

Read and trim files for original 800 HHs

May 3, 2022

10:09

Seiro Ito

Contents

I	Read from a list	2
I.1	Read village data file	2
I.2	Assign arms	2
II	Sample selection and treatment assignment	5
II.1	Create cumulative values in admin file	5
II.2	Merge roster with admin files: ar.1	6
II.3	Merge admin files with roster: arA	8
II.4	Merge village level info with ar.1; ar	8
II.5	Merge village level info with arAll; arA	10
II.6	Attach o1600	11
II.7	Define statuses	12
III	Merge admin-roster with other files	15
III.1	Choosing sample in admin-roster	15
III.2	Attach variables from admin-roster to other files	16

This file reads data from a list `data_read_in_a_list_with_baseline_patched.rds` and `c:/data/GUK/received/cleaned` merges all non-roster files with roster-adin (ar), attaches village level information, and saves in `c:/data/GUK/analysis/save/EstimationMemo/`.

I Read from a list

In reading raw files, I added ID information (`./ID/ID_Updated_received_from_Abu.dta`) to all pages. I further added HH ID information from the admin file. (code omitted)

Use roster file as a base, pick 800 o800 by referring to JDS data.

- One cannot base ar, arA because they are admin files that do not include rejecters.

```
jds <- fread(paste0(pathreceived, "DataForJDS.prn"))
ros[, o800 := 0L]
ros[hhid %in% jds[grepl("trea", treat), hhid], o800 := 1L]
ass[, o800 := 0L]
ass[hhid %in% jds[grepl("trea", treat), hhid], o800 := 1L]
```

Roster etries:

```
addmargins(table0(ros[o800 == 1L, .(AssignOriginal, tee = 1:N),
by = .(survey, hhid)] [tee == 1, .(AssignOriginal, survey)]))
```

	survey				
AssignOriginal	1	2	3	4	Sum
traditional	140	134	134	132	540
large	180	171	174	173	698
large grace	180	172	174	171	697
cow	190	180	180	177	727
<NA>	110	86	83	55	334
Sum	800	743	745	708	2996

I.1 Read village data file

```
library(readstata13)
vr <- read.dta13(paste0(pathcleaned, "RCT_village.dta"),
generate.factors = T, nonint.factors = T)
vr <- data.table(vr)
vr[, GroupStatus := "accepted"]
vr[grepl("De", comment), GroupStatus := "group rejection"]
vr[grepl("Ero", comment), GroupStatus := "erosion"]
setnames(vr, c("comment", "randomization"), c("GroupComment", "VArm"))
```

I.2 Assign arms

probgp

	group.id	randomization0	comment
1:	70205	large grace	eroded
2:	70314	large grace	denial
3:	70317	large	denial
4:	70319	large	denial
5:	70539	traditional	denial
6:	70544	traditional	eroded

7:	70858	traditional	denial
8:	71064	cattle	eroded
9:	81483	traditional	denial
10:	81697	traditional	denial
11:	817102	traditional	eroded

There are NAs in arm assignment. Need to fill from village level info. Apply on AssignOriginal in roster file.

```
ros[, Arm := AssignOriginal]
ros[, Arm := factor(Arm, labels = armsC)]
for (gg in unique(ros[is.na(AssignOriginal) & gid %in% vr[, groupid], gid]))
  ros[is.na(AssignOriginal) & gid == gg, Arm := vr[groupid == gg, VArm]]
addmargins(table0(ros[o800 == 1L, .(Arm, tee = 1:N),
  by = .(survey, hhid)][tee == 1, .(survey, Arm)]))
```

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	200	200		200	200	800
2	190	191		172	190	743
3	188	193		174	190	745
4	168	192		171	177	708
Sum	746	776		717	757	2996

```
ass[, Arm := AssignOriginal]
ass[, Arm := factor(Arm, labels = armsC)]
ass[is.na(gid), gid := substr(hhid, 1, 5)]
for (gg in unique(ass[is.na(AssignOriginal) & gid %in% vr[, groupid], gid]))
  ass[is.na(AssignOriginal) & gid == gg, Arm := vr[groupid == gg, VArm]]
addmargins(table0(ass[o800 == 1L, .(Arm, tee = 1:N),
  by = .(survey, hhid)][tee == 1, .(survey, Arm)]))
```

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	184	189		189	179	741
2	188	191		171	188	738
3	188	192		174	190	744
4	168	192		171	177	708
Sum	728	764		705	734	2931

Check how many baseline asset samples survive through rounds.

```
ass[, InBase := F]
ass[hhid %in% hhid[survey == 1], InBase := T]
addmargins(table0(ass[o800 == 1L & InBase, .(Arm, tee = 1:N),
  by = .(survey, hhid)][tee == 1, .(survey, Arm)]))
```

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	184	189		189	179	741
2	174	181		161	169	685
3	176	182		164	169	691
4	157	182		161	156	656
Sum	691	734		675	673	2773

```
ass[, InBase := NULL]
```

Incorporate credit and borrowing data. In rd 1, out_bal = NA for all entries. Only out.ngo, out_rel, loan_amount_lender have entries. Create YBal where Y=NGO, relative, and lender.

Then sum all: OutBal as outstanding balance for relatives, NGOs, and money lenders (code omitted). In rd1, OutBal = out_rel+out_ngo+loan_amount_lender, in rd 2, 4, OutBal = out_bal+sum(out_bal_X)+sum(loan_amount_lender_X). In rd 3, there is no data.

