

## Read risk preference data of rd 2

August 19, 2021

17:17

Seiro Ito

### Risk preference:

1. A: 1000 BDT vs. B: Low 1000 BDT, High 1500 BDT
2. A: 1000 BDT vs. B: Low 875 BDT, High 1500 BDT
3. A: 1000 BDT vs. B: Low 750 BDT, High 1500 BDT
4. A: 1000 BDT vs. B: Low 625 BDT, High 1500 BDT
5. A: 1000 BDT vs. B: Low 500 BDT, High 1500 BDT

### Time preference 1:

1. A: 1000 BDT tomorrow vs. B: 1000 BDT in 3 months
2. A: 1000 BDT tomorrow vs. B: 1100 BDT in 3 months
3. A: 1000 BDT tomorrow vs. B: 1200 BDT in 3 months
4. A: 1000 BDT tomorrow vs. B: 1400 BDT in 3 months
5. A: 1000 BDT tomorrow vs. B: 1600 BDT in 3 months
6. A: 1000 BDT tomorrow vs. B: 2000 BDT in 3 months

### Time preference 2:

1. A: 1000 BDT in a year (12 months) vs. B: 1000 BDT in 1 year and 3 months (15 months)
2. A: 1000 BDT in a year (12 months) vs. B: 1200 BDT in 1 year and 3 months (15 months)
3. A: 1000 BDT in a year (12 months) vs. B: 1400 BDT in 1 year and 3 months (15 months)
4. A: 1000 BDT in a year (12 months) vs. B: 1600 BDT in 1 year and 3 months (15 months)
5. A: 1000 BDT in a year (12 months) vs. B: 2000 BDT in 1 year and 3 months (15 months)

```
library(readstata13)
rk <- read.dta13(paste0(path234, "only_2/Risk pref_Section 13.dta"))
, generate.factors = T, nonint.factors = T)
rk <- data.table(rk)
ToFacCols <- colnames(rk)[-c(1, ncol(rk))]
rk[, (ToFacCols) := lapply(.SD, factor), .SDcols = ToFacCols]
summary(rk)
```

	id	rp_1_bag_a	rp_2_bag_a	rp_3_bag_a	rp_4_bag_a	rp_5_bag_a		
Min.	:7.01e+06	: 2	: 2	: 2	: 2	: 2		
1st Qu.	:7.04e+06	a: 46	a: 133	a:1481	a:1964	a:1952		
Median	:7.13e+06	b:2035	b:1948	b: 600	b: 117	b: 129		
Mean	:1.65e+10							
3rd Qu.	:9.81e+09							
Max.	:9.91e+10							
NA's	:2							
	respondent_mid	tp_1_1_a	tp_1_2_a	tp_1_3_a	tp_1_4_a	tp_1_5_a	tp_1_6_a	tp_2_1_a
2	:1874	: 2	: 2	: 2	: 2	: 2	: 2	: 2
1	: 193	a:2070	a:2001	a:1451	a: 435	a: 36	a: 28	a:2061

3	:	8	b:	11	b:	80	b:	630	b:	1646	b:	2045	b:	2053	b:	20
4	:	2														
6	:	2														
(Other):	:	2														
NA's	:	2														
tp_2_2_a		tp_2_3_a		tp_2_4_a		tp_2_5_a		panel_no		year						
:	2	:	2	:	2	:	2	2:2083		Min.	:	2014				
a:1956		a: 930		a: 57		a: 24				1st Qu.:	:	2014				
b: 125		b:1151		b:2024		b:2057				Median	:	2014				
										Mean	:	2014				
										3rd Qu.:	:	2014				
										Max.	:	2014				

```
# add o800 variable
ar <- readRDS(paste0(pathsaveHere , DataFileNames[3], "Trimmed.rds"))
ar00 <- unique(ar[o800 == 1L, .(hhid, o800)])
setnames(rk, "id", "hhid")
setkey(rk, hhid); setkey(ar00, hhid)
rk <- ar00[rk]
rk <- rk[!is.na(hhid)]
rk[is.na(o800), o800 := 0L]
table(rk[, .(n = 1:N, o800), by = hhid][n==1, o800])
```

