GUI Data Workshop

Worksheet 4a - Family Structure -

Infant Cohort - Wave 1 (at 9 months), Wave 2 (at 3 years) and Wave 3 (at 5 years)

This document provides worked examples of some very basic commands which can be used to explore and analyse the GUI data using SPSS drop-down menus. It includes detailed screen shots of how the run the analysis using SPPS menus.

This worksheet is based on the matched Wave 1 (at 9 months), Wave 2 (at 3 years) and Wave 3 (at 5 years) files – please see Information Sheet 4c for details on how to match the files.

Please note this worksheet is based on SPSS Version 22.

Convention A naming – drop down menus¹

Exercise 1: Cross-sectional - Family Structure at Wave 1 (at 9 months)

1. First you will need to weight the data. Select Data \rightarrow Weight Cases

ta Gl	JI Data	_InfantC	ohortWa	ave1Wave2W	/ave3.sav.sa	av [DataSet]	1] - IBM :	SPSS S	Statistic	s Data	Edito
<u>F</u> ile	<u>E</u> dit	View	<u>D</u> ata	Transform	Analyze	<u>G</u> raphs	Utilities	Ad	d- <u>o</u> ns	Wind	ow
			🧔 Def	ine <u>V</u> ariable F	Properties			1	*,		
			🔏 Set	Measuremer	nt Level for	Unknown					
1		1	ing Cot	y Data Prope	erties			-	Pa	artnor	
			Nev 📄	v Custom Att	ri <u>b</u> ute				10	arunei	
1			🗟 Def	ine Dates				4.92			.00
2	2		🔡 Def	ine <u>M</u> ultiple R	esponse S	ets		4.55		1	.00
3	3		🔡 Ider	ntify D <u>u</u> plicate	cases			8.14		1	.00
4	ł		🔁 Cor	npare Datase	ets			5.14		1	.00
5	5		Sor	t Cases				4.21			.00
6	5		Sor	t Variables				8.37		1	.00
7	·			vana <u>b</u> ies				6.51		1	.00
8	3			nspose				1.99			.00
9)		Mer	rge Files			•	1.61		1	.00
1	0		Res	structure				2.32		1	.00
1	1		🔁 Agg	regate				1.05			.00
1:	2		Rak	e Weights				8.41		1	.00
1	3		Pro	pensity Score	e Matching.			1.61		1	.00
1.	4		Cas	e Control Ma	tching			1.38		1	.00
1	5		🔂 Cor	v Dataset	-			.96			.00
1	6		Soli	t into Files				1.21		1	.00
1	7			+ File				4.59		1	.00
1	8		= Spli	t <u>F</u> ile				8.29		1	.00
1	9		E Sele	ect Cases				1.19		1	.00
2	0		<u>ी W</u> e	ight Cases				6.5 2			.00
2	1		2100.0	0	.89			5.92		1	.00
2	2		2200.0	0	.68			4.49			.00

¹ See Variable Naming Conventions And Longitudinal Data Dictionary For Wave 1 And Wave 2 Of The Infant Cohort Of Growing Up In Ireland document for further details on variable naming conventions.
 Worksheet 4a – Family Structure – Infant Cohort - Wave 1 (at 9 months), Wave 2 (at 3 years) and Wave 3 (at 5 years)

2. The following dialog box will appear.

ta Weight Cases	
 ✓ ID ✓ WGT_9MTH ✓ GROSS_9MTH ✓ Partner ✓ Int_type MMA4 MMa5ap1 ✓ MMagep1 	 ● Do not weight cases ○ Weight cases by Frequency Variable: ○ Weight cases ○ Event Status: Do not weight cases ○ Reset Cancel Help

3. Click on 'Weight cases by'. Browse through the list of variables on the left and highlight the one you want ('*WGT_9MTH'*) and click on the arrow in the middle. Click on 'OK'. Note that this is a cross-sectional analysis of the 11,134 cases in Wave 1 so we use the Wave 1 weight.

ta Weight Cases	X
 ✓ ID ✓ GROSS_9MTH ✓ Partner ✓ Int_type ✓ MMA4 ✓ MMa5ap1 ✓ MMa5rmp1 ✓ OK Paste 	○ Do not weight cases ⑨ Weight cases by Frequency Variable: ✓ WGT_9MTH Current Status: Do not weight cases Reset Cancel Help

ta *G	🔚 *GUI Data_InfantCohortWave1Wave2Wave3.sav.sav [DataSet1] - IBM SPSS Statistics Data Editor										
File	<u>E</u> dit	View	<u>D</u> ata	Transform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	Add- <u>o</u> n	s <u>W</u> indow	Help	
				IC 3	Repor	ts		•	1	- A	
<u> </u>		N <u>-</u> Y			D <u>e</u> scr	iptive Stati	istics	•	123 <u>F</u> requencie	es	
				WOT	Ta <u>b</u> le:	S		•	Descriptive	es	tuno
			ID.	wgi_	Co <u>m</u> p	are Means	6	•	4 Explore		type
1			100.0	00	<u>G</u> ener	al Linear N	lodel	•	<u>₩</u> Crosstabs		2
2	2		200.0	00	Gener	alized Line	ear Models	•	TURF Ana	lysis	1
3	;		300.0	00	Mixed	Models		•	🔽 Ratio		1
4	ļ		400.0	00	<u>C</u> orre	late		•	P-P Plots		1
5	5		500.0	00	<u>R</u> egre	ssion			Q-Q Plots		2
6	;		600.0	00	L <u>o</u> glin	ear			1.00		1
7	<u> </u>		700.0	00	Class	i <u>t</u> y 			1.00		3
8	}		800.0	00	Dimer	nsion Redu	uction	•	.00		2
9)		900.0	00	Sc <u>a</u> le			•	1.00		1
1	0		1000.0	00	<u>N</u> onpa	arametric 1	Fests	•	1.00		3
1	1		1100.0	00	Forec	asting		•	.00		2
1	2		1200.0	00	<u>S</u> urviv	al		•	1.00		1
1	3		1300.0	00	M <u>u</u> ltip	le Respon	se	•	1.00		1

4. To run the frequency, select Analyse \rightarrow Descriptive Statistics \rightarrow Frequencies

5. The following dialog box will appear:



6. Browse through the list of variables on the left and highlight the one(s) you want

Frequencies						
 ✓ ID ✓ WGT_9MTH ✓ GROSS_9MTH ✓ Int_type MMA4 MMa5ap1 ✓ MMagep1 	Variable(s): ier	Charts Charts Format Style			
✓ Display frequency tables OK Paste Reset Cancel Help						

('partner' – partner in household) and click on the arrow in the middle.

7. Click 'OK' and the output will show you the frequency table for that variable.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 No partner	1645	14.8	14.8	14.8
	1.00 Has partner	9489	85.2	85.2	100.0
	Total	11134	100.0	100.0	

Partner Partner in household

Exercise 2: Cross-sectional - Family Structure at Wave 2 (at 3 years)

First you will need to weight the data, using the Wave 2 weight this time.

- 1. Data \rightarrow Weight Cases \rightarrow WGT_3YR
- 2. Analyse \rightarrow Descriptive Statistics \rightarrow Frequencies \rightarrow b2_partner

A weighted frequency of this will give us the following output:

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 no	1418	14.5	14.5	14.5
	1 yes	8375	85.5	85.5	100.0
	Total	9793	100.0	100.0	

b2_partner PCG has spouse/partner living in the hsd at Wave 2

Note that although this frequency is run on the full matched file of 11,134 cases, results are only shown for the 9,793 Wave 2 cases. You will see the following warning in the output window which refers to this (if you have set SPSS to display a log. To do this, click Edit \rightarrow Options \rightarrow Viewer \rightarrow Item \rightarrow Log \rightarrow Contents are initially Shown):

>Warning # 3211

On at least one case, the value of the weight variable was zero, negative, or missing. Such cases are invisible to statistical procedures and graphs which need positively weighted cases, but remain on the file and are processed by non-statistical facilities such as LIST and SAVE.

Exercise 3: Cross-sectional - Family Structure at Wave 3 (at 5 years)

First you will need to weight the data, using the Wave 3 weight this time. Note we want to analyse all 9,001 families that participated in Wave 3 (not the subset that participated at all three waves) so we use '*WGT_5Yra*'. See Information Sheet 03c for more information on the weights at Wave 3.

1. Data \rightarrow Weight Cases \rightarrow WGT_5YRa

- 2. Analyse \rightarrow Descriptive Statistics \rightarrow Frequencies \rightarrow b3_partner
- A weighted frequency of this will give us the following output:

b3_partner PCG has spouse/partner living in the hsd at Wave 3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00 No	1259	14.0	14.0	14.0
	1.00 Yes	7742	86.0	86.0	100.0
	Total	9001	100.0	100.0	

Exercise 4: Changes in Family Structure across Waves

The cross-sectional analyses above show the proportions of children in one and two-parent households for each wave. The results are summarized in Figure 1 below. A total of 14.8% of 9 month olds were in one parent households, 14.5% of 3 year old children were in one parent households and 14% of 5 year old children were in one parent households.





From these cross-sectional analyses it appears that family structure is reasonably stable across waves with similar proportions of children in one parent households at each wave. However, these are separate analyses and do not tell us anything about changes in family structure from wave to wave.

To look at changes in family structure between Wave 1 and Wave 3, we need to do a crosstabulation of the two variables in each wave. In order to do this, we will be analysing only the 9,001 cases who responded in both Wave 1 and Wave 3, and using the Wave 3 a weight (*'WGT_5Yra'*). The Wave 3 a weight adjusts the data to make it representative of all children who were resident in Ireland at Wave 1 and who continue to be resident in Ireland

at Wave 3 (accounting for those who no longer live in Ireland at 5 years of age or who have deceased since 9 months of age).

1. Data \rightarrow Weight Cases \rightarrow WGT_5YRa

2. Analyse \rightarrow Descriptive Statistics \rightarrow Crosstabs

U *O	GUI Dat	a_Infant	tCohortV	Vave1Wave2V	Vave3.sav.s	av [DataSe	et1] - IBM S	SPSS Sta	tistics Data Editor	
File	<u>E</u> dit	View	<u>D</u> ata	<u>T</u> ransform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	Add- <u>o</u> r	ns <u>W</u> indow <u>H</u> el	р
				IC 31	Repor	ts		•	🖁 🕎 🔜 🌢	2
<u> </u>		- <u>1</u>			D <u>e</u> scr	iptive Stati	istics	•	Erequencies	
			10	MOT	Tables	S		•	Descriptives	
			ID	wgi_	Co <u>m</u> p	are Means	5	•	A Explore	typ
1	1		100 (0	Gener	al Linear M	lodel	•	Crosstabs	
	2		200.0	00	Gener	alized Line	ear Models	*	TURF Analysis	
3	3		300.0	00	Mixed	Models		•	Ratio	
4	1		400.0	00	<u>C</u> orrel	ate		•	P-P Plots	
5	5		500.0	00	<u>R</u> egre	ssion		•		
6	6		600.0	00	L <u>o</u> glin	ear		•		
7	7		700.0	00	Class	i <u>f</u> y		•	1.00	
8	3		800.0	00	<u>D</u> imer	nsion Redu	uction	•	.00	
ç)		900.0	00	Sc <u>a</u> le			•	1.00	
1	0		1000.0	00	<u>N</u> onpa	arametric 1	Fests	•	1.00	
1	1		1100.0	00	Forec	asting		•	.00	
1	2		1200.0	00	<u>S</u> urviv	al		•	1.00	
1	3		1300.0	00	M <u>u</u> ltipl	e Respon	se	•	1.00	

3. The following dialog box will appear.

ta Crosstabs		X
 ID WGT_9MTH GROSS_9MTH Partner Int_type MMA4 MMa5ap1 MMa5prp1 MMa5rrp1 MMa5pesp1 MMa5pesp1 MMa5ap2 	Row(s): Column(s): Layer 1 of 1 Previous Next	Statistics Cells Eormat Style
Display clustered <u>b</u> ar charts Suppress <u>t</u> ables OK	Display layer variables in table layers Paste Reset Cancel Help	5

4. Browse and highlight the variables you want to cross-reference and move them to the row and column boxes as applicable. Move 'partner' to the 'Rows' box and 'b3_partner' to the 'Columns' box.

ta Crosstabs		X
 b2_scgsens b2_scgsens wave2 xxwave1 xxwave2 xxwave3 WGT_5YRa GROSS_5YRa WGT_5YRb GROSS_5YRb xpcgstatph3 xscgstatph3 bpc3A3a 	Row(s): Partner Column(s): b3_partner Layer 1 of 1 Previous Next	Statistics C <u>e</u> lls Format Sty <u>l</u> e
Display clustered <u>b</u> ar charts Suppress <u>t</u> ables OK	Display layer variables in table layers Paste Reset Cancel Help	3

5. Select 'Cells' and the following dialog box will appear

ta Cros	Cros Crosstabs: Cell Display							
€ b2	Counts Observed Expected Hide small counts Less than 5	z-test Compare column proportions Adjust p-values (Bonferroni m	ethod)	stics ells mat yle				
 ✓ W(✓ GF ✓ W(✓ GF ✓ GF ✓ xpc 	Percentages	Residuals Unstandardized Standardized Adjusted standardized						
Sup	Noninteger Weights	© Round case <u>w</u> eights s						
	Conti	nue Cancel Help						

6. Under 'Percentages', click row and total



7. Click 'Continue' and then 'OK' and you will get the following output:

			b3_partner PCG has spouse/partner living in the hsd at Wave 3		
			.DO No	1.00 Yes	Total
Partner Partner in household	.00 No partner	Count	A 881	412	1293
		% within Partner Partner in household	68.1%	B 31.9%	109.0%
		% of Total	9.8%	4.6%	C 14.4%
	1.00 Has partner	Count	377	7330	707
		% within Partner Partner in household	4.9%	95.1%	100.0%
		% of Total	4.2%	81.4%	85.6%
Total		Count	1258	7742	9000
		% within Partner Partner in household	14.0%	86.0%	100.0%
		% of Total	14.0%	86.0%	100.0%

Partner Partner in household * b3_partner PCG has spouse/partner living in the hsd at Wave 3 Crosstabulation

Note:

- A. The first row of information shows the number of children in each of the cells. For e.g. 881 children were in one parent households at Wave 1 and also at Wave 3
- B. The second row of information shows the percentage of children in each one parent household status category at Wave 1 who are in each of the Wave 3 one parent household status categories. For e.g. 31.9% of children who were in one parent households at Wave 1 were in two parent households at Wave 3.
- C. The third row of information shows the percentage of all children in each of the cells. For e.g. the category of children who were in one parent households at Wave 1 and were in two parent households at Wave 3 accounts for 4.6% of all children.

8. A new variable can be created to classify children according to their status with regards to changes in family structure between Wave 1 and Wave 3. The easiest way to do this is using SPSS syntax and the code to do so is displayed below.

COMPUTE famstructure = 0. IF partner = 0 and b3_partner = 0 famstructure = 1. IF partner = 0 and b3_partner = 1 famstructure = 2. IF partner = 1 and b3_partner = 0 famstructure = 3. IF partner = 1 and b3_partner = 1 famstructure = 4. VARIABLE LABELS famstructure "Changes in family structure from Wave 1 to Wave 3". VALUE LABELS famstructure 1 "One-parent at 9 mths / One-parent at 5 yrs" 2 "One-parent at 9 mths / One-parent at 5 yrs" 3 "Two-parent at 9 mths / One-parent at 5 yrs" 4 "Two-parent at 9 mths / Two-parent at 5 yrs". FREQUENCIES famstructure.

These changes in family structure from Wave 1 to Wave 3 are summarized in Figure 2 below. Although the previous cross-sectional results shown in Figure 1 above suggest that there is little variation in family structure across waves, the longitudinal results in Figure 2 show that 4.6% of children who were in one parent households at Wave 1 were in two parent households at Wave 3 and 4.2% of those in two-parent households at Wave 1 were in one-parent households at Wave 3. These changes in family structure may be an important consideration when looking at outcomes at Wave 3.



Figure 2: Changes in Family Structure from Wave 1 to Wave 3

9. As just one example of how children's outcomes may vary according to changes in family structure across waves we can look at the mean SDQ Total Difficulties score at 5 years across each of the four categories of family structure.

🔄 *GUI Data_InfantCohortWave1Wave2Wave3.sav.sav [DataSet1] - IBM SPSS Statistics Data Editor														
File	<u>E</u> dit	View	<u>D</u> ata	Transform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	Add-on	ns <u>W</u> ir	ndow	<u>H</u> elp			
					Repor	ts intius Stati	otia o) k				1		2
				Name	Descr	iptive Statis	SUCS	ŗ	Deci	mals	L	abel		Valu
16	92	bsc3J5	5 a 2		l ables	6		•	0		15a A	llow par	11	Yes)
16	93	bsc3.J5	5h2		Co <u>m</u> p	are Means		•	Mea	ns				(es)
16	Q <u>/</u>	hsc 3.14	5i2		Gener	al Linear M	odel	•	1 One	-Sam	ole T Te	st		
16	05	bec31	512		Gener	ali <u>z</u> ed Line	ar Models	•	lnde	nende	nt-Sami	oles T Tes	ł	(00)
16	30 806	bcc2 15k2			Mixed	Models		•	Deired Complete T Test			Cocl		
	90	DSC3J3KZ		Correlate		•	maired-Samples T Test				esj			
16	i97	bsc3J6		<u>_</u>	<u>o</u> onolato			🔄 One-Way ANOVA				/ery		
16	698	bsc3J7	7		Regre	ssion		P 1	0		J7. Su	pervisio	. {1,	Yes}
16	699	bsc3J8	3		L <u>o</u> gline	ear		•	0		J8. Nu	mber of	. {15	, 11
17	'00	bsc3J9)		Classi	fy		•	0		J9. Nu	mber of	. {10	, 6 tc
17	'01	bsc3j1	Dacre		Dimer	ision Redu	ction	•	0		J10. A	cres far	. {50	, 50
17	/02	bsc3J11		Sc <u>a</u> le			•	0		J11. S	CG ever.	. {1,	Yes}	
			~			· ·			-					

10. Analyze \rightarrow Compare Means \rightarrow Means

11. Move 'b3_sdqtotaldiffs' into the Dependent List box and move 'famstructure' into the Independent List box. Click 'ok'.

ta Means	X			
Ø ID Ø WGT_9MTH Ø GBOSS_9MTH	Dependent List:			
 ✓ Partner ✓ Int_type MMA4 MMa5ap1 ✓ MMagep1 	Layer 1 of 1 Previous Independent List:			
MMa5rmp1 MMa5rcp1 OK Paste Reset Cancel Help				

12. This will give the following results.

Report

b3_sdqtotaldiffs SDQ Total difficulties score - Wave 3

famstructure Changes in family structure from Wave 1 to Wave 3	Mean	N	Std. Deviation	
1.00 One-parent at 9 mths / One-parent at 5 yrs	9.8342	879	5.59772	
2.00 One-parent at 9 mths / Two-parent at 5 yrs	10.1501	412	5.83796	
3.00 Two-parent at 9 mths / One-parent at 5 yrs	8.4332	377	5.03921	
4.00 Two-parent at 9 mths / Two-parent at 5 yrs	6.9493	7327	4.60224	
Total	7.4399	8995	4.90551	

These results are summarized in Figure 3 below which shows the mean SDQ Total Difficulties score of 5 year old children across the four categories of changes in family structure between Wave 1 and Wave 3. We can see from this that the lowest scores (fewest problems) are experienced by children in stable two parent families (two parent at Wave 1 and Wave 3). The highest scores (most problems) are experienced by those who change from being in a one parent family at Wave 1 to a two parent family at Wave 3. Obviously, more investigation is needed into this before conclusions could be made but it nicely demonstrates the value of using longitudinal data to examine children's outcomes.



Figure 3: SDQ Total Difficulties score by changes in family structure from Wave 1 to Wave 3

Convention A naming – SPSS syntax

WEIGHT by WGT_9MTH. FREQUENCIES partner.

WEIGHT by WGT_3YR. FREQUENCIES b2_partner.

WEIGHT by WGT_5YRa. FREQUENCIES b3_partner.

WEIGHT by WGT_5YRa. CROSSTABS TABLES partner by b3_partner / CELLS count row total.

COMPUTE famstructure = 0. IF partner = 0 and b3_partner = 0 famstructure = 1. IF partner = 0 and b3_partner = 1 famstructure = 2. IF partner = 1 and b3_partner = 0 famstructure = 3. IF partner = 1 and b3_partner = 1 famstructure = 4. VARIABLE LABELS famstructure "Changes in family structure from Wave 1 to Wave 3". VALUE LABELS famstructure 1 "One-parent at 9 mths / One-parent at 5 yrs" 2 "One-parent at 9 mths / Two-parent at 5 yrs" 3 "Two-parent at 9 mths / One-parent at 5 yrs" 4 "Two-parent at 9 mths / Two-parent at 5 yrs". FREQUENCIES famstructure.

MEANS TABLES b3_sdqtotaldiffs by famstructure.

Convention B naming

To conduct the above analysis using the Convention B naming files, replace the following

variables:

Convention A naming	Convention B naming
WGT_9MTH	azwg01
WGT_3YR	bzwg01
WGT_5YRa	czwg01
partner	azID04
b2_partner	bzID04
b3_partner	czID04
b3_sdqtotaldiffs	cdED20f