  1410018
```

Now the weighted number of incidents by *relationship* (capped at 5, capped at 98th percentile, uncapped). Relationship is stored in an ONS variable called **VIOLGRPNR**.

```
MEANS TABLES= numinc number number_uncapped BY violgrpnr
  /CELLS=SUM.
```

```
+ Relationship        Number of incidents cap at 5      Number of incidents (capped at 98%ile)     Uncapped number of i
+ Domestic          188477                            199961                                   212131
+ Stranger          483434                            510624                                   519137
+ Acquaintance      476078                            527987                                   678750
+ Total            1147989                           1238572                                  1410018
```

Next, the weighted number of incidents by *sex* and *relationship* (capped at 5, capped at 98th percentile, uncapped).

```
MEANS TABLES= numinc number number_uncapped BY sex BY violgrpnr
```

```
  /CELLS=SUM.
```

```
+ Sex     Relationship        Number of incidents cap at 5      Number of incidents (capped at 98%ile)     Uncapped num
+_____________________________________________________________________________________________________________________
+  Male   Domestic            47824                             47824                                     47824
+           Stranger             347352                            368636                                    377150
+           Acquaintance         283913                            323588                                    401852
+           Total                679089                            740048                                    826826
+ Female   Domestic            140653                            152137                                    164306
+           Stranger             136082                            141988                                    141988
+           Acquaintance         192166                            204399                                    276898
+           Total                468901                            498524                                    583192
+ Total    Domestic            188477                            199961                                    212131
+           Stranger             483434                            510624                                    519137
+           Acquaintance         476078                            527987                                    678750
+           Total               1147989                           1238572                                   1410018
+
```

Next, the weighted number of incidents by *sex* and *type of violence* (severity) (capped at 5, capped at 98th percentile, uncapped).

```
MEANS TABLES=numinc number number_uncapped BY sex  BY VIOLTYPE
  /CELLS=SUM.
```

```
+ Sex     Type           Number of incidents cap at 5   Number of incidents (capped at 98%ile)     Uncapped number of i
+_____________________________________________________________________________________________________________________
+ Male   Serious Wounding     56947                          56947                                     56947
+          Other Wounding      126641                            159483                                    167996
+          Common Assault      390007                            418125                                    496389
+        Attempted Assault 105494                           105494                                    105494
+          Total               679089                            740048                                    826826
+ Female Serious Wounding     18924                          18924                                     18924
```

```
+          Other Wounding       87845                    105235                    117404
+          Common Assault      297850                    310083                    382582
+          Attempted Assault    64282                     64282                    64282
+          Total               468901                    498524                    583192
+ Total  Serious Wounding     75870                     75870                     75870
+          Other Wounding      214486                    264718                    285401
+          Common Assault      687857                    728209                    878972
+          Attempted Assault 169776                     169776                    169776
+          Total              1147989                   1238572                   1410018
+
```

And, finally, the weighted number of incidents by *sex*, *relationship* and *type of violence* (severity) (capped at 5, capped at 98th percentile, uncapped).

```
MEANS TABLES= numinc number number_uncapped BY sex BY violgrpnr BY VIOLTYPE
  /CELLS=SUM.
```

Report

Sum

Adult number 1 (respondent): Sex	CSEW Type of violence	VIOLTYPE	Number of incidents in series (5 maximum)	Number of valid incidents (after removing those outside England & Wales and outside 12 month reference period)	Uncapped number of valid incidents (after removing those outside England & Wales and outside 12 month reference period)
Male	Domestic	Serious Wounding	1707	1707	1707
		Other Wounding	5606	5606	5606
		Common Assault	31391	31391	31391
		Attempted Assault	9120	9120	9120
		Total	47824	47824	47824
	Stranger	Serious Wounding	36026	36026	36026
		Other Wounding	89672	110956	119470
		Common Assault	173533	173533	173533
		Attempted Assault	48121	48121	48121
		Total	347352	368636	377150
	Acquaintance	Serious Wounding	19214	19214	19214
		Other Wounding	31363	42921	42921
		Common Assault	185084	213202	291465
		Attempted Assault	48252	48252	48252
		Total	283913	323588	401852
	Total	Serious Wounding	56947	56947	56947
		Other Wounding	126641	159483	167996
		Common Assault	390007	418125	496389
		Attempted Assault	105494	105494	105494
		Total	679089	740048	826826
Female	Domestic	Serious Wounding	8782	8782	8782
		Other Wounding	50916	62400	74570
		Common Assault	75072	75072	75072
		Attempted Assault	5882	5882	5882
		Total	140653	152137	164306

	Stranger	Serious Wounding	1635	1635	1635
		Other Wounding	15497	21402	21402
		Common Assault	92013	92013	92013
		Attempted Assault	26938	26938	26938
		Total	136082	141988	141988
	Acquaintance	Serious Wounding	8506	8506	8506
		Other Wounding	21432	21432	21432
		Common Assault	130765	142998	215497
		Attempted Assault	31462	31462	31462
		Total	192166	204399	276898
	Total	Serious Wounding	18924	18924	18924
		Other Wounding	87845	105235	117404
		Common Assault	297850	310083	382582
		Attempted Assault	64282	64282	64282
		Total	468901	498524	583192
Total	Domestic	Serious Wounding	10490	10490	10490
		Other Wounding	56522	68007	80176
		Common Assault	106463	106463	106463
		Attempted Assault	15002	15002	15002
		Total	188477	199961	212131
	Stranger	Serious Wounding	37661	37661	37661
		Other Wounding	105168	132358	140872
		Common Assault	265546	265546	265546
		Attempted Assault	75059	75059	75059
		Total	483434	510624	519137
	Acquaintance	Serious Wounding	27720	27720	27720
		Other Wounding	52795	64353	64353
		Common Assault	315849	356200	506963
		Attempted Assault	79715	79715	79715
		Total	476078	527987	678750
	Total	Serious Wounding	75870	75870	75870
		Other Wounding	214486	264718	285401
		Common Assault	687857	728209	878972
		Attempted Assault	169776	169776	169776
		Total	1147989	1238572	1410018

We now move onto estimating the number ovcitims of violent crime. Data is stored as crime events - each respondent is allowed up to six. So we need to aggregate the data to produce one value per case, disaggregated by any factors under consideration First we need to aggregate by *sex* and *relationship*. a respondent may have multiple violent events carried out by perptrators with differing relationships. VICT is aggregated into VICT_max which takes the value 1 or zero for each sex-relationship combination. Weighted sums of this variable are then produced.