0	1
1336	742

```
# reshape to get the switching point
setnames(rk,
  grepout("rp", colnames(rk))
  , paste0("rp.", 1:5))
setnames(rk,
  grepout("tp_1", colnames(rk))
  , paste0("tp1.", 1:6))
setnames(rk,
  grepout("tp_2", colnames(rk))
  , paste0("tp2.", 1:5))
rk[, c("rp.6", "tp2.6") := NA]
# drop duplicates (keep 1st appearance)
rk[, n := 1:N, by = hhid]
rk <- rk[n==1, ]
rk[, n := NULL]
# find switching points
rkL <- reshape(rk, direction = "long", idvar = c("hhid", "respondent_mid", "o800"),
  varying = grepout("\\.\\d$", colnames(rk)))
setnames(rkL, "time", "q")
setkey(rkL, hhid, q)
# expect: b => a (2 => 1) at some point or never, not a => b (1 => 2)
rkL[, rp := factor(rp, levels = c("a", "b", NA))]
rkL[, tp1 := factor(tp1, levels = c("a", "b", NA))]
rkL[, tp2 := factor(tp2, levels = c("a", "b", NA))]
rkL[, rp := as.numeric(rp)]
rkL[, tp1 := as.numeric(tp1)]
rkL[, tp2 := as.numeric(tp2)]
rkL[, Drp := c(NA, diff(rp)), by = hhid]
rkL[, Dtp1 := c(NA, diff(tp1)), by = hhid]
rkL[, Dtp2 := c(NA, diff(tp2)), by = hhid]
```

```
# check if there is only 1 switch
rkL[, Num.rp := sum(!is.na(Drp) & Drp != 0), by = hhid]
rkL[, Num.tp1 := sum(!is.na(Dtp1) & Dtp1 != 0), by = hhid]
rkL[, Num.tp2 := sum(!is.na(Dtp2) & Dtp2 != 0), by = hhid]
table(rkL[, .(n = 1:N, Num.rp, o800), by = hhid][o800 == 1L & n == 1L, Num.rp])
```

0	1	2
16	712	14

```
table(rkL[, .(n = 1:N, Num.tp1, o800), by = hhid][o800 == 1L & n == 1L, Num.tp1])
```

0	1	2	3
1	735	5	1

```
table(rkL[, .(n = 1:N, Num.tp2, o800), by = hhid][o800 == 1L & n == 1L, Num.tp2])
```

0	1	2	3
7	733	1	1

```
# Change.rp for non-switcher or single switcher
# single switcher b => a (rp), a => b (tp1, tp2)
rkL[Num.rp == 1, Change.rp := (1:N)[Num.rp == 1 & !is.na(Drp) & Drp == -1], by = hhid]
rkL[Num.tp1 == 1, Change.tp1 := (1:N)[Num.tp1 == 1 & !is.na(Dtp1) & Dtp1 == 1], by = hhid]
rkL[Num.tp2 == 1, Change.tp2 := (1:N)[Num.tp2 == 1 & !is.na(Dtp2) & Dtp2 == 1], by = hhid]
# non switcher (always a or b)
rkL[hhid %in% hhid[Num.rp == 0], Change.rp := rp]
rkL[hhid %in% hhid[Num.tp1 == 0], Change.tp1 := tp1]
rkL[hhid %in% hhid[Num.tp2 == 0], Change.tp2 := tp2]
# multiple switcher (irrational)
rkL[Num.rp > 1, Change.rp := -1, by = hhid]
rkL[Num.tp1 > 1, Change.tp1 := -1, by = hhid]
rkL[Num.tp2 > 1, Change.tp2 := -1, by = hhid]
# single a => b (rp, irrational), b => a (tp1, tp2, irrational)
rkL[hhid %in% hhid[Num.rp == 1 & !is.na(Drp) & Drp == 1], Change.rp := -2]
rkL[hhid %in% hhid[Num.tp1 == 1 & !is.na(Dtp1) & Dtp1 == -1], Change.tp1 := -2]
rkL[hhid %in% hhid[Num.tp2 == 1 & !is.na(Dtp2) & Dtp2 == -1], Change.tp2 := -2]
rkL[, Change.rp := factor(as.integer(Change.rp))]
rkL[, Change.tp1 := factor(as.integer(Change.tp1))]
rkL[, Change.tp2 := factor(as.integer(Change.tp2))]
# Quantify the difference in switch point
# To do so, compare expected value differences in each option
# E.g., if a=1000, b=c(1000, 1500)*1/2, 1250-1000=250
rkL[, c("RiskPrefVal", "TimePref1Val", "TimePref2Val") := as.numeric(NA)]
rkL[Change.rp == 2, RiskPrefVal := (1500+875)/2-1000]
rkL[Change.rp == 3, RiskPrefVal := (1500+750)/2-1000]
rkL[Change.rp == 4, RiskPrefVal := (1500+625)/2-1000]
rkL[Change.rp == 5, RiskPrefVal := (1500+500)/2-1000]
rkL[Change.tp1 == 2, TimePref1Val := 100]
rkL[Change.tp1 == 3, TimePref1Val := 200]
rkL[Change.tp1 == 4, TimePref1Val := 400]
rkL[Change.tp1 == 5, TimePref1Val := 600]
rkL[Change.tp1 == 6, TimePref1Val := 1000]
rkL[Change.tp2 == 2, TimePref2Val := 200]
```

```
rkL[Change.tp2 == 3, TimePref2Val := 400]
rkL[Change.tp2 == 4, TimePref2Val := 600]
rkL[Change.tp2 == 5, TimePref2Val := 1000]
summary(rkL[q==1 & o800 == 1L, ])
```

