



```

name: <unnamed>
log: C:\data\AsianBaro\logofcleaningforEntropy.smcl
log type: smcl
opened on: 10 Sep 2024, 15:52:07

```

```

1 . *do file for Trial 1 of Entropy with Ordinal Variables
2 . *filename Trialslordinal.do
3 . * Wendy Olsen
4 . * grateful thanks to Mr Ziyang Zhou - Univ. of Manchester
5 . * Univ of Manchester 2024
6 .
7 . * Stata 18
8 . ***Data cleaning***
9 . cd "C:\data\AsianBaro"
C:\data\AsianBaro
10.
11. *Part 2: FURTHER CLEANING
12. *Clean variables in the AB data for 2019 India
13. * make income have ten bins, ordinal cuts on a continuous variable.
14.
15. *Objective: Entropy and Regression Analysis (conferences 2024)
16. *Wendy Olsen
17. * Univ. of Manchester 2024
18.
19. * file origins. First, in \raw\ directory, the AsianBaro2019.dta was revised to giv
> e clean income. This file was named AsianBaro2019rev1.dta. I have omitted this step
> and placed the cleaning codes here, below.
20. * Then, in R the encoding took place and entropy was measured & compared.
21. * Now, we are making explicit all the coding so that regression can be done by anybo
> dy, using this file alone.
22.
23. use "C:/data/AsianBaro/raw/AsianBaro2019.dta", clear
24. merge 1:1 IDnumber using "data\tmpAgeEduc.dta", update replace

```

Result	Number of obs	
Not matched	0	
Matched	5,318	
not updated	5,318	(_merge==3)
missing updated	0	(_merge==4)
nonmissing conflict	0	(_merge==5)

```

25.
26. *Aims 1. clean raw income, and Aim 2 clean the raw opinion variables. Output the da
> taset in a well named form.
27. tab sel4

```

sel4 What group your household income on average is?	Freq.	Percent	Cum.
Lowest quintile	1,293	24.31	24.31
2 nd	985	18.52	42.84
3rd	2,063	38.79	81.63
4th	353	6.64	88.27
Highest quintile	20	0.38	88.64
Decline to answer	604	11.36	100.00
Total	5,318	100.00	

28. tab SE14a

se14a Does the total income of your household allow you to satisfactorily cover y	Freq.	Percent	Cum.
Not applicable	50	0.94	0.94
Our income covers the needs well, we ca	412	7.75	8.69
Our income covers the needs well, we ca	890	16.74	25.42
Our income covers the needs all right,	1,501	28.22	53.65
Our income does not cover the needs, th	1,464	27.53	81.18
Our income does not cover the needs, th	882	16.59	97.76
Do not understand the question	35	0.66	98.42
Decline to answer	84	1.58	100.00
Total	5,318	100.00	

29. tab Se14, nol

se14 What group your household income on average is?	Freq.	Percent	Cum.
1	1,293	24.31	24.31
2	985	18.52	42.84
3	2,063	38.79	81.63
4	353	6.64	88.27
5	20	0.38	88.64
9	604	11.36	100.00
Total	5,318	100.00	

30. tab SE14a, nol

se14a Does the total income of your household allow you to satisfactorily cover y	Freq.	Percent	Cum.
0	50	0.94	0.94
1	412	7.75	8.69
2	890	16.74	25.42
3	1,501	28.22	53.65
4	1,464	27.53	81.18
5	882	16.59	97.76
7	35	0.66	98.42
9	84	1.58	100.00
Total	5,318	100.00	

31. *check education as it arrives to us

32. tab SE5

se5 Education	Freq.	Percent	Cum.
No formal education	870	16.36	16.36
Incomplete primary/elementary	375	7.05	23.41
Complete primary/elementary	547	10.29	33.70
Incomplete secondary/high school: techn	336	6.32	40.02
Complete secondary/high school: technic	587	11.04	51.05
Incomplete secondary/high school	288	5.42	56.47
Complete secondary/high school	794	14.93	71.40
Some university education	319	6.00	77.40
University education completed	647	12.17	89.56
Post-graduate degree	191	3.59	93.16

Decline to answer	364	6.84	100.00
Total	5,318	100.00	

33. *and check the main source after recodes were done in R.

34. tab edu1_1 edu2_1

edu1_1	edu2_1	Total
0	3,709	3,709
1	1,609	1,609
Total	5,318	5,318

35. tab edu1_4 edu2_5

edu1_4	edu2_5	Total
0	2,780	3,937
1	1,381	1,381
Total	4,161	5,318

36. tab edu1_5 edu2_3

edu1_5	edu2_3	Total
0	2,156	4,161
1	0	1,157
Total	2,156	5,318

37. *the conclusion of the recoding is later on in this file.

38.

39. gen income = 5

40. *note, there are 92 cases which had missing Se14 and also no substantive answer to S > E14a so these were coded to the category 5 income.

41. replace income = 1 if Se14==1
(1,293 real changes made)

42. replace income=3 if Se14==2
(985 real changes made)

43. replace income=5 if Se14==3
(0 real changes made)

44. replace income=7 if Se14==4
(353 real changes made)

45. replace income=9 if Se14==5
(20 real changes made)

46. *Note in Se14, the 9 refers to "Decline to Answer", 11% of respondents.

47. *Note in SE14a, the subjective options are:

48. *#1=Our income covers the needs well, we can save a lot.

49. *#2=Our income covers the needs well, we can save.

50. *#3=Our income covers the needs all right, without much difficulties.
 51. *#4=Our income does not cover the needs, there are difficulties.
 52. *#5=Our income does not cover the needs, there are great difficulties.
 53. *#0=Not applicable.
 54. tab Sel4 SE14a, col row cell

Key
<i>frequency</i>
<i>row percentage</i>
<i>column percentage</i>
<i>cell percentage</i>

sel4 What group your household > actorily income on average		sel4a Does the total income of your household allow you to satis cover y Our income Our income Our income Our income Our income						
is?		Not appli	Our incom	Our incom	Our incom	Our incom	Our incom	Our incom
> Do not un		Total						
Lowest quintile			5	52	106	240	458	426
> 1	1,293							
> 0.08	100.00	0.39	4.02	8.20	18.56	35.42	32.95	
> 2.86	24.31	10.00	12.62	11.91	15.99	31.28	48.30	
> 0.02	24.31	0.09	0.98	1.99	4.51	8.61	8.01	
2 nd			1	30	122	317	342	165
> 1	985							
> 0.10	100.00	0.10	3.05	12.39	32.18	34.72	16.75	
> 2.86	18.52	2.00	7.28	13.71	21.12	23.36	18.71	
> 0.02	18.52	0.02	0.56	2.29	5.96	6.43	3.10	
3rd			5	253	449	718	462	150
> 6	2,063							
> 0.29	100.00	0.24	12.26	21.76	34.80	22.39	7.27	
> 17.14	38.79	10.00	61.41	50.45	47.83	31.56	17.01	
> 0.11	38.79	0.09	4.76	8.44	13.50	8.69	2.82	
4th			3	34	76	96	74	17
> 25	353							
> 7.08	100.00	0.85	9.63	21.53	27.20	20.96	4.82	
> 71.43	6.64	6.00	8.25	8.54	6.40	5.05	1.93	
> 0.47	6.64	0.06	0.64	1.43	1.81	1.39	0.32	
Highest quintile			0	7	6	5	2	0
> 0	20							
> 0.00	100.00	0.00	35.00	30.00	25.00	10.00	0.00	
> 0.00	0.38	0.00	1.70	0.67	0.33	0.14	0.00	
> 0.00	0.38	0.00	0.13	0.11	0.09	0.04	0.00	