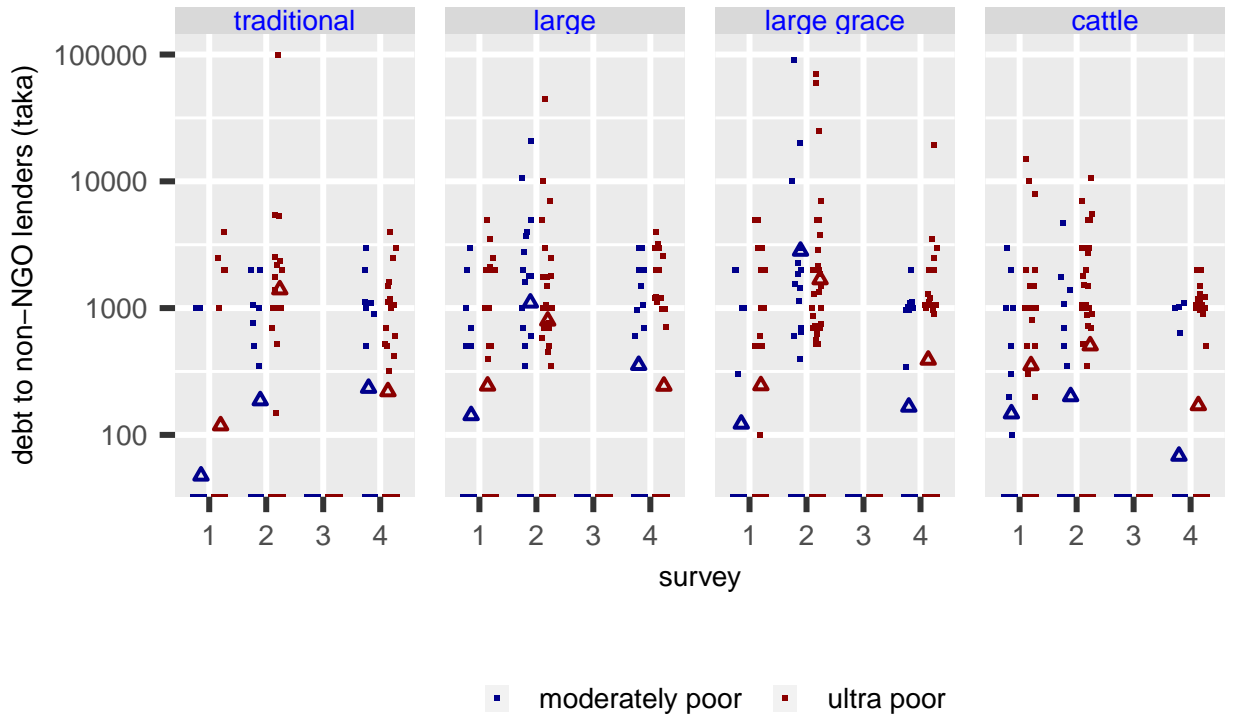
TABLE 1: NUMBER OF OBSERVATIONS IN OTHER BORROWING

	lender	round 1	round 2	round 4
any debt		425	1298	1053
NGO		14	1057	572
GUK			34	871
relatives		258	265	132
money lenders		157	291	191
non-NGO		411	489	318

TABLE 1 shows the growing numbers of NGO debts. This should include GUK borrowing. I created NetOutBal = OutBal-GUK borrowing to get just non-GUK debts, but this does not give any insight because households do not always report lender as GUK.

We define NonNGOBal as non-NGO debt (relative + money lender), median debts are 500, 1520, NA, 1116, mean debts are 1256, 3686, NaN, 1376 in round 1, 2 and 4. Mean debts by arm are 2150, 2163, 2521, NaN for traditional, large, large grace, and cattle.

FIGURE 1: INDEBTEDNESS TO NON-NGOs



Note: Sum of debts to relatives and money lenders in the last 12 months of survey. Each dots represent one observation, filled triangles show the group means of indebtedness.

FIGURE 1 gives borrowing from friends/relatives and money lenders. If we exclude loans from GUK, they are the only sources of borrowing for the households in our study area. Both box plots and means show an increased borrowing in round 2, but it decreased in round 4 to the pre-intervention level. This suggests the repayment schedule that we intended to adopt the heifer growth cycle is still not suited to the actual cash flow profiles, and such gaps might have induced the borrowers to get liquidity elsewhere for installments during the second round. If the households could foresee the

gap in cash flows, they might have not participated the program. If the households could not foresee the gap and did not have credit access other than GUK, they might have ended up as a delinquent borrower.

Define shock variables. FloodInRd1 is reported flood damage in code_1 in round 1. (code omitted)

Description of data files:

- ad Administrative data: Up to [-24, 48] months after first loan disbursement. This file has not been used in read_cleaned_data.rnw. o800 is imported from JDS file.
- ros roster to condition the initial status prior to participation. This is the base of all files, and includes variables o800 and Arm.
- sch Schooling panel with attrition. Aged 6-18 in rd1. Enrolled={0,1} is defined for children aged 6-18 in rd1 by referencing to currently_enrolled and age information.
- ass MergedAssets. Merged from several tables: household assets (houses, durables), productive assets (machines, tools), and ownership and contract (land holding; operated area, owned area).
- lvo Livestock holding.
- lab Labour incomes.
- far Farming revenues (no costs reported).
- con Household consumption. Food expenditure asks both bought and consumed volumes and prices. We impute consumption values by using median prices. All quantity is set to annualised quantity.
- obr Other borrowing (pages under “credit and borrowing”) from relatives and money lenders. To be merged with MergedAssets in this file.
- shk Shocks. Merged with all other files.

II Sample selection and treatment assignment

II.1 Create cumulative values in admin file

Read admin files.

```
adw3 <- readRDS(paste0(path1234, "admin_data_wide2.rds"))
adw3[, MemNum := 1:N, by = .(hhid, Year)]
adw3[, tee := 1:N, by = hhid]
```

Number of meetings in recorded in admin file: there are 1999 entries that have 48 meetings.