hhid		o800		respondent_mid		panel_no		year	
Min.	: 7010101	Min.	:1	1	: 79	2:742		Min.	:2014
1st Qu.:	7042215	1st Qu.:	:1	2	:659			1st Qu.:	2014
Median	: 7065208	Median	:1	3	: 2			Median	:2014
Mean	:12003222	Mean	:1	4	: 1			Mean	:2014
3rd Qu.:	8147815	3rd Qu.:	:1	6	: 0			3rd Qu.:	2014
Max.	:81710520	Max.	:1	7	: 1			Max.	:2014
				21:	0				

  

q		rp		tp1		tp2		Drp	
Min.	:1	Min.	:1.00	Min.	:1.00	Min.	:1.00	Min.	: NA
1st Qu.:	:1	1st Qu.:	:2.00	1st Qu.:	:1.00	1st Qu.:	:1.00	1st Qu.:	NA
Median	:1	Median	:2.00	Median	:1.00	Median	:1.00	Median	: NA
Mean	:1	Mean	:1.98	Mean	:1.01	Mean	:1.01	Mean	:NaN
3rd Qu.:	:1	3rd Qu.:	:2.00	3rd Qu.:	:1.00	3rd Qu.:	:1.00	3rd Qu.:	NA
Max.	:1	Max.	:2.00	Max.	:2.00	Max.	:2.00	Max.	: NA
								NA's	:742

  

Dtp1		Dtp2		Num.rp		Num.tp1		Num.tp2	
Min.	: NA	Min.	: NA	Min.	:0.000	Min.	:0.00	Min.	:0.000
1st Qu.:	NA	1st Qu.:	NA	1st Qu.:	:1.000	1st Qu.:	:1.00	1st Qu.:	:1.000
Median	: NA	Median	: NA	Median	:1.000	Median	:1.00	Median	:1.000
Mean	:NaN	Mean	:NaN	Mean	:0.997	Mean	:1.01	Mean	:0.995
3rd Qu.:	NA	3rd Qu.:	NA	3rd Qu.:	:1.000	3rd Qu.:	:1.00	3rd Qu.:	:1.000
Max.	: NA	Max.	: NA	Max.	:2.000	Max.	:3.00	Max.	:3.000
NA's	:742	NA's	:742						

  

Change.rp		Change.tp1		Change.tp2		RiskPrefVal		TimePref1Val	
-2:	13	4	:396	-2:	2	Min.	: 0.0	Min.	: 100
-1:	14	3	:170	-1:	2	1st Qu.:	62.5	1st Qu.:	200
2 :	34	5	:140	1 :	5	Median	:125.0	Median	: 400
3 :	487	2	: 24	2 :	31	Mean	:110.8	Mean	: 384
4 :	191	-1	: 6	3 :	364	3rd Qu.:	:125.0	3rd Qu.:	400
5 :	3	-2	: 4	4 :	322	Max.	:187.5	Max.	:1000
		(Other):	2	5 :	16	NA's	:27	NA's	:10

  

TimePref2Val	
Min.	: 200
1st Qu.:	400
Median	: 400
Mean	: 492
3rd Qu.:	600
Max.	:1000
NA's	:9

```
rkL ← rkL[q==1, ]
rsk ← rkL[, c("hhid", "respondent_mid", grepout("Ch|Val", colnames(rkL))), with = F]
setnames(rsk, c("respondent_mid", grepout("Ch", colnames(rsk))),
  c("rskmid", gsub("Change.", "", grepout("Ch", colnames(rsk)))))
ToNum ← grepout("^p", colnames(rsk))
rsk[, (ToNum) := lapply(.SD, function(x) as.numeric(as.character(x))), .SDcols = ToNum]
setnames(rsk, ToNum, c("RiskPrefIndex", "TimePref1Index", "TimePref2Index"))
rsk[, PresentBias := 0L]
rsk[TimePref1Index > TimePref2Index, PresentBias := 1L]
## keep only rational respondents of risk preferences
#rsk ← rsk[RiskPrefIndex ≥ 0, ]
saveRDS(rkL,
  paste0(pathsaveHere, "RiskPreferencesRawData.rds"))
```

```
)  
saveRDS(rsk ,  
  paste0(pathsaveHere , "RiskPreferences.rds")  
)
```

For Change.rp:

- 2-5    Single switch point from b to a.
- 0      No switch.
- 1     Multiple switch (irrational).
- 2     Single switch from a to b (irrational).

For Change.tp1, Change.tp2:

- 2-6    Single switch point from a to b.
- 0      No switch.
- 1     Multiple switch (irrational).
- 2     Single switch from b to a (irrational).