Comparing outcomes between groups

March 15, 2017
07:27

Seiro Ito

# Contents

I Read 2

II Treatment through time 2

III Food consumption (23B in rd 1, 3B in rds 2, 3) 2

IV Credit use 3

V New loans 4

VI Assets 6

VII livestock 9

VIII livestock 9

There are a few key variables.

- receivedCredit: The actual treatment status. This is time-invariant. It is T if a subject is receives a loan in our observation period.
- assignment: The original treatment assignement. This is time-invariant. This differs from receivedCredit because in our design everyone is deemed to get treated at the end (but there are subjects who opted out of a loan but remained in a group).
- disbursed: If a subject received a loan. This is time-variant.
- elapsed: The number of days since receiving a loan at rd 3 interview date. This is time-invariant as it is computed only at rd 3. This defines the eventual treatment dose and should be our main covariate.

Following results are obtained.

Figure 1 This plots the mean of "3 meals per day" in each round. Left panel is control vs. treated in receivedCredit (actual assignmemt). Right panel is control vs. treated in assignment (original assignment). The question is changed in rd 2 onwards so direct comparison across rd 1 and 2,3 are not tenable. However, one sees a rising "3 meals per day", but almost paralel trend between rd 2 and 3 (which are comparable).

Table 1 This is a first-difference (linear probability) estimation result of "3 meals per day" on disbursed, arm, their interactions, assignment, elapsed days, using the post treatment data of rds 2, 3. It shows positive impacts of receiving a loan (disbursed under the covariate name "credit") aafter controlling for arm. control/treated are relative to "lost to flood" or "rejections," so it is not surprising to have better food intake.

Figure 2 Livestock is the main stated usage of loans.

Figure 3 Work hours seem to get longer.

Figure 4 New loans increased in rd 2, whose recall period corresponds to the timing disbursement.

Figure 5 Asset holding by receivedCredit= T/F and rds = 1,2,3. Asset holding is computed with rd 1 asset holding, asset addition in rd 2, 3, while assuming an annual rate of 5% depreciation. receivedCredit not randomised allocation as loan receipt must be agreed by subjects. It shows that the loan receivers have higher mean asset in rd 3, but not in rd 1 or 2, where the latter is good. Red dotted lines are medians, blue dotted lines are means.

Figure 6 Asset holding by disbursed=T/F and rds=1,2,3. This is also an endogenous switch. The basic picture is the same as Figure 5.

Figure 7 Asset holding by elapsed days grouped into "early receivers" and "late receivers" according to median elapsed day. This is (roughly) a randomised switch. It shows increasing asset levels, but no mean or median difference between early and late receivers.

Figure 8 Asset holding by elapsed days and arms. Not much to see here.

Figure 9 Difference in group-average asset holding between the treated and the control in assignment by difference in elapsed and arms. So it is group differences among original treatment assignment (treated - control) within the same cluster. It controls for cluster FE, and dose and outcome differences are taken between randomly assigned treatment status. This should be one of our main comparisons. (This is not DID: If I plot the first-difference version of the plots between rds, it will be double difference estimates.) When I draw loess curves, I see no trend over various elapsed day ("treatment dose") differences. It hints a zero gradient of dose levels.

Figure 10 Same identification idea as Figure 9 on livestock values. Again, we do not see markedly strong impacts, but we see some differences in terms of dispersion. cow arm has smaller variations around the loess curves in rds 2 and 3 relative to the large grace arm. traditional arm has a simlar pattern as large grace.

Figure 11 Livestock holding by elapsed days. Substantial heterogeneity in rd 3 but no statistical significant changes.

Figure 13 Total asset holding by elapsed days. Substantial heterogeneity in rd 3 but no statistical significant changes.

Table 3 DID estimation of total asset holding by elapsed days. Zero impact.

Please ignore the texts hereafter as they are my memos in programming.

# I Read

List folder names and read files.

# II Treatment through time

Treatment assignment file. There are 220 cases of attrition who are group rejections (140) and lost to flood (80).

tr0 and tr1 are based on the information at rd 3.

Merge interview dates with treatment assignment info tr1.