Decline to answer	36	36	131	125	126	124
> 2	604					
> 0.33	100.00	5.96	5.96	21.69	20.70	20.53
> 5.71	11.36	72.00	8.74	14.72	8.33	8.61
> 0.04	11.36	0.68	0.68	2.46	2.35	2.37
Total	50	412	890	1,501	1,464	882
> 35	5,318	0.94	7.75	16.74	28.22	27.53
> 0.66	100.00	100.00	100.00	100.00	100.00	100.00
> 100.00	100.00	0.94	7.75	16.74	28.22	27.53
> 0.66	100.00					

sel4 What group your household income on average is?	sel4a Does the total income of your household allow you to satisfactor ily cover y Decline t	Total
Lowest quintile	5 0.39 5.95 0.09	1,293 100.00 24.31 24.31
2 nd	7 0.71 8.33 0.13	985 100.00 18.52 18.52
3rd	20 0.97 23.81 0.38	2,063 100.00 38.79 38.79
4th	28 7.93 33.33 0.53	353 100.00 6.64 6.64
Highest quintile	0 0.00 0.00 0.00	20 100.00 0.38 0.38
Decline to answer	24 3.97 28.57 0.45	604 100.00 11.36 11.36
Total	84 1.58 100.00 1.58	5,318 100.00 100.00 100.00

55. *#7=Do not understand the question. <1% of respondents, but they gave an income question
> tile answer.
56. *#9=Decline to answer. <2% of respondents --> income is coded '5' for 24 of these, i
> e 0.45% of sample.
57. replace income=2 if Sel4==9& SE14a==5
(124 real changes made)
58. replace income=4 if Sel4==9& SE14a==4
(126 real changes made)
59. replace income=6 if Sel4==9& SE14a==3
(125 real changes made)
60. replace income=8 if Sel4==9& SE14a==2
(131 real changes made)
61. replace income=10 if Sel4==9 & SE14a==1
(36 real changes made)
62. tab income SE14a, nol

		se14a Does the total income of your household allow you to satisfactorily cover your income						
		9	0	1	2	3	4	5
		Total						
	1	5	52	106	240	458	426	
> 1	2	5	1,293	0	0	0	0	124
> 0	3	0	124	30	122	317	342	165
> 1	4	7	985	0	0	0	126	0
> 0	5	0	126	0	0	0	126	0
> 8	6	41	2,125	253	449	718	462	150
> 0	7	0	125	0	0	125	0	0
> 0	8	0	125	0	0	125	0	0
> 25	9	28	353	34	76	96	74	17
> 0	10	0	131	0	131	0	0	0
> 0		0	20	7	6	5	2	0
> 0		0	36	36	0	0	0	0
Total		84	5,318	412	890	1,501	1,464	882

63. tab income

income	Freq.	Percent	Cum.
1	1,293	24.31	24.31
2	124	2.33	26.65
3	985	18.52	45.17
4	126	2.37	47.54
5	2,125	39.96	87.50
6	125	2.35	89.85
7	353	6.64	96.48
8	131	2.46	98.95
9	20	0.38	99.32
10	36	0.68	100.00
Total	5,318	100.00	

```
64. tab income Se14, nol
```

income		What group your household income on average is?					To	
> tal		1	2	3	4	5	9	
1		1,293	0	0	0	0	0	1,
> 293		2	0	0	0	0	124	
> 124		3	0	985	0	0	0	
> 985		4	0	0	0	0	126	
> 126		5	0	0	2,063	0	62	2,
> 125		6	0	0	0	0	125	
> 125		7	0	0	0	353	0	
> 353		8	0	0	0	0	131	
> 131		9	0	0	0	20	0	
> 20		10	0	0	0	0	36	
> 36								
Total		1,293	985	2,063	353	20	604	5,
> 318								

```
65. tab income SE5, row
```

Key
<i>frequency</i> <i>row percentage</i>

Age	Income	No. of university graduates	se5 Education						
			formal Total	Incomplet Total	Complete	Incomplet	Complete	Incomplet	Complete
> 45	1	30	370	142	148	85	110	70	1
> 21	2.32		28.62	10.98	11.45	6.57	8.51	5.41	11.45
			100.00						
> 14	2	1	56	8	19	6	7	4	
> 29	0.81		45.16	6.45	15.32	4.84	5.65	3.23	11.45
			100.00						
> 41	3	61	169	68	123	87	110	44	1
> 31	6.19		17.16	6.90	12.49	8.83	11.17	4.47	14.29
			100.00						
> 19	4	4	28	18	19	7	9	1	
> 08	3.17		22.22	14.29	15.08	5.56	7.14	0.79	15.32
			100.00						
> 03	5	147	177	85	187	133	280	118	4
			2,125						

> 96	6.92	8.33 100.00	4.00	8.80	6.26	13.18	5.55	18.
> 17	6	10	19 125	10	9	10	14	9
> 60	8.00	15.20 100.00	8.00	7.20	8.00	11.20	7.20	13.
> 37	7	54	20 353	24	28	3	42	28
> 48	15.30	5.67 100.00	6.80	7.93	0.85	11.90	7.93	10.
> 12	8	10	27 131	13	10	4	10	11
> 16	7.63	20.61 100.00	9.92	7.63	3.05	7.63	8.40	9.
> 4	9	0	1 20	1	2	1	2	1
> 00	0.00	5.00 100.00	5.00	10.00	5.00	10.00	5.00	20.
> 2	10	2	3 36	6	2	0	3	2
> 56	5.56	8.33 100.00	16.67	5.56	0.00	8.33	5.56	5.
> 94	Total	870 5,318	375	547	336	587	288	7
> 93	6.00	16.36 100.00	7.05	10.29	6.32	11.04	5.42	14.

income	se5 Education			Total
	Universit	Post-grad	Decline t	
1	47 3.63	12 0.93	134 10.36	1,293 100.00
2	2 1.61	2 1.61	5 4.03	124 100.00
3	85 8.63	20 2.03	77 7.82	985 100.00
4	10 7.94	7 5.56	4 3.17	126 100.00
5	387 18.21	104 4.89	104 4.89	2,125 100.00
6	17 13.60	5 4.00	5 4.00	125 100.00
7	82 23.23	21 5.95	14 3.97	353 100.00
8	10 7.63	15 11.45	9 6.87	131 100.00
9	6 30.00	1 5.00	1 5.00	20 100.00
10	1 2.78	4 11.11	11 30.56	36 100.00

Total	647	191	364	5,318
	12.17	3.59	6.84	100.00

```

66. *To confirm, high decile income is higher income household.
67. *And high educ response so far is indicated only in categoricals.
68.
69.
70. *comment, this income distribution rises from lowest decile 1 to highest 10.
71.
72. *there are a high percent of people declaring in the lowest 3 deciles (42%, raw)
73.
74.
75. label variable income "Household Income Decile - Subjective"

76. *aim: set up incl etc. so that they are cumulative coding of income. And on the co
    > nverse, i.income is the distinct coding of income.
77. gen incl = 0

78. gen inc2=0

79. gen inc3=0

80. gen inc4=0

81. gen inc5=0

82. gen inc6 = 0

83. gen inc7=0

84. gen inc8=0

85. gen inc9=0

86. gen inc10=0

87. replace incl=1 if income==1
    (1,293 real changes made)

88. replace inc2=1 if income==2
    (124 real changes made)

89. replace inc3=1 if income==3
    (985 real changes made)

90. replace inc4=1 if income==4
    (126 real changes made)

91. replace inc5=1 if income==5
    (2,125 real changes made)

92. replace inc6=1 if income==6
    (125 real changes made)

93. replace inc7=1 if income==7
    (353 real changes made)

94. replace inc8=1 if income==8
    (131 real changes made)

```

95. replace inc9=1 if income==9
(20 real changes made)

96. replace inc10=1 if income==10
(36 real changes made)

97. tab inc10 income

		Household Income Decile - Subjective							
		8	1	2	3	4	5	6	
		Total							
> 7	inc10								
> 53	0	131	1,293	124	985	126	2,125	125	3
> 53	1	131	5,282	0	0	0	0	0	
> 0		0	36						
> 53	Total	131	1,293	124	985	126	2,125	125	3
> 53		131	5,318						