1999
48

Add rolling means.

```
# add rolling means
library(zoo)
rollvars <- c("value.missw", "value.repay", "value.NetSaving", "OtherNetSaving", "OtherRep
ad0[, (paste0("RM", rollvars)) := lapply(.SD, rollmean, k = 6, na.pad = TRUE),
  by = hhid, .SDcols = rollvars]
# lag rolling means by 3 months to get previous 6 month averages
```

```
ad0[, (paste0("RM", rollvars)) := shift(.SD, n=3, type = "lag"),
      by = hhid, .SDcols = paste0("RM", rollvars)]
```

II.2 Merge roster with admin files: ar.1

Create `adbase` (baseline fixed characteristics `creditstatus`, `Mem`, `povertystatus`, `DisDate1` taken from `ad0` (=c:/data/GUK/received/cleaned_by_RA/clean_panel_data_by_section/admin_data_wide2.rds). Merge it with roster. This gives fixed characteristics of membership attached with roster. Name the resulting data as `ar.0`.

Check the completeness of roster file.

```
addmargins(table0(ros[o800 == 1L, .(teenum = 1:N, Arm)],
  by = .(hhid, survey))[teenum == 1, .(survey, Arm)], 2)
```

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	200	200		200	200	800
2	190	191		172	190	743
3	188	193		174	190	745
4	168	192		171	177	708

Tabulation of arms with `ar.0` for `o800`. There are 0 NAs which will be filled in with `RCT_village.dta` with `ar`, `arAll` in the next subsection.

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	200	200		200	200	800
2	190	191		172	190	743
3	188	193		174	190	745
4	168	192		171	177	708

Create `adrest`: Time-variant characteristics in `ad0`. Merge with `ar.0`. Name resulting data as `ar.1`. `ar.1` is roster with fixed and variable characteristics found in admin data.

```
setkey(ar.0, groupid, hhid, Year, Month)
setkey(adrest, groupid, hhid, Year, Month)
ar.1 <- adrest[ar.0]
ar1vars <- c("Arm", "groupid", "creditstatus", "DisDate1", "Mship", "Mem")
for (i in ar1vars)
  ar.1[, (i) := eval(parse(text=
    paste0(i, "[!is.na(", i, ")][1]"))
  ), by = hhid]
for (i in grepout("Arm|rand|Assi", ar1vars)) {
  ar.1[grepl("largeG", eval(parse(text=i))), (i) := "large grace"]
  ar.1[, (i) := factor(eval(parse(text=i)),
    levels = c("traditional", "large", "large grace", "cattle"))]
}
setorder(ar.1, hhid, survey, IntDate, mid)
ar.1[, MemNum := 1:N, by = .(hhid, survey, IntDate)]
```

Errors in `Mstatus` in `ar.1`. Correct to `oldMember`. (No corresponding entry in `arAll` because it is data only for members.)

	hhid	survey	CumRepaid	CumNetSaving	TradGroup	Date	DisDate1
1:	7137220	1	NA	NA	<NA>	<NA>	2013-11-01
2:	7137220	2	1750	0	planned	2014-10-01	2013-11-01
3:	7137220	3	4250	385	planned	2015-11-01	2013-11-01
	creditstatus		Mship	Mstatus			

```

1:      Yes oldMember iRejection
2:      Yes oldMember iRejection
3:      Yes oldMember iRejection

```

There are 0 members (oldMember in Mstatus), 20 members (newGroup in Mstatus), 3 members (iReplacement in Mstatus) who did not borrow but only saved. This is identified by DisDate1 == NA & EverSaved & creditstatus == No.

```

      Arm      DisDate1  EverSaved  creditstatus      Mstatus
traditional:23  Min.    :NA      Mode:logical  Yes: 0      gErosion    : 0
large          : 0  1st Qu.:NA      TRUE:23      No :23      gRejection  : 0
large grace: 0  Median :NA                                iRejection  : 0
cattle        : 0  Mean   :NA                                iReplacement: 3
                                     3rd Qu.:NA      newGroup    :20
                                     Max.    :NA      oldMember    : 0
                                     NA's    :23
Mship
oldMember      : 0
newMember      :23
quitMembership: 0

```

There are also members who were offered membership but never took up. This is identified by DisDate1 == NA & !EverSaved & !EverRepaid.

```

      DisDate1  EverSaved  EverRepaid  creditstatus      Mstatus
Min.    :NA      Mode :logical  Mode :logical  Yes:0      gErosion    :0
1st Qu.:NA      FALSE:2      FALSE:2      No :2      gRejection  :0
Median  :NA                                iRejection  :2
Mean    :NA                                iReplacement:0
3rd Qu.:NA                                newGroup    :0
Max.    :NA                                oldMember    :0
NA's    :2
Arm
traditional:2
large      :0
large grace:0
cattle     :0

```

Create BorrowerStatus to indicate these guys (DisDate1 == NA & EverSaved & creditstatus == No) as a pure saver. 2 entries with DisDate1 == NA & !EverSaved & !EverRepaid are people who quit so set as quit membership.

```

      BorrowerStatus
Mstatus  borrower pure saver quit membership  Sum
gErosion      80      0      0      0      80
gRejection    140      0      0      0     140
iRejection    157      0      2      2     159
iReplacement  112      3      0      0     115
newGroup      388     20      0      0     408
oldMember    1221      0      0      0    1221
Sum          2098     23      2      2    2123

```

In roster + admin (base: roster): Tabulate hhid observations by survey round and Arm before supplementing with AssignOriginal and VArm. Note: 0 observations with NA are also pointed in read_cleaned_data.rnw and are going to be dealt with in the next subsection.

```

      Arm
survey traditional large large grace cattle  Sum
1      605      504      507      507  2123
2      585      485      447      466  1983
3      582      487      452      472  1993
4      540      483      447      444  1914

```

Sum	2312	1959	1853	1889	8013
-----	------	------	------	------	------

II.3 Merge admin files with roster: arA

Create arAll: admin data ad0 with period 1 roster data. (Roster information is added only if matched with admin HH IDs.) We define rd = 4 even if Date > IntDate.4.