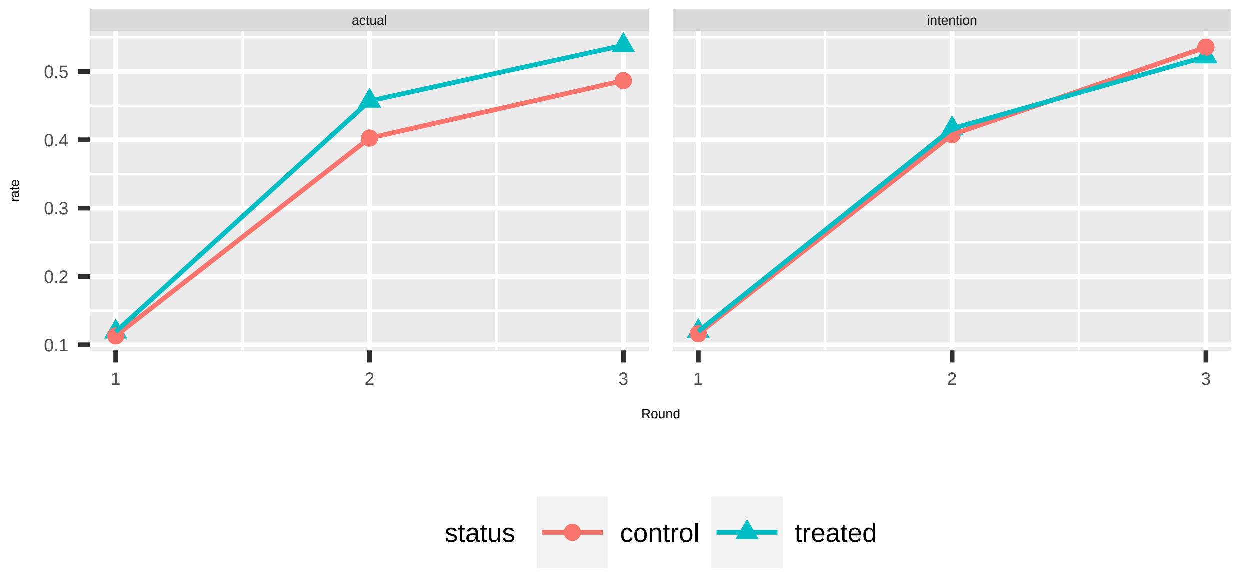


Figure 1 3 meals per day

## III Food consumption (23B in rd 1, 3B in rds 2, 3)

There is pure duplication in rd 1 files. Drop them.

There is only one HH that reports food intake of a non-head member. I will drop this non-head member.

Drop hhid = NA.

Merge treatment assignment info.

Some gids are missing in sec3b. Check if the merge is done correctly. Check if this is due to hhid = 980... cases. Strip leading 980/990 and see if the matched observations have variables originally from tr0.

Rd 1 seems to be merged OK. Rd 2, 3 show that there are duplicated hhid so drop all entries with duplication.

There still remains unmatched observations as seen in NAs in assignment (found in Sec 3B files but not in identification files.) We drop these observations.

Three meals. In rd 1, we ask for all the members about the number of times they eat meals, during monga and off-monga seasons. On average, there is only 1 out of 1 HH members reponding to the question, which are all HH head members. In rds 2 and 3, we ask a blanket question if all the members eat three times a day for the whole year. So rd 1 question is more likely to be responded as “3 times” than in rd 2, 3 questions, *cetris paribus*. So observing more “3 times” responses in the latter rds indicate that there may be improvements in household food intake.

Combine rd 1 original and additional into a single file, then put into a list with rds 2, 3.

Given the questions are different, it is not surprising that we have different proportion of subjects with three meals per day. Despite this limitation, we have an increasing food consumption security which is promising.

Form data for regression.

3 meals per day in regular times for rd 1. For rd 2, 3, yes to the question.

Rescale days by 100. Note that assignment has empty observations who either group rejected or lost to flood. They form the reference group for assignment (control, treated).

m3data: Rd 2-3 data on three meals per day.

Table 1: FD estimates of three meals per day, round 2, 3

| rn | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| (Intercept) | 0.118* | −0.060 | 0.081 | −0.060 | 0.141** | 0.051 |
| | (0.062) | (0.111) | (0.066) | (0.111) | (0.055) | (0.055) |
| large | −0.014 | | 0.067 | | −0.008 | 0.124 |
| | (0.099) | | (0.096) | | (0.097) | (0.086) |
| large grace | 0.093 | | 0.162 | | 0.087 | 0.197** |
| | (0.090) | | (0.098) | | (0.084) | (0.091) |
| cow | −0.079 | | −0.051 | | −0.069 | −0.009 |
| | (0.101) | | (0.101) | | (0.098) | (0.098) |
| lost to flood | −0.043 | | −0.005 | | | |
| | (0.133) | | (0.134) | | | |
| control | | 0.211* | | 0.236** | | |
| | | (0.117) | | (0.119) | | |
| treated | | 0.176 | | 0.178 | | |
| | | (0.117) | | (0.141) | | |
| credit | | | 0.147** | −0.062 | | 0.142** |
| | | | (0.064) | (0.062) | | (0.065) |
| large * credit | | | −0.367*** | | | −0.423*** |
| | | | (0.126) | | | (0.125) |
| large grace * credit | | | −0.230** | | | −0.265*** |
| | | | (0.090) | | | (0.091) |
| cow * credit | | | −0.108 | | | −0.122 |
| | | | (0.134) | | | (0.136) |
| treated * credit | | | | −0.006 | | |
| | | | | (0.086) | | |
| elapsed days * 100 | | | | | −0.001 | 0.006 |
| | | | | | (0.006) | (0.009) |
| $R^2$ | 0.01 | 0.009 | 0.022 | 0.011 | 0.009