		Household Income Decile - Subjective		Total
		9	10	
inc10	0	20	0	5,282
	1	0	36	36
Total		20	36	5,318

98. *Cumulative coding of income (scheme 2)

99. gen inc2_1 = 0

100 gen inc2_2=0

101 gen inc2_3=0

102 gen inc2_4=0

103 gen inc2_5=0

104 gen inc2_6 = 0

105 gen inc2_7=0

106 gen inc2_8=0

107 gen inc2_9=0

108 gen inc2_10=0

109 replace inc2_1=1
(5,318 real changes made)

110 replace inc2_2=1 if inlist(income, 2,3,4,5,6,7,8,9, 10)
(4,025 real changes made)

111 replace inc2_3=1 if inlist(income, 3,4,5,6,7,8, 9, 10)
(3,901 real changes made)

```

112 replace inc2_4=1 if inlist(income, 4,5,6,7, 8, 9, 10)
    (2,916 real changes made)

113 replace inc2_5=1 if inlist(income, 5,6,7,8,9,10)
    (2,790 real changes made)

114 replace inc2_6=1 if inlist(income, 6,7,8,9,10)
    (665 real changes made)

115 replace inc2_7=1 if inlist(income, 7,8,9,10)
    (540 real changes made)

116 replace inc2_8=1 if inlist(income, 8,9,10)
    (187 real changes made)

117 replace inc2_9=1 if inlist(income, 9,10)
    (56 real changes made)

118 replace inc2_10=1 if income==10
    (36 real changes made)

119 tab inc2_8 income

```

		Household Income Decile - Subjective						
		8	1	2	3	4	5	6
> 7	inc2_8	Total						
	0	1,293	124	985	126	2,125	125	3
> 53	1	0	5,131	0	0	0	0	0
> 0		131	187					
	Total	1,293	124	985	126	2,125	125	3
> 53		131	5,318					

		Household Income Decile - Subjective		Total
inc2_8		9	10	
0		0	0	5,131
1		20	36	187
Total		20	36	5,318

```

120 *create a distinct indicator of income.
121 gen byte incomecat = income

122 de incomecat, detail

```

Variable name	Storage type	Display format	Value label	Variable label
incomecat	byte	%8.0g		

```

123 *it is simply income.
124
125 *Create a cumulative age schema.

```

```

126 gen age2_1=0
127 gen age2_2=0
128 gen age2_3=0
129 gen age2_4=0
130 gen age2_5=0
131 gen age2_6=0
132 gen age2_7=0
133 gen age2_8=0
134 gen age2_9=0
135 gen age2_10=0
136 replace age2_1 = 1 if age91to98==1 | age83to90==1 | age75to82==1 | age67to74==1 | ag
> e59to66==1 | age51to58==1 | age43to50==1 | age35to42==1 | age27to34==1 | age18to26==
> 1
(5,318 real changes made)
137 replace age2_2 = 1 if age91to98==1 | age83to90==1 | age75to82==1 | age67to74==1 | ag
> e59to66==1 | age51to58==1 | age43to50==1 | age35to42==1 | age27to34==1
(4,388 real changes made)
138 replace age2_3 = 1 if age91to98==1 | age83to90==1 | age75to82==1 | age67to74==1 | ag
> e59to66==1 | age51to58==1 | age43to50==1 | age35to42==1
(3,364 real changes made)
139 replace age2_4 = 1 if age91to98==1 | age83to90==1 | age75to82==1 | age67to74==1 | ag
> e59to66==1 | age51to58==1 | age43to50==1
(2,273 real changes made)
140 replace age2_5 = 1 if age91to98==1 | age83to90==1 | age75to82==1 | age67to74==1 | ag
> e59to66==1 | age51to58==1
(1,340 real changes made)
141
142 replace age2_6 = 1 if age91to98==1 | age83to90==1 | age75to82==1 | age67to74==1 | ag
> e59to66==1
(848 real changes made)
143 replace age2_7 = 1 if age91to98==1 | age83to90==1 | age75to82==1 | age67to74==1
(391 real changes made)
144 replace age2_8 = 1 if age91to98==1 | age83to90==1 | age75to82==1
(155 real changes made)
145 replace age2_9=1 if age91to98==1 | age83to90==1
(44 real changes made)
146 replace age2_10 = 1 if age91to98==1
(14 real changes made)
147 *create a distinct age schema. Just use i.agecat.
148 gen age1_1=0

```

```
149 gen age1_2=0
150 gen age1_3=0
151 gen age1_4=0
152 gen age1_5=0
153 gen age1_6=0
154 gen age1_7=0
155 gen age1_8=0
156 gen age1_9=0
157 gen age1_10=0

158 *work up the replacement carefully.
159 replace age1_1 =1 if age18to26==1
    (930 real changes made)

160 replace age1_2 =1 if age27to34==1
    (1,024 real changes made)

161 replace age1_3 =1 if age35to42==1
    (1,091 real changes made)

162 replace age1_4 =1 if age43to50==1
    (933 real changes made)

163 replace age1_5 =1 if age51to58==1
    (492 real changes made)

164 replace age1_6 =1 if age59to66==1
    (457 real changes made)

165 replace age1_7 =1 if age67to74==1
    (236 real changes made)

166 replace age1_8 =1 if age75to82==1
    (111 real changes made)

167 replace age1_9 =1 if age83to90==1
    (30 real changes made)

168 replace age1_10 =1 if age91to98==1
    (14 real changes made)

169 gen agecat=0

170 replace agecat=1 if age1_1==1
    (930 real changes made)

171 replace agecat=2 if age1_2==1
    (1,024 real changes made)

172 replace agecat=3 if age1_3==1
    (1,091 real changes made)
```

```

173 replace agecat=4 if age1_4==1
    (933 real changes made)

174 replace agecat=5 if age1_5==1
    (492 real changes made)

175 replace agecat=6 if age1_6==1
    (457 real changes made)

176 replace agecat=7 if age1_7==1
    (236 real changes made)

177 replace agecat=8 if age1_8==1
    (111 real changes made)

178 replace agecat=9 if age1_9==1
    (30 real changes made)

179 replace agecat=10 if age1_10==1
    (14 real changes made)

180 tab agecat SE6, col

```

Key
<i>frequency</i>
<i>column percentage</i>

agecat		se6 Religion							
> an	None	Islam	Hindu	Buddhist	Jain	Sikhism	Others	Christi	
		Total							
1		118	737	7	0	22	21		
> 20	5	930							
> 39	50.00	18.02	17.00	17.95	0.00	18.97	51.22	17.	
		17.49							
2		122	850	6	1	22	4		
> 19	0	1,024							
> 52	0.00	18.63	19.60	15.38	16.67	18.97	9.76	16.	
		19.26							
3		149	881	14	2	19	5		
> 20	1	1,091							
> 39	10.00	22.75	20.32	35.90	33.33	16.38	12.20	17.	
		20.52							
4		115	760	5	1	23	5		
> 23	1	933							
> 00	10.00	17.56	17.53	12.82	16.67	19.83	12.20	20.	
		17.54							
5		55	413	2	2	8	1		
> 10	1	492							
> 70	10.00	8.40	9.52	5.13	33.33	6.90	2.44	8.	
		9.25							
6		52	378	1	0	11	3		
> 10	2	457							
> 70	20.00	7.94	8.72	2.56	0.00	9.48	7.32	8.	
		8.59							
7		28	194	1	0	7	1		

> 5	0	236							
> 35	0.00	4.27	4.44	4.47	2.56	0.00	6.03	2.44	4.
<hr/>									
> 5	8	13	111	89	2	0	2	0	
> 35	0.00	1.98	2.09	2.05	5.13	0.00	1.72	0.00	4.
<hr/>									
> 3	9	3	30	22	1	0	1	0	
> 61	0.00	0.46	0.56	0.51	2.56	0.00	0.86	0.00	2.
<hr/>									
> 0	10	0	14	12	0	0	1	1	
> 00	0.00	0.00	0.26	0.28	0.00	0.00	0.86	2.44	0.
<hr/>									
> 15	Total	655	5,318	4,336	39	6	116	41	1
> 00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.