Some HHs in admin file are not found in roster.

[1]	9807042003	9807042011	9807042514	9807042706	9807042710	9807054106
[7]	9807054304	9807054520	9807064605	9807064607	9807064612	9807064617
[13]	9807064619	9807065207	9807065208	9807065212	9807065306	9807065307
[19]	9807065313	9807065315	9807065316	9807065319	9807075702	9807085904
[25]	9807085914	9807086106	9807086107	9807106513	9807106517	9807106518
[31]	9807126819	9807126820	9807127103	9807127105	9807127106	9807127108
[37]	9807137203	9807137204	9807137206	9807137217	9807137218	9808169816
[43]	9907065108	9907075402	9907075405	9907075406	9907075407	9907075410
[49]	9907075411	9907075413	9907075418	9907075419	9907075420	98081710308
[55]	98081710317	99070210905	99070210906	99070211813	99070310702	99070311402
[61]	99070311403	99070311405	99070311407	99070311408	99070311411	99070311412
[67]	99070311415	99070311416	99070311419	99070311501	99070311502	99070311505
[73]	99070311507	99070311508	99070311509	99070311511	99070311513	99070311517
[79]	99070311520	99070712701	99070712703	99070712704	99070712707	99070712708
[85]	99070712710	99070712713	99070712714	99070712716	99070712720	99071010811
[91]	99071010813	99071010814	99071010819	99081711206	99081711207	99081711208

Refer to ID file c:/data/GUK/received/cleaned_by_RA/clean_panel_data.by_section/ID.rds to see their Mstatus. They are all new groups and individual replacing members who are not originally included in the baseline survey.

Mstatus	Assign					Sum
	traditional	large	large	grace	cow	
iReplacement	14	4		11	15	44
newGroup	34	4		4	10	52
Sum	48	8		15	25	96

Drop these from arAll who are missing in survey but found in admin. (They are kept in ar). This results in reduction in observations with 48 meetings. There are 1903 households who have 48 rows in data.

In arAll, nonmembers (gRejection, gErosion) are not included.

Mstatus	Mship			Sum
	oldMember	newMember	quitMembership	
gErosion	0	0	0	0
gRejection	0	0	0	0
iRejection	1	0	159	160
iReplacement	0	115	0	115
newGroup	0	408	0	408
oldMember	1220	0	0	1220
Sum	1221	523	159	1903

II.4 Merge village level info with ar.1; ar

Create ar: ar.1 + vr (RCT_village.dta). Roster as base + admin.

BorrowerStatus	AttritIn				Sum
	2	3	4	9	
borrower	46	30	390	1514	1980
pure saver	0	0	24	729	753

quit membership	6	9	8	483	506
Sum	52	39	422	2726	3239

```

      EverRepaid
DisDate3NA <NA>
      TRUE 1980

```

```
< table of extent 0 x 0 >
```

I use arm VArm from village level information. Tabulation of AssignOriginal against VArm shows complementarity so I can use one variable to fill in NAs in another.

	VArm					
Arm	traditional	large	large	grace	cattle	Sum
traditional	200	0		0	0	200
large	0	200		0	0	200
large grace	0	0		200	0	200
cattle	0	0		0	200	200
Sum	200	200		200	200	800

	VArm						
Arm	traditional	large	large	grace	cattle	<NA>	Sum
traditional	439	0		0	0	166	605
large	0	408		0	0	96	504
large grace	0	0		411	0	96	507
cattle	0	0		0	457	50	507
Sum	439	408		411	457	408	2123

Tabulation of Arm after supplementing with VArm.

```

ar[, ArmBefore := Arm]
ar[is.na(Arm) & !is.na(VArm), Arm := VArm]

```

```
addmargins(table0(ar[o800 == 1L & MemNum == 1, .(survey, Arm)]))
```

	Arm					
survey	traditional	large	large	grace	cattle	Sum
1	200	200		200	200	800
2	190	191		172	190	743
3	188	193		174	190	745
4	168	192		171	177	708
Sum	746	776		717	757	2996

```
addmargins(table0(ar[MemNum == 1, .(survey, Arm)]))
```

	Arm					
survey	traditional	large	large	grace	cattle	Sum
1	605	504		507	507	2123
2	585	485		447	466	1983
3	582	487		452	472	1993
4	540	483		447	444	1914
Sum	2312	1959		1853	1889	8013

Below is what is supplemented from VArm of village level information to the 0 NAs.

```

ar[MemNum == 1 & survey == 1 & is.na(ArmBefore),
  BorrowerStatus := "quit membership"]
table0(ar[MemNum == 1 & survey == 1 & is.na(ArmBefore) & !is.na(VArm),
  .(BorrowerStatus, Arm)])

```

```
< table of extent 3 x 0 >
```

```
table0(ar[o800 == 1L & MemNum == 1 & survey == 1 & is.na(ArmBefore) & !is.na(VArm),
.(BorrowerStatus, Arm))
```

```
< table of extent 3 x 0 >
```

```
ar[, ArmBefore := NULL]
```

II.5 Merge village level info with arAll; arA

Create arA: arAll (admin data as base + roster) + vr (village randomisation)

Tabulation of BorrowerStatus in arA at round 1.

BorrowerStatus	Arm					Sum
	traditional	large	large	grace	cattle	
borrower	383	452		445	415	1695
pure saver	49	0		0	0	49
quit membership	53	12		22	72	159
Sum	485	464		467	487	1903

Tabulation of Mstatus in arA at round 1.

Mstatus	Arm					Sum
	traditional	large	large	grace	cattle	
gErosion	0	0		0	0	0
gRejection	0	0		0	0	0
iRejection	53	12		22	72	159
iReplacement	39	8		11	57	115
newGroup	166	96		96	50	408
oldMember	227	348		338	308	1221
Sum	485	464		467	487	1903

Tabulation of Mstatus in ar at round 1.