```
DATASET DECLARE datasetagg1.
SORT CASES BY rowlabel sex violgrpnr.
AGGREGATE
  /OUTFILE='datasetagg1'
  /PRESORTED
  /BREAK=rowlabel sex violgrpnr
    /VICT_max=MAX(VICT) / C11IndivWgtAGG1=MAX (C11IndivWgt)/.
DATASET ACTIVATE datasetagg1.
WEIGHT BY C11IndivWgtAGG1.
MEANS TABLES=VICT_max  BY sex  BY violgrpnr
  /CELLS=SUM.
```

```
+  Sex    Relationship  estimated number of victims in population
+   Male    Domestic             33063
+           Stranger            275266
+           Acquaintance        166781
+           Total               475111
+ Female    Domestic             79006
+           Stranger            105707
+           Acquaintance        121905
+           Total               306618
+ Total   Domestic          112069
+           Stranger            380973
+         Acquaintance      288686
+         Total             781728
```

Now we produce the number of victims disaggregated by sex and type of violence (severity). WE are really only interested in the number of victims of serious violence cross-classified by sex.

```
DATASET CLOSE datasetagg1.

DATASET activate dataset1.
DATASET DECLARE datasetagg2.
SORT CASES BY rowlabel  sex violtype.
AGGREGATE
  /OUTFILE='datasetagg2'
  /PRESORTED
  /BREAK=rowlabel sex violtype
    /VICT_max = MAX(VICT) / C11IndivWgtAGG2=MAX (C11IndivWgt)/.
DATASET ACTIVATE datasetagg2.
WEIGHT BY C11IndivWgtAGG2.
MEANS TABLES=VICT_max    BY SEX BY violtype
  /CELLS=SUM.
```

```
+    Sex     Type   estimated number of victims in population
+   Male    Serious Wounding      50301
+           Other Wounding       82823
+           Common Assault      275312
+           Attempted Assault     74307
+           Total                482743
+ Female    Serious Wounding      15214
+           Other Wounding       58136
+           Common Assault      196779
+           Attempted Assault     44578
+           Total                314707
+ Total   Serious Wounding     65515
+           Other Wounding      140959
+           Common Assault      472091
```

```
+           Attempted Assault    118885
+           Total                797450
```

Finally we estimate the number of violent victims by sex.

```
DATASET CLOSE datasetagg2.

DATASET activate dataset1.
DATASET DECLARE datasetagg3.
SORT CASES BY rowlabel  sex.
AGGREGATE
  /OUTFILE='datasetagg3'
  /PRESORTED
  /BREAK=rowlabel sex
    /VICT_max = MAX(VICT) / C11IndivWgtAGG3=MAX (C11IndivWgt)/.
DATASET ACTIVATE datasetagg3.
WEIGHT BY C11IndivWgtAGG3.
MEANS TABLES=VICT_max    BY sex
  /CELLS=SUM.
```

```
+    Sex        estimated number of victims in population
+   Male       462327
+ Female       303849
+  Total       766176
```

We now copy the relevant figures for 2019/20 into the EXCEL spreadsheet. Ther are numerous worksheets - one for eacch type of disaggregation. So worksheet *maledomestic* contains the estimates for male domestic violence. Row 25 contains the estimated totals over 16 years of data (2004/5 to 2019/20). Columns movave3 and movave4 are calculated within the spreadsheet and contain the three-year and four-year moving averages respectively.