181 *Strong bias to low age groups in the unweighted Asian Barometers.

182

183 *The AB 2019 dataset has case weights. We have not used them here.

184 isid IDnumber

185

186

187 *Descriptives

188 tab Q63

63 When a mother-in-law and a daughter-in-law come into conflict, even if the mo	Freq.	Percent	Cum.
Strongly agree	1,372	25.80	25.80
Somewhat agree	1,464	27.53	53.33
Somewhat disagree	1,031	19.39	72.72
Strongly disagree	959	18.03	90.75
Do not understand the question	77	1.45	92.20
Can't choose	220	4.14	96.33
Decline to answer	195	3.67	100.00
Total	5,318	100.00	

189 tab Q69

69 If one could have only one child, it is more preferable to have a boy than a	Freq.	Percent	Cum.
Strongly agree	999	18.79	18.79
Somewhat agree	1,219	22.92	41.71
Somewhat disagree	1,005	18.90	60.61
Strongly disagree	1,600	30.09	90.69
Do not understand the question	76	1.43	92.12
Can't choose	192	3.61	95.73
Decline to answer	227	4.27	100.00
Total	5,318	100.00	

190 tab Q146

146 Women should not be involved in politics as much as men	Freq.	Percent	Cum.
Strongly agree	1,262	23.73	23.73
Somewhat agree	1,234	23.20	46.93
Somewhat disagree	924	17.37	64.31
Strongly disagree	1,366	25.69	90.00
Do not understand the question	54	1.02	91.01
Can't choose	163	3.07	94.08
Decline to answer	315	5.92	100.00
Total	5,318	100.00	

191 gen female=0

192 replace female=1 if SE2==2
(2,357 real changes made)

193 hist w
(bin=37, start=.64050809, width=.09194254)

194 tab Region

Region	Freq.	Percent	Cum.
Andhra Pradesh	240	4.51	4.51
Assam	153	2.88	7.39
Bihar	401	7.54	14.93
Gujarat	231	4.34	19.27
Haryana	162	3.05	22.32
Karnataka	327	6.15	28.47
Kerala	160	3.01	31.48
Madhya Pradesh	315	5.92	37.40
Maharashtra	439	8.25	45.66
Odisha	187	3.52	49.17
Punjab	160	3.01	52.18
Rajasthan	224	4.21	56.39
Tamil Nadu	404	7.60	63.99
Uttar Pradesh	909	17.09	81.08
West Bengal	367	6.90	87.98
Delhi	156	2.93	90.92
Jharkhand	160	3.01	93.93
Chhattisgarh	163	3.07	96.99
Telangana	160	3.01	100.00
Total	5,318	100.00	

195 tab Level

Rural or Urban	Freq.	Percent	Cum.
Rural	4,219	79.33	79.33
Urban	1,099	20.67	100.00
Total	5,318	100.00	

196 tab Q63

63 When a mother-in-law and a daughter-in-law come into conflict, even if the mo	Freq.	Percent	Cum.
Strongly agree	1,372	25.80	25.80
Somewhat agree	1,464	27.53	53.33
Somewhat disagree	1,031	19.39	72.72
Strongly disagree	959	18.03	90.75
Do not understand the question	77	1.45	92.20
Can't choose	220	4.14	96.33
Decline to answer	195	3.67	100.00
Total	5,318	100.00	

197 tab Q63 , nol

63 When a mother-in-law and a daughter-in-law come into conflict, even if the mo	Freq.	Percent	Cum.
1	1,372	25.80	25.80
2	1,464	27.53	53.33
3	1,031	19.39	72.72
4	959	18.03	90.75
7	77	1.45	92.20
8	220	4.14	96.33
9	195	3.67	100.00
Total	5,318	100.00	

198 *Variable based on "Q63. When a mother-in-law and a daughter-in-law come into conflict, even if the mother-in-law is in the wrong, the husband should still persuade his wife to obey his mother."
 199 gen rural=0

200 replace rural=1 if Level==1
 (4,219 real changes made)

201 gen edu=0

202 replace edu=1 if SE5==1|SE5==2
 (1,245 real changes made)

203 replace edu=2 if SE5==3
 (547 real changes made)

204 replace edu=3 if SE5==4|SE5==6
 (624 real changes made)

205 replace edu=4 if SE5==5|SE5==7
 (1,381 real changes made)

```
206 replace edu=5 if SE5==8|SE5==9|SE5==10
    (1,157 real changes made)
```

```
207 replace edu=1 if SE5==99
    (364 real changes made)
```

```
208 tab edu
```

edu	Freq.	Percent	Cum.
1	1,609	30.26	30.26
2	547	10.29	40.54
3	624	11.73	52.28
4	1,381	25.97	78.24
5	1,157	21.76	100.00
Total	5,318	100.00	

```
209 label define educ 1 "Below Primary" 2 "Primary" 3 "Incomplete Secondary" 4 "Complete
> Secondary" 5 "Higher Educ", modify
```

```
210 label values edu educ
```

```
211 *Note. This is the original recoding of Education but we could alternatively receive
> the one-hot encoded variables as vectors from R using merge. The one that was in R
> is called Edu (sic)
```

```
212 *But. When we use R we kept the edu1 and edu2 encoded blocks in dataframes. So one
> has to start combining etc. before a single write.csv and it's not really worth it.
```

```
213
```

```
214
```

```
215 *Scheme 1 distinct coding - it arrived from R.
```

```
216 *gen edu1_1=0
```

```
217 *gen edu1_2=0
```

```
218 *gen edu1_3=0
```

```
219 *gen edu1_4=0
```

```
220 *gen edu1_5=0
```

```
221 *replace edu1_1=1 if edu==1
```

```
222 *replace edu1_2=1 if edu==2
```

```
223 *replace edu1_3=1 if edu==3
```

```
224 *replace edu1_4=1 if edu==4
```

```
225 *replace edu1_5=1 if edu==5
```

```
226 *Scheme 2 cumulative coding
```

```
227 *gen edu2_1=0
```

```
228 *gen edu2_2=0
```

```
229 *gen edu2_3=0
```

```
230 *gen edu2_4=0
```

```
231 *gen edu2_5=0
```

```
232 *replace edu2_1=1 if inlist(edu, 1)
```

```
233 *replace edu2_2=1 if inlist(edu, 1,2)
```

```
234 *replace edu2_3=1 if inlist(edu, 1,2,3)
```

```
235 *replace edu2_4=1 if inlist(edu, 1,2,3,4)
```

```
236 *replace edu2_5=1 if inlist(edu, 1,2,3,4,5)
```

```
237 tab edu2_4 edu
```

edu2_4	Below Pri	Primary	edu Incomplet	Complete	Higher Ed	Total
0	1,609	547	624	0	0	2,780
1	0	0	0	1,381	1,157	2,538
Total	1,609	547	624	1,381	1,157	5,318

238 tab edu1_2 edu

edu1_2	Below Pri	Primary	edu Incomplet	Complete	Higher Ed	Total
0	1,609	0	624	1,381	1,157	4,771
1	0	547	0	0	0	547
Total	1,609	547	624	1,381	1,157	5,318

239 tab edu2_2 edu

edu2_2	Below Pri	Primary	edu Incomplet	Complete	Higher Ed	Total
0	1,609	0	0	0	0	1,609
1	0	547	624	1,381	1,157	3,709
Total	1,609	547	624	1,381	1,157	5,318

240 tab edu2_5 edu

edu2_5	Below Pri	Primary	edu Incomplet	Complete	Higher Ed	Total
0	1,609	547	624	1,381	0	4,161
1	0	0	0	0	1,157	1,157
Total	1,609	547	624	1,381	1,157	5,318