Mstatus	Arm					Sum
	traditional	large	large	grace	cattle	
gErosion	40	0		20	20	80
gRejection	80	40		20	0	140
iRejection	53	12		22	72	159
iReplacement	39	8		11	57	115
newGroup	166	96		96	50	408
oldMember	227	348		338	308	1221
Sum	605	504		507	507	2123

adw3 idfu[adw2]: admin data adw2 + idfu (arm information)

ad0 Selected columns of adw3.

Base: roster.

ar.0 adbase[ros]: ros (33223, 37) + invariant portion of admin data ad0 (1999, 5).

ar.1 adrest[ar.0]: ar.0 (33223, 40) + variable portion of admin data ad0 (95952, 45).

ar vr[ar.1]: ar.1 (33223, 83) + vr (RCT_village.dta) (80, 4), resulting in (33223, 87). Number of individuals: 2123.

Base: admin. This has a smaller number of individuals because admin data do not include individuals who left the group.

ar.00 ros.00W[ad0]: ad0 (95952, 49) + ros.00W (survey round info) (2123, 5).

arAll ros.0[ar.00]: ar.00 (admin data with survey round info) (95952, 50) + ros.0 (roster only with first observed round) (2123, 11).

arA vr[arAll]: arAll (admin data as base + roster) (91344, 63) + vr (village randomisation) (80, 4), resulting in (91344, 68). Number of individuals: 1903.

II.6 Attach o1600

Create o1600 to indicate the original 1600 HHs. Tabulation of total observations in roster ar by o1600 and survey.

	survey			
o1600	1	2	3	4
0	2503	2510	2543	2457
1	6130	5817	5843	5420

Tabulation of total observations in roster ar by o1600 and survey after restricting to 1 obs per HH.

	survey			
o1600	1	2	3	4
0	627	611	616	607
1	1496	1372	1377	1307
Sum	2123	1983	1993	1914

Tabulation of o800 for ar.

	Arm					
survey	traditional	large	large	grace	cattle	Sum
1		200	200	200	200	800
2		190	191	172	190	743
3		188	193	174	190	745
4		168	192	171	177	708

At rd 1.

	Arm					
Mstatus	traditional	large	large	grace	cattle	Sum
gErosion		20	0	10	10	40
gRejection		40	20	10	0	70
iRejection		31	9	13	37	90
iReplacement		0	0	0	0	0
newGroup		0	0	0	0	0
oldMember		109	171	167	153	600
Sum		200	200	200	200	800

At rd 4. There are 92 attritions.

	Arm					
Mstatus	traditional	large	large	grace	cattle	Sum
gErosion		0	0	0	0	0
gRejection		0	0	0	0	0
iRejection		61	28	11	30	130
iReplacement		0	0	0	0	0
newGroup		0	0	0	0	0
oldMember		107	164	160	147	578
Sum		168	192	171	177	708

II.7 Define statuses

Check AttritIn consistency. Define the observed largest survey rounds and tabulate against AttritIn.

	AttritIn				
Tee	2	3	4	9	Sum
1	41	0	0	0	41
2	0	14	0	0	14
3	0	0	37	0	37
4	0	0	0	708	708
Sum	41	14	37	708	800

	AttritIn				
Mstatus	2	3	4	9	Sum
gErosion	12	0	28	0	40
gRejection	11	4	0	55	70
iRejection	10	4	1	75	90
iReplacement	0	0	0	0	0
newGroup	0	0	0	0	0
oldMember	8	6	8	578	600
Sum	41	14	37	708	800

Tabulation for arA. It has survey == 5 which are meetings after the rd 4 interview. arA has fewer observations per meeting than ar when only using 1 obs per rd,

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	12	21		30	49	112
2	167	343		342	346	1198
3	165	341		338	335	1179
4	165	343		343	342	1193
Sum	509	1048		1053	1072	3682

but more observations per round because there are multiple meetings per round.

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	445	845		967	1886	4143
2	3054	6197		6221	6156	21628
3	2220	4650		4607	4596	16073
4	2681	5588		5485	5602	19356
Sum	8400	17280		17280	18240	61200

Tabulation of o800 for arA. It has a smaller number of obs than ar because it does not include rejecters or flood evacuees.

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	9	10		14	18	51
2	134	171		172	180	657
3	133	170		169	175	647
4	132	173		171	177	653

Refer to ar to see who are missing in arA.

	Arm					
Mstatus	traditional	large	large	grace	cattle	Sum
gErosion	20	0		10	10	40
gRejection	40	20		10	0	70
iRejection	0	0		0	0	0
iReplacement	0	0		0	0	0

newGroup	0	0	0	0	0
oldMember	0	0	0	0	0
Sum	60	20	20	10	110

Initial period obs matches with loan recipients of ar.

	Arm					
Mstatus	traditional	large	large	grace	cattle	Sum
gErosion	0	0		0	0	0
gRejection	0	0		0	0	0
iRejection	31	9		13	37	90
iReplacement	0	0		0	0	0
newGroup	0	0		0	0	0
oldMember	109	171		167	153	600
Sum	140	180		180	190	690

In ar, there are 114 cases of group rejections in GroupStatus classified as individual rejections in Mstatus. Overwrite Mstatus with GroupStatus in these cases, which results in the below:

```
addmargins(table0(ar[tee == 1 & MemNum == 1, .(Mstatus, GroupStatus)]))
```

	GroupStatus				
Mstatus	accepted	erosion	group	rejection	Sum
gErosion	0	80		0	80
gRejection	0	0		140	140
iRejection	159	0		0	159
iReplacement	115	0		0	115
newGroup	408	0		0	408
oldMember	1221	0		0	1221
Sum	1903	80		140	2123

```
ar[grepl("iR", Mstatus) & grepl("rej", GroupStatus), Mstatus := "gRejection"]
ar[, tee:= 1:N, by = hhid]
addmargins(table0(ar[tee == 1 & MemNum == 1, .(GroupStatus, Arm)]))
```