241 tab edu1_3 edu

edu1_3	Below Pri	Primary	edu Incomplet	Complete	Higher Ed	Total
0	1,609	547	0	1,381	1,157	4,694
1	0	0	624	0	0	624
Total	1,609	547	624	1,381	1,157	5,318

242 tab edu2_3 edu

edu2_3	Below Pri	Primary	edu Incomplet	Complete	Higher Ed	Total
0	1,609	547	0	0	0	2,156
1	0	0	624	1,381	1,157	3,162
Total	1,609	547	624	1,381	1,157	5,318

243 *

244 gen state=Region

245 gen muslim=0

246 replace muslim=1 if SE6==40
(655 real changes made)

247 label define muslim 0 "No" 1 "Muslim", modify

248 label values muslim muslim

249 summ Q63, detail

63 When a mother-in-law and a daughter-in-law come
into conflict, even if the mo

Percentiles		Smallest		
1%	1	1		
5%	1	1		
10%	1	1	Obs	5,318
25%	1	1	Sum of wgt.	5,318
50%	2		Mean	2.873825
		Largest	Std. dev.	2.014539
75%	4	9		
90%	4	9	Variance	4.058367
95%	8	9	Skewness	1.6391
99%	9	9	Kurtosis	5.347439

250 tab Q63

63 When a mother-in-law and a daughter-in-law come into conflict, even if the mo		Freq.	Percent	Cum.
Strongly agree		1,372	25.80	25.80
Somewhat agree		1,464	27.53	53.33
Somewhat disagree		1,031	19.39	72.72
Strongly disagree		959	18.03	90.75
Do not understand the question		77	1.45	92.20
Can't choose		220	4.14	96.33
Decline to answer		195	3.67	100.00
Total		5,318	100.00	

251 *Preferring that a daughter-in-law NOT concede to husband's mother is a high value o
> n this ordinal scale.

252 gen op1=Q63

253 replace op1=4 if Q63==3
(1,031 real changes made)

254 replace op1=5 if Q63==4
(959 real changes made)

255 *take care of Do not understand the question 7 Can't choose 8 and Decline to answer
> 9.

256 replace op1=3 if Q63==7|Q63==8|Q63==9
(492 real changes made)

257 tab op1

op1	Freq.	Percent	Cum.
1	1,372	25.80	25.80
2	1,464	27.53	53.33
3	492	9.25	62.58
4	1,031	19.39	81.97
5	959	18.03	100.00
Total	5,318	100.00	

258 gen opinDIL=op1

259

260 * N0tes: Do not understand the question | 7

261 *> 7 1.45 92.20

262 * Can't choose | 22

263 *> 0 4.14 96.33

264 * Decline to answer | 19

265 *> 5 3.67 100.00

266

267 tab Q69, nol

69 If one could have only one child, it is more preferable to have a boy than a			
	Freq.	Percent	Cum.
1	999	18.79	18.79
2	1,219	22.92	41.71
3	1,005	18.90	60.61
4	1,600	30.09	90.69
7	76	1.43	92.12
8	192	3.61	95.73
9	227	4.27	100.00
Total	5,318	100.00	

268 *Preferring to have girl children is a high value on the ordinal scale.

269 gen op2=Q69

270 replace op2=4 if Q69==3

(1,005 real changes made)

271 replace op2=5 if Q69==4

(1,600 real changes made)

272 *take care of Do not understand the question 7 Can't choose 8 and Decline to answer
> 9.

273 replace op2=3 if Q69==7|Q69==8|Q69==9

(495 real changes made)

274 tab op2

op2			
	Freq.	Percent	Cum.
1	999	18.79	18.79
2	1,219	22.92	41.71
3	495	9.31	51.02
4	1,005	18.90	69.91
5	1,600	30.09	100.00
Total	5,318	100.00	

275 gen opinBOY=op2

276

277 summ Q146, detail

146 Women should not be involved in politics as
much as men

Percentiles		Smallest		
1%	1	1		
5%	1	1		
10%	1	1	Obs	5,318
25%	2	1	Sum of wgt.	5,318
50%	3		Mean	3.099473
		Largest	Std. dev.	2.119009
75%	4	9		
90%	7	9	Variance	4.490197
95%	9	9	Skewness	1.47824
99%	9	9	Kurtosis	4.8102

278 tab Q146, nol

146 Women should not be involved in politics as much as men	Freq.	Percent	Cum.
1	1,262	23.73	23.73
2	1,234	23.20	46.93
3	924	17.37	64.31
4	1,366	25.69	90.00
7	54	1.02	91.01
8	163	3.07	94.08
9	315	5.92	100.00
Total	5,318	100.00	

279 * Preferring to have women engaged in politics is a high value on the ordinal scale.

280 gen op3=Q146

281 replace op3=4 if Q146==3
(924 real changes made)

282 replace op3=5 if Q146==4
(1,366 real changes made)

283 *take care of Do not understand the question 7 Can't choose 8 and Decline to answer
> 9.

284 replace op3=3 if Q146==7|Q146==8|Q146==9
(532 real changes made)

285 tab op3

op3	Freq.	Percent	Cum.
1	1,262	23.73	23.73
2	1,234	23.20	46.93
3	532	10.00	56.94
4	924	17.37	74.31
5	1,366	25.69	100.00
Total	5,318	100.00	

```
286 gen opinFEM=op3
```

```
287 table (female), stat(mean op1-op3) nformat(%9.1f)
```

	op1	opinDIL	op2	opinBOY	op3
female					
0	2.7	2.7	3.2	3.2	2.9
1	2.9	2.9	3.1	3.1	3.0
Total	2.8	2.8	3.2	3.2	3.0

```
288 format op1-op3 %9.1f
```

```
289 corr op1-op3, noformat
      (obs=5,318)
```

	op1	opinDIL	op2	opinBOY	op3
op1	1.0000				
opinDIL	1.0000	1.0000			
op2	0.2356	0.2356	1.0000		
opinBOY	0.2356	0.2356	1.0000	1.0000	
op3	0.2046	0.2046	0.1869	0.1869	1.0000

```
290 save "C:/data/AsianBaro/data/AsianBaro2019revForEntropy.dta", replace
file C:/data/AsianBaro/data/AsianBaro2019revForEntropy.dta saved
```

291 *Note: opinFEM is the one about women in politics, FEM being the view that women sho
> uld be engaged in politics. opinDIL is the one about daughters-i
> n-law: Variable based on "Q63. When a mother-in-law and a daughter-in- law come into
> conflict, even if the mother- in-law is in the wrong, the husband should still pers
> uade his wife to obey his mother. And, opinBOY is the view about preferr
> ing to have a boy child, if only 1 child can be had.

292

293

294 *Figure 0 Income and Age by Education in Asia Barometers 2019 India

295 *Table for Figure 0

```
296 collect: table (edu) (agecat) (incomecat), statistic (mean agecat) statistic(perce
> nt edu) statistic(n income) nformat (%9.0f mean) nformat (%9.0f n)
```

```
incomecat = 1
```