	Arm					
GroupStatus	traditional	large	large	grace	cattle	Sum
accepted	485	464		467	487	1903
erosion	40	0		20	20	80
group rejection	80	40		20	0	140
Sum	605	504		507	507	2123

Define BStatus.

```
datas <- c("ar", "arA")
for (i in 1:length(datas))
{
  dd <- get(datas[i])
  dd[, BStatus := BorrowerStatus]
  dd[grepl("gRe", Mstatus), BStatus := "group rejection"]
  dd[grepl("iRej", Mstatus), BStatus := "individual rejection"]
  dd[grepl("gEr", Mstatus), BStatus := "rejection by flood"]
  dd[, BStatus := factor(BStatus, levels = c("borrower", "pure saver",
    "individual rejection", "group rejection", "rejection by flood"))]
  assign(datas[i], dd)
}
addmargins(table0(ar[o800 == 1L & MemNum == 1 & survey == 1, .(BStatus, AttritIn)]))
```

	AttritIn				
BStatus	2	3	4	9	Sum

borrower	8	6	8	578	600
pure saver	0	0	0	0	0
individual rejection	10	4	1	75	90
group rejection	11	4	0	55	70
rejection by flood	12	0	28	0	40
Sum	41	14	37	708	800

```
if (any(ar[, is.na(BStatus)]))
  addmargins(table0(ar[is.na(BStatus)&tee == 1&o800==1L, .(Mstatus, BorrowerStatus)]))
```

If any: 0 NAs in BStatus are borrowers. Correct it.

For o800:

BStatus	GroupStatus				Sum
	accepted	erosion	group	rejection	
borrower	600	0		0	600
pure saver	0	0		0	0
individual rejection	90	0		0	90
group rejection	0	0		70	70
rejection by flood	0	40		0	40
Sum	690	40		70	800

Arm	AttritIn				Sum
	2	3	4	9	
traditional	8	4	20	168	200
large	5	2	1	192	200
large grace	23	3	3	171	200
cattle	5	5	13	177	200
Sum	41	14	37	708	800

BStatus	AttritIn				Sum
	2	3	4	9	
borrower	8	6	8	578	600
pure saver	0	0	0	0	0
individual rejection	10	4	1	75	90
group rejection	11	4	0	55	70
rejection by flood	12	0	28	0	40
Sum	41	14	37	708	800

If we exclude twice or double disbursements in traditional (24 members).

Arm	AttritIn				Sum
	2	3	4	9	
traditional	8	4	20	144	176
large	5	2	1	192	200
large grace	23	3	3	171	200
cattle	5	5	13	177	200
Sum	41	14	37	684	776

For traditional arm.

BStatus	AttritIn				Sum
	2	3	4	9	
borrower	1	0	1	83	85
pure saver	0	0	0	0	0
individual rejection	4	1	1	25	31
group rejection	1	3	0	36	40
rejection by flood	2	0	18	0	20
Sum	8	4	20	144	176

III Merge admin-roster with other files

III.1 Choosing sample in admin-roster

In ar: Keep if Mstatus includes strings old, iRej, gEro, gRej, & TradGroup does not include strings tw (relaxing DisDate1 is before 2015-01-01). There are 776 HHs at the baseline. [This the admin data used in this note](#). This also shows a lower attrition rate for large arm.

```
addmargins(table0(ar[o800 == 1L & grepl("old|iRej|^g", Mstatus) &
!grepl("tw", TradGroup) & MemNum == 1, .(survey, Arm)]))
```

	Arm					
survey	traditional	large	large	grace	cattle	Sum
1	176	200		200	200	776
2	166	191		172	190	719
3	164	193		174	190	721
4	144	192		171	177	684
Sum	650	776		717	757	2900

In ar, as one can see below, gRejection is more frequent in traditional and large, while there is none in cattle. traditional, cattle have more frequent iRejection. So traditional was disliked both at group and individual levels, large was disliked as a group, cattle was disliked at an individual level, and large grace were well received at both group and individual levels. This indicates attractiveness of a grace period at least at the group level, and a large cash form (over small cash or in-kind) at the individual level.

	Arm					
Mstatus	traditional	large	large	grace	cattle	Sum
gErosion		40	0		20	80
gRejection	80	40		20	0	140
iRejection	53	12		22	72	159
iReplacement	39	8		11	57	115
newGroup	166	96		96	50	408
oldMember	227	348		338	308	1221
Sum	605	504		507	507	2123

In ar, for o800 we have:

	Arm					
Mstatus	traditional	large	large	grace	cattle	Sum
gErosion	20	0		10	10	40
gRejection	40	20		10	0	70
iRejection	31	9		13	37	90
iReplacement	0	0		0	0	0
newGroup	0	0		0	0	0
oldMember	109	171		167	153	600
Sum	200	200		200	200	800

arA is used in saving and repayment regressions. Contrast it with arA:

	Arm					
Mstatus	traditional	large	large	grace	cattle	Sum
gErosion	0	0		0	0	0
gRejection	0	0		0	0	0
iRejection	31	9		13	37	90
iReplacement	0	0		0	0	0
newGroup	0	0		0	0	0
oldMember	109	171		167	153	600
Sum	140	180		180	190	690

Create LYear.

Save roster-admin data to c:/data/GUK/analysis/save/EstimationMemo/.

```
saveRDS(ar, paste0(pathsaveHere, "RosterAdminData.rds"))
saveRDS(arA, paste0(pathsaveHere, "AllMeetingsRosterAdminData.rds"))
fwrite(ar, paste0(pathsaveHere, "RosterAdminData.prn"), sep = "\t", quote = F)
fwrite(arA, paste0(pathsaveHere, "AllMeetingsRosterAdminData.prn"), sep = "\t", quote = F)
```

Schooling.

III.2 Attach variables from admin-roster to other files

Attach Arm, TradGroup, Mem, ObPattern, AttritIn, o1600, Mstatus, BorrowerStatus, BStatus, creditstatus, povertystatus, RMvalue.repay, RMvalue.NetSaving, RMOtherNetSaving, RMOtherRepaid, HHsize, HeadLiteracy, IntDate, DisDate1 from ar.

```
vartoattach <- c("Arm", "TradGroup", "Mem",
  "ObPattern", "AttritIn", "o1600",
  "Mstatus", "BorrowerStatus", "BStatus",
  "creditstatus", "povertystatus", "RMvalue.repay",
  "RMvalue.NetSaving", "RMOtherNetSaving", "RMOtherRepaid",
  "HHsize", "HeadLiteracy", "IntDate", "DisDate1")
dfiles <- c("ass", "s1", "lvo", "lvoL", "lvp", "lab", "far", "con", "obr", "shk")
for (j in 1:length(dfiles)) {
  dd <- get(dfiles[j])
  if (!any(grepl("groupid", colnames(dd)))) {
    dd[, groupid := as.integer(as.numeric(as.character(gid)))]
    dd[, gid := NULL]
  }
  dd[, Year := as.numeric(format(as.Date(IntDate), "%Y"))]
  dd[, Month := as.character(format(as.Date(IntDate), "%B"))]
  dd[Year <= 2010, Year := Year + 10]
  # attach o800
  dd[, o800 := 0L]
  dd[hhid %in% jds[grepl("trea", treat), hhid], o800 := 1L]
  # drop all variables in each page before copying from ar0
  dd[, (vartoattach) := NULL]
  setorder(dd, groupid, hhid, survey, Year, Month)
  setkey(dd, groupid, hhid, survey)
  if (j < length(dfiles)) dd <- ar0[dd]
  assign(dfiles[j], dd)
}
```

[1] "dummyHadCows"	"dummyHadCows.Time3"
[3] "dummyHadCows.Time4"	"dummyTraditional.dummyHadCows"
[5] "dummyLarge.dummyHadCows"	"dummyLargeGrace.dummyHadCows"
[7] "dummyCattle.dummyHadCows"	"dummyTraditional.dummyHadCows.Time3"
[9] "dummyLarge.dummyHadCows.Time3"	"dummyLargeGrace.dummyHadCows.Time3"
[11] "dummyCattle.dummyHadCows.Time3"	"dummyTraditional.dummyHadCows.Time4"
[13] "dummyLarge.dummyHadCows.Time4"	"dummyLargeGrace.dummyHadCows.Time4"
[15] "dummyCattle.dummyHadCows.Time4"	

```
dummyHadCows.Time3
Min.      :0.0000
1st Qu.   :0.0000
Median    :0.0000
Mean      :0.0617
```



```
3rd Qu.:0.0000
Max.    :1.0000
```

Check number of HHs in assets by o1600:

```
addmargins(table(ass[, .(creditstatus, survey, o1600)]))
```

```
, , o1600 = 0

      survey
creditstatus  1    2    3    4  Sum
Yes      571  587  593  586 2337
No       23   23   23   21   90
Sum      594  610  616  607 2427

, , o1600 = 1

      survey
creditstatus  1    2    3    4  Sum
Yes     1012 1040 1052 1039 4143
No       172  150  155  154  631
Sum     1184 1190 1207 1193 4774

, , o1600 = Sum

      survey
creditstatus  1    2    3    4  Sum
Yes     1583 1627 1645 1625 6480
No       195  173  178  175  721
Sum     1778 1800 1823 1800 7201
```

```
addmargins(table(ass[o800 == 1, .(survey, creditstatus)]))
```

```
      creditstatus
survey  Yes    No   Sum
1       555    84   639
2       580    72   652
3       585    76   661
4       578    75   653
Sum     2298   307 2605
```

```
#table0(ass[o1600 == 0L, .(creditstatus, survey)])
```

Check number of HHs in schooling by o1600:

```
table(s1[, .(Schooling, survey, o1600)])
```

```
, , o1600 = 0

      survey
Schooling  1    2    3    4
primary0512 528  427  361  202
junior1315  133  129  140  204
high1618    94   94   93  111

, , o1600 = 1

      survey
Schooling  1    2    3    4
primary0512 1318  911  659  322
junior1315  307  279  427  499
high1618    202  198  179  225
```

```
addmargins(table(s1[o800 == 1, .(survey, Schooling)]))
```

	Schooling			
survey	primary0512	junior1315	high1618	Sum
1	695	159	110	964
2	483	147	105	735
3	344	230	90	664
4	165	264	115	544
Sum	1687	800	420	2907

Check number of o800 HHs in ar:

```
ar[, tee := as.integer(1:.N), by = .(hhid, survey)]
addmargins(table0(ar[tee == 1 & o800 == 1L, .(survey, Arm)]))
```

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	200	200		200	200	800
2	190	191		172	190	743
3	188	193		174	190	745
4	168	192		171	177	708
Sum	746	776		717	757	2996

Check number of o800 HHs in arA:

```
arA[, tee := as.integer(1:.N), by = .(hhid, survey)]
addmargins(table0(arA[tee == 1 & o800 == 1L, .(survey, Arm)]))
```

Arm						
survey	traditional	large	large	grace	cattle	Sum
1	9	10		14	18	51
2	134	171		172	180	657
3	133	170		169	175	647
4	132	173		171	177	653
Sum	408	524		526	550	2008

Number of observations differ between ar and arA because the latter does not include rejecters.

	traditional	large	large	grace	cattle	traditional	large	large	grace
gErosion	20	0		10	10	0	0		0
gRejection	40	20		10	0	0	0		0
iRejection	31	9		13	37	31	9		13
iReplacement	0	0		0	0	0	0		0
newGroup	0	0		0	0	0	0		0
oldMember	109	171		167	153	109	171		167
	cattle								
gErosion	0								
gRejection	0								
iRejection	37								
iReplacement	0								
newGroup	0								
oldMember	153								

Original 800 households in arA (members only).

```
[1] TRUE
```

	EverRepaid	
Mstatus	TRUE	Sum
gErosion	0	0
gRejection	0	0
iRejection	0	0

iReplacement	0	0
newGroup	0	0
oldMember	600	600
Sum	600	600

What is relevant in estimation is observations by LoanYear, total of 600.

	Arm					
LoanYear	traditional	large	large	grace	cattle	Sum
1	109	171		167	153	600
2	109	171		167	153	600
3	109	171		167	153	600
4	109	171		167	153	600
Sum	436	684		668	612	2400

If we restrict to planned in TradGroup, number of observation becomes 576.

	Arm					
LoanYear	traditional	large	large	grace	cattle	Sum
1	85	171		167	153	576
2	85	171		167	153	576
3	85	171		167	153	576
4	85	171		167	153	576
Sum	340	684		668	612	2304

ObPattern in original 800.

	Arm					
ObPattern	traditional	large	large	grace	cattle	Sum
0111	0	2		3	1	6
1000	1	5		1	1	8
1010	0	1		0	0	1
1011	0	0		0	0	0
1100	0	1		3	2	6
1110	1	0		3	3	7
1111	107	162		157	146	572
Sum	109	171		167	153	600

BorrowerStatus pattern in original 800.

	Arm					
BorrowerStatus	traditional	large	large	grace	cattle	Sum
borrower		109	171		167	153 600
pure saver		0	0		0	0 0
quit membership		0	0		0	0 0
Sum		109	171		167	153 600

BStatus pattern in original 800.

	Arm					
BStatus	traditional	large	large	grace	cattle	Sum
borrower		109	171		167	153 600
pure saver		0	0		0	0 0
individual rejection		0	0		0	0 0
group rejection		0	0		0	0 0
rejection by flood		0	0		0	0 0
Sum		109	171		167	153 600

BorrowerStatus pattern in original 800 with only planned in TradGroup.

	Arm					
BorrowerStatus	traditional	large	large	grace	cattle	Sum
borrower		85	171		167	153 576
pure saver		0	0		0	0 0

quit membership	0	0	0	0	0
Sum	85	171	167	153	576

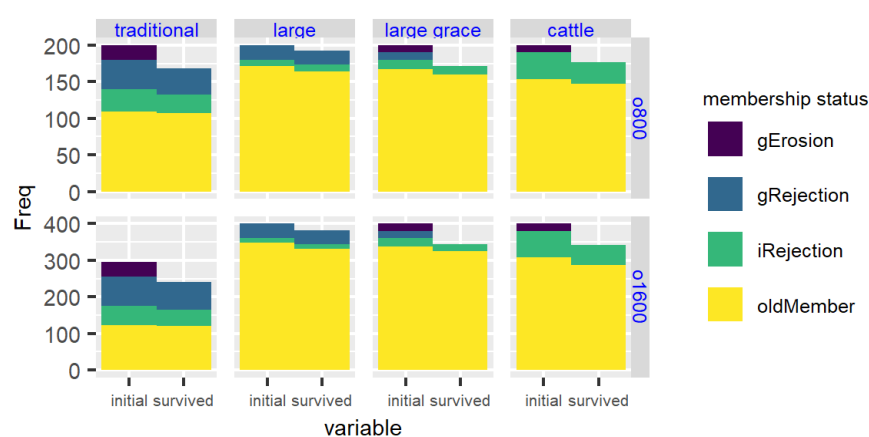
BStatus pattern in original 800 with only planned in TradGroup.

BStatus	Arm					Sum
	traditional	large	large grace	cattle		
borrower	85	171	167	153		576
pure saver	0	0	0	0		0
individual rejection	0	0	0	0		0
group rejection	0	0	0	0		0
rejection by flood	0	0	0	0		0
Sum	85	171	167	153		576

Below tabulates attrition pattern in ar for 800 and 1600 households.

pdf
2

FIGURE 2: ATTRITION AND MEMBERSHIP STATUS AMONG ORIGINAL 800 AND 1600 HOUSEHOLDS



Source: Survey and administrative data. ar

Note: Top panel: Membership status and respective non-attrition in o800. Bottom panel: Membership status and respective non-attrition in o1600.

Save all data in c:/data/GUK/analysis/save/EstimationMemo/.

```
fwrite(s1, paste0(pathsaveHere, "RosterAdminSchoolingData.prn"), sep = "\t", quote = F)
fwrite(ass, paste0(pathsaveHere, "AssetAdminData.prn"), sep = "\t", quote = F)
fwrite(lvoL, paste0(pathsaveHere, "LivestockLongAdminData.prn"), sep = "\t", quote = F)
fwrite(lvo, paste0(pathsaveHere, "LivestockAdminData.prn"), sep = "\t", quote = F)
fwrite(lvp, paste0(pathsaveHere, "LivestockProductsAdminData.prn"), sep = "\t", quote = F)
fwrite(lab, paste0(pathsaveHere, "LabourIncomeAdminData.prn"), sep = "\t", quote = F)
fwrite(far, paste0(pathsaveHere, "FarmRevenueAdminData.prn"), sep = "\t", quote = F)
fwrite(con, paste0(pathsaveHere, "ConsumptionAdminData.prn"), sep = "\t", quote = F)
fwrite(obr, paste0(pathsaveHere, "OtherBorrowingAdminData.prn"), sep = "\t", quote = F)
fwrite(shk, paste0(pathsaveHere, "Shocks.prn"), sep = "\t", quote = F)
```