					ageca					
> t					1	2	3	4	5	6
>	7	8	9	10	Total					

edu

 \succ

Below Primary

>

Mean

$$>$$

agecat

$$>$$

Percent

 γ

edu

```
> 0.31 0.21 0.04 0.03
   Number of nonmissing values
```

 \succ

Household Income Decile - Subjective

 \succ

Primary

 \succ

Mean

 \succ

>	agecat					1	2	3	4	5	6
>	7	8	9		3						
>	Percent										
>	edu					0.29	0.35	0.43	0.35	0.23	0.13
>	0.05	0.03	0.01		1.86						
>	Number of nonmissing values										
>	Household Income Decile - Subjective					23	28	34	28	18	10
>	4	2	1		148						
>	Incomplete Secondary										
>	Mean										
>	agecat					1	2	3	4	5	6
>	7				3						
>	Percent										
>	edu					0.55	0.89	0.68	0.49	0.13	0.13
>	0.06				2.93						
>	Number of nonmissing values										
>	Household Income Decile - Subjective					29	47	36	26	7	7
>	3				155						
>	Complete Secondary										
>	Mean										
>	agecat					1	2	3	4	5	6
>	7	8		10	3						
>	Percent										
>	edu					1.69	1.49	1.59	0.76	0.40	0.10
>	0.20	0.15		0.05	6.42						
>	Number of nonmissing values										
>	Household Income Decile - Subjective					67	59	63	30	16	4
>	8	6		2	255						
>	Higher Educ										
>	Mean										
>	agecat					1	2	3	4	5	6
>	8				2						
>	Percent										
>	edu					1.07	0.72	0.54	0.22	0.13	0.09
>	0.03				2.80						
>	Number of nonmissing values										
>	Household Income Decile - Subjective					34	23	17	7	4	3
>	1				89						
>	Total										
>	Mean										
>	agecat					1	2	3	4	5	6
>	7	8	9	10	4						
>	Percent										
>	edu					3.93	3.95	3.90	2.76	1.28	1.09
>	0.62	0.42	0.05	0.08	18.08						
>	Number of nonmissing values										
>	Household Income Decile - Subjective					206	236	257	241	108	125
>	64	43	7	6	1293						

incomecat = 2

				agecat					
				1	2	3	4	5	6
7	8	9	Total						
edu									
> Below Primary									
> Mean									
> agecat									
7	8	9	5	1	2	3	4	5	6
> Percent									
> edu									
0.04	0.03	0.01	0.43	0.02	0.03	0.06	0.11	0.06	0.09
> Number of nonmissing values									
> Household Income Decile - Subjective									
6	4	2	69	3	5	9	17	9	14
> Primary									
> Mean									
> agecat									
7			3	1	2	3	4	5	6
> Percent									
> edu									
0.01			0.24	0.04	0.05	0.05	0.04	0.03	0.03
> Number of nonmissing values									
> Household Income Decile - Subjective									
1			19	3	4	4	3	2	2
> Incomplete Secondary									
> Mean									
> agecat									
			3	1	2	3	4		
> Percent									
> edu									
			0.19	0.04	0.02	0.08	0.06		
> Number of nonmissing values									
> Household Income Decile - Subjective									
			10	2	1	4	3		
> Complete Secondary									
> Mean									
> agecat									
			3	1	2	3	4	5	6
> Percent									
> edu									
			0.53	0.08	0.15	0.13	0.10	0.05	0.03
> Number of nonmissing values									
> Household Income Decile - Subjective									
			21	3	6	5	4	2	1
> Higher Educ									
> Mean									
> agecat									
				1	2				

>	Percent	2									
>	edu				0.06	0.09					
>	Number of nonmissing values	0.16									
>	Household Income Decile - Subjective				2	3					
>	Total	5									
>	Mean										
>	agecat				1	2	3	4	5	6	
>	Percent	7	8	9	4						
>	edu				0.23	0.35	0.31	0.30	0.13	0.14	
>	Number of nonmissing values	0.05	0.03	0.01	1.55						
>	Household Income Decile - Subjective				13	19	22	27	13	17	
>		7	4	2	124						

incomecat = 3

> t												ageca
>	7	8	9	10	Total							
>												
>	edu											
>	Below Primary											
>	Mean											
>	agecat					1	2	3	4	5	6	
>	Percent	7	8	9	4							
>	edu					0.14	0.25	0.40	0.44	0.27	0.22	
>	Number of nonmissing values	0.20	0.04	0.02	1.98							
>	Household Income Decile - Subjective					22	39	64	70	43	35	
>	Primary	32	6	3	314							
>	Mean											
>	agecat					1	2	3	4	5	6	
>	Percent	7	8	9	4							
>	edu					0.14	0.29	0.47	0.21	0.23	0.15	
>	Number of nonmissing values	0.04	0.01	0.01	1.55							
>	Household Income Decile - Subjective					11	23	37	17	18	12	
>	Incomplete Secondary	3	1	1	123							
>	Mean											
>	agecat					1	2	3	4	5	6	
>	Percent	7	8		3							

>	0.02	0.01	0.31						
>	Number of nonmissing values								
>	Household Income Decile - Subjective								
>	3	1	50	3	6	5	17	6	9
>	Primary								
>	Mean								
>	agecat								
>		9	4	1	2	3	4	5	6
>	Percent								
>	edu								
>				0.04	0.06	0.01	0.03	0.05	0.04
>	Number of nonmissing values								
>	Household Income Decile - Subjective								
>	1	19	3	5	1	2	4	3	
>	Incomplete Secondary								
>	Mean								
>	agecat								
>	7		4			3	4	5	6
>	Percent								
>	edu								
>						0.06	0.02	0.04	0.02
>	0.02		0.15						
>	Number of nonmissing values								
>	Household Income Decile - Subjective								
>	1	8				3	1	2	1
>	Complete Secondary								
>	Mean								
>	agecat								
>	8		3	1	2	3	4	5	
>	Percent								
>	edu								
>				0.20	0.15	0.18	0.05	0.10	
>	0.03		0.71						
>	Number of nonmissing values								
>	Household Income Decile - Subjective								
>	1	28	8	6	7	2	4		
>	Higher Educ								
>	Mean								
>	agecat								
>			2	1	2	3	4		
>	Percent								
>	edu								
>				0.16	0.25	0.13	0.13		
>	Number of nonmissing values								
>	Household Income Decile - Subjective								
>		21	5	8	4	4			
>	Total								
>	Mean								
>	agecat								
>	7	8	9	4	1	2	3	4	5
>	Percent								
>	edu								
>				0.42	0.50	0.40	0.33	0.23	0.11
>	0.04	0.03	0.02	2.07					
>	Number of nonmissing values								

```

>
> Household Income Decile - Subjective | 19 25 20 26 16 13
> 4 1 2 126

```

```

incomecat = 5

```

```

> t
>
> 7 8 9 10 Total | 1 2 3 4 5 6
>
>
> edu
> Below Primary
> Mean
> agecat
> 7 8 9 10 5 | 1 2 3 4 5 6
> Percent
>
> edu
> 0.21 0.11 0.06 0.03 2.30 | 0.16 0.16 0.31 0.47 0.38 0.43
> Number of nonmissing values
>
> Household Income Decile - Subjective
> 33 17 9 4 366 | 25 25 50 75 60 68
> Primary
> Mean
> agecat
> 7 8 10 4 | 1 2 3 4 5 6
> Percent
>
> edu
> 0.08 0.03 0.01 2.35 | 0.24 0.24 0.69 0.44 0.28 0.35
> Number of nonmissing values
>
> Household Income Decile - Subjective
> 6 2 1 187 | 19 19 55 35 22 28
> Incomplete Secondary
> Mean
> agecat
> 7 8 4 | 1 2 3 4 5 6
> Percent
>
> edu
> 0.25 0.08 4.74 | 0.53 1.02 1.02 0.91 0.62 0.32
> Number of nonmissing values
>
> Household Income Decile - Subjective
> 13 4 251 | 28 54 54 48 33 17
> Complete Secondary
> Mean
> agecat
> 7 8 9 3 | 1 2 3 4 5 6
> Percent
>
> edu
> 0.68 0.25 0.08 17.20 | 2.39 3.95 4.10 3.27 1.49 0.98
> Number of nonmissing values
>
> Household Income Decile - Subjective | 95 157 163 130 59 39

```

					1	2	3	4	5	6
>	27	10	3	683						
>	Higher Educ									
>	Mean									
>	agecat									
>	7	8	10	2						
>	Percent									
>	edu				6.58	5.95	3.68	1.95	1.04	0.47
>	0.31	0.06	0.03	20.08						
>	Number of nonmissing values									
>	Household Income Decile - Subjective				209	189	117	62	33	15
>	10	2	1	638						
>	Total									
>	Mean									
>	agecat									
>	7	8	9	10	3					
>	Percent									
>	edu				9.90	11.32	9.81	7.04	3.80	2.56
>	1.52	0.52	0.13	0.07	46.68					
>	Number of nonmissing values									
>	Household Income Decile - Subjective				376	444	439	350	207	167
>	89	35	12	6	2125					

incomecat = 6

				agecat						
				1	2	3	4	5	6	
>	7	8	Total							
<hr/>										
edu										
>	Below Primary									
>	Mean									
>	agecat									
>	7	8	5		1	2	3	4	5	6
>	Percent									
>	edu				0.01	0.03	0.02	0.01	0.04	0.04
>	0.05	0.02	0.21							
>	Number of nonmissing values									
>	Household Income Decile - Subjective				2	5	3	1	6	6
>	8	3	34							
>	Primary									
>	Mean									
>	agecat									
>		4			1	2	3	4	5	6
>	Percent									
>	edu				0.01	0.03	0.01	0.03	0.01	0.03
>		0.11								
>	Number of nonmissing values									
>	Household Income Decile - Subjective				1	2	1	2	1	2
>		9								
>	Incomplete Secondary									

>						
>	agecat					
>	7	8	9	4		
>	Percent					
>	edu					
>	0.01	0.03	0.01	0.37		
>	Number of nonmissing values					
>	Household Income Decile - Subjective					
>	1	5	1	58		
>	Primary					
>	Mean					
>	agecat					
>	7			4		
>	Percent					
>	edu					
>	0.01			0.35		
>	Number of nonmissing values					
>	Household Income Decile - Subjective					
>	1			28		
>	Incomplete Secondary					
>	Mean					
>	agecat					
>	7	8		4		
>	Percent					
>	edu					
>	0.08	0.02		0.59		
>	Number of nonmissing values					
>	Household Income Decile - Subjective					
>	4	1		31		
>	Complete Secondary					
>	Mean					
>	agecat					
>	7	8		4		
>	Percent					
>	edu					
>	0.13	0.05		1.99		
>	Number of nonmissing values					
>	Household Income Decile - Subjective					
>	5	2		79		
>	Higher Educ					
>	Mean					
>	agecat					
>	7	8		3		
>	Percent					
>	edu					
>	0.03			4.94		
>	Number of nonmissing values					
>	Household Income Decile - Subjective					
>	1			157		
>	Total					
>	Mean					
>	agecat					
>	7					


```
incomecat = 10
```

			agecat						
			1	2	3	4	5	6	
7	8	Total							
edu									
> Below Primary									
> Mean									
> agecat			1	2	3	4	5	6	
> 7 4									
> Percent									
> edu			0.02	0.01	0.01	0.03	0.03	0.03	
> 0.01 0.13									
> Number of nonmissing values									
> Household Income Decile - Subjective			3	1	2	4	4	5	
> 1 20									
> Primary									
> Mean									
> agecat			1						
> 7 4									
> Percent									
> edu			0.01						
> 0.01 0.03									
> Number of nonmissing values									
> Household Income Decile - Subjective			1						
> 1 2									
> Incomplete Secondary									
> Mean									
> agecat						3	4		
> 4									
> Percent									
> edu						0.02	0.02		
> 0.04									
> Number of nonmissing values									
> Household Income Decile - Subjective						1	1		
> 2									
> Complete Secondary									
> Mean									
> agecat			1	2					5
> 2									
> Percent									
> edu			0.03	0.08					0.03
> 0.13									
> Number of nonmissing values									
> Household Income Decile - Subjective			1	3					1
> 5									
> Higher Educ									
> Mean									
> agecat			1	2				4 5	

>	Percent	8	4						
>	edu			0.06	0.03		0.03	0.06	
>	Number of nonmissing values	0.03	0.22						
>	Household Income Decile - Subjective			2	1		1	2	
>	Total	1	7						
>	Mean								
>	agecat			1	2	3	4	5	6
>	Percent	7	8	4					
>	edu			0.12	0.11	0.03	0.08	0.11	0.03
>	Number of nonmissing values	0.02	0.03	0.54					
>	Household Income Decile - Subjective			7	5	3	6	7	5
>	Total	2	1	36					

incomecat = Total

							age				
cat							1	2	3	4	5
>	6	7	8	9	10	Total					
edu											
>	Below Primary										
>	Mean										
>	agecat						1	2	3	4	5
>	Percent	6	7	8	9	10					
>	edu						0.78	1.08	1.61	2.22	1.32
>	Number of nonmissing values	.61	0.86	0.45	0.15	0.05					1
>	Household Income Decile - Subjective						124	171	255	353	210
>	Primary	256	137	71	24	8	1609				
>	Mean										
>	agecat						1	2	3	4	5
>	Percent	6	7	8	9	10					
>	edu						0.81	1.06	1.80	1.22	0.88
>	Number of nonmissing values	.79	0.21	0.06	0.04	0.01	6.89				0
>	Household Income Decile - Subjective						64	84	143	97	70
>	Incomplete Secondary	63	17	5	3	1	547				
>	Mean										
>	agecat						1	2	3	4	5
>	Percent	6	7	8							

```

>
>      edu
> .79  0.47  0.13      11.79 | 1.62  2.72  2.74  2.27  1.04  0
>      Number of nonmissing values
>
>      Household Income Decile - Subjective
> 42   25   7      624 | 86   144   145   120   55
>      Complete Secondary
>
>      Mean
>
>      agecat
> 6   7   8   9   10   3 | 1   2   3   4   5
>      Percent
>
>      edu
> .61  1.11  0.58  0.08  0.08  34.78 | 6.75  7.91  8.21  5.92  2.54  1
>      Number of nonmissing values
>
>      Household Income Decile - Subjective
> 64   44   23   3   3   1381 | 268  314  326  235  101
>      Higher Educ
>
>      Mean
>
>      agecat
> 6   7   8   10   2 | 1   2   3   4   5
>      Percent
>
>      edu
> .01  0.41  0.16      0.06  36.42 | 12.21  9.79  6.99  4.03  1.76  1
>      Number of nonmissing values
>
>      Household Income Decile - Subjective
> 32   13   5      2   1157 | 388  311  222  128  56
>      Total
>
>      Mean
>
>      agecat
> 6   7   8   9   10   3 | 1   2   3   4   5
>      Percent
>
>      edu
> .82  3.07  1.38  0.26  0.20  100.00 | 22.17  22.55  21.34  15.66  7.55  5
>      Number of nonmissing values
>
>      Household Income Decile - Subjective
> 457   236   111   30   14   5318 | 930  1024  1091  933  492

```

```

297 collect export "EducationIncomeAgeIndia2019.xlsx", cell(C6) sheet(Figure0table) modi
> fy
(collection Table exported to file EducationIncomeAgeIndia2019.xlsx)
298 putexcel set EducationIncomeAgeIndia2019.xlsx, sheet(Figure0table) modify
299 putexcel A1 ="Household Income, Age and Education of Respondent, Info for Figure 0"
file EducationIncomeAgeIndia2019.xlsx saved

```

```

300 putexcel A2 ="EducationIncomeAgeIndia2019"
    file EducationIncomeAgeIndia2019.xlsx saved

301 putexcel (a3) = "University of Manchester"
    file EducationIncomeAgeIndia2019.xlsx saved

302 putexcel C6:L6 , txtwrap
    file EducationIncomeAgeIndia2019.xlsx saved

303 putexcel close

304 putexcel save

305 * Bar charts of original education levels reported, with age mean and Income Decile.
>
306 gen educLabel=edu

307 lab var educLabel "Education Level of Adult"

308 lab define educLabel 1 "" 2 "Pr" 3 "" 4 "Se" 5 "Uni", modify

309 lab val educLabel educLabel

310
311 lab var income "Household Income Decile"

312 lab define inclab 1 "Worst-Off" 10 "Best-Off", modify

313 lab val income inclab

314 graph hbar age , over(educLabel) over(income) bar(1, color(sand)) bar(2,color(gold
> )) blabel(bar, format(%9.1f) size(2)) ytitle("Education Levels Clustered; Age Sh
> ows Numeric Mean") title("Mean Age Within Education Group, By Income Decile") subtit
> le("High Income Is at Bottom")

315 graph export "results\EducandAgeByIncomeIndia.wmf", replace
    file C:\data\AsianBaro\results\EducandAgeByIncomeIndia.wmf saved as Windows Metafile
    format

316 graph export "results\EducandAgeByIncomeIndia.jpg", replace
    file results\EducandAgeByIncomeIndia.jpg written in JPEG format

317
318 *Figure 1 Income by Education in Asia Barometers 2019 India
319 *Table for Figure 1
320 collect: table (edu) (incomecat), statistic(percent edu) statistic(n income) nfor
> mat (%9.0f mean) nformat (%9.0f n)

```

				incomecat						
				1	2	3	4	5	6	7
8	9	10	Total							
edu										
> Below Primary										
> Percent										
> edu				4.07	0.43	1.98	0.31	2.30	0.21	0.37
> 0.31	0.02	0.13	10.13							
Number of nonmissing values										
> Household Income Decile				646	69	314	50	366	34	58
> 49	3	20	1609							
> Primary										
> Percent										
>										

edu	1.86	0.24	1.55	0.24	2.35	0.11	0.35
> 0.13 0.03 0.03 6.89							
Number of nonmissing values							
>							
Household Income Decile	148	19	123	19	187	9	28
> 10 2 2 547							
Incomplete Secondary							
>							
Percent							
>							
edu	2.93	0.19	2.47	0.15	4.74	0.36	0.59
> 0.28 0.04 0.04 11.79							
Number of nonmissing values							
>							
Household Income Decile	155	10	131	8	251	19	31
> 15 2 2 624							
Complete Secondary							
>							
Percent							
>							
edu	6.42	0.53	6.32	0.71	17.20	0.78	1.99
> 0.55 0.15 0.13 34.78							
Number of nonmissing values							
>							
Household Income Decile	255	21	251	28	683	31	79
> 22 6 5 1381							
Higher Educ							
>							
Percent							
>							
edu	2.80	0.16	5.23	0.66	20.08	1.01	4.94
> 1.10 0.22 0.22 36.42							
Number of nonmissing values							
>							
Household Income Decile	89	5	166	21	638	32	157
> 35 7 7 1157							
Total							
>							
Percent							
>							
edu	18.08	1.55	17.55	2.07	46.68	2.47	8.23
> 2.37 0.45 0.54 100.00							
Number of nonmissing values							
>							
Household Income Decile	1293	124	985	126	2125	125	353
> 131 20 36 5318							

321 collect export "EducationIncomeIndia2019.xlsx", cell(C6) sheet(Figure1table) modify
(collection **Table** exported to file [EducationIncomeIndia2019.xlsx](#))

322 putexcel set EducationIncomeIndia2019.xlsx, sheet(Figure1table) modify

323 putexcel A1 ="Income and Education of Respondent, Info for Figure 2"
file **EducationIncomeIndia2019.xlsx** saved

324 putexcel A2 ="EducationIncomeIndia2019"
file **EducationIncomeIndia2019.xlsx** saved


```
325 putexcel (a3) = "University of Manchester"
file EducationIncomeIndia2019.xlsx saved
```

```
326 putexcel C6:L6 , txtwrap
file EducationIncomeIndia2019.xlsx saved
```

```
327 putexcel close
```

```
328 putexcel save
```

```
329 * Bar charts of original education levels reported, with age mean and Income Decile.
>
```

```
330
```

```
331 lab var educLabel "Education Level of Adult"
```

```
332 lab define educLabel 1 "Less" 2 "Prim" 3 "Some" 4 "Secy" 5 "Uni", modify
```

```
333 lab val educLabel educLabel
```

```
334 tab educLabel income
```

Education Level of Adult		Household Income Decile							
> 7		Worst-Off	Total	2	3	4	5	6	
		8							
	Less		646	69	314	50	366	34	
> 58		49	1,609						
	Prim		148	19	123	19	187	9	
> 28		10	547						
	Some		155	10	131	8	251	19	
> 31		15	624						
	Secy		255	21	251	28	683	31	
> 79		22	1,381						
	Uni		89	5	166	21	638	32	1
> 57		35	1,157						
Total									
> 53		131	1,293	124	985	126	2,125	125	3
			5,318						

Education Level of Adult		Household Income Decile		Total
		9	Best-Off	
	Less	3	20	1,609
	Prim	2	2	547
	Some	2	2	624
	Secy	6	5	1,381
	Uni	7	7	1,157
Total		20	36	5,318

```
335 lab var income "Household Income Decile"
```

```
336 lab define inclab 1 "Worst-Off" 10 "Best-Off", modify
```

```
337 lab val income inclab
```

```

338 graph hbar edul_1 edul_2 edul_3 edul_4 edul_5, over(income) stack percentages bar(
> 1, color(sand)) bar(2,color(gold)) blabel(bar, format(%9.2f) size(2)) ytitle("Edu
> cation Levels") title("India, % in Education Group, By Income Decile")

339 graph export "results\EducByIncomeIndia.wmf", replace
file C:\data\AsianBaro\results\EducByIncomeIndia.wmf saved as Windows Metafile
format

340 graph export "results\EducByIncomeIndia.jpg", replace
file results\EducByIncomeIndia.jpg written in JPEG format

341
342 *Figure 2 Income Breakdown in Asia Barometers 2019 India
343 *Table for Figure 2
344 collect: table (incomecat), statistic(percent incomecat) statistic(n income) nfor
> mat (%9.0f mean) nformat (%9.0f n)

```

	Percent incomecat	Number of nonmissing values Household Income Decile
incomecat		
1	6.33	1293
2	1.21	124
3	14.46	985
4	2.47	126
5	52.00	2125
6	3.67	125
7	12.09	353
8	5.13	131
9	0.88	20
10	1.76	36
Total	100.00	5318

```

345 collect export "IncomeIndia2019.xlsx", cell(C6) sheet(Figure2table) modify
(collection Table exported to file IncomeIndia2019.xlsx)

346 putexcel set IncomeIndia2019.xlsx, sheet(Figure2table) modify

347 putexcel A1 ="Income of Respondent, Info for Figure 2"
file IncomeIndia2019.xlsx saved

348 putexcel A2 ="IncomeIndia2019"
file IncomeIndia2019.xlsx saved

349 putexcel (a3) = "University of Manchester"
file IncomeIndia2019.xlsx saved

350 putexcel C6:L6 , txtwrap
file IncomeIndia2019.xlsx saved

351 putexcel close

352 putexcel save

353 * Bar charts of original education levels reported, with age mean and Income Decile.
>
354 lab var income "Household Income Decile"

```

```
355 lab define inclab 1 "Worst-Off" 10 "Best-Off", modify
356 lab val income inclab
357 *graph hbar (percent) i.income, bar(1, color(sand)) bar(2,color(gold))   blabel(bar,
>   format(%9.2f) size(2))   title("India, % in Income Decile")
358
359 ** Part 3 Aggregate Exercise - see separate do file.
360 * Hypothesis. Using simulation, the MSE of H is higher for ordinal education than f
> or cumulative education when it is multinomial in 5 categories.
361 * We emulated education in five levels from the Asian Barometers, unweighted.
362 *this dataset has nothing in comment with the rest of the data.
363 *use "\data\Edtmp.dta", clear
364
365
366 save "C:/data/AsianBaro/data/AsianBaro2019revForEntropy.dta", replace
file C:/data/AsianBaro/data/AsianBaro2019revForEntropy.dta saved
367 log close
      name: <unnamed>
      log: C:\data\AsianBaro\logofcleaningforEntropy.smcl
      log type: smcl
      closed on: 10 Sep 2024, 15:52:33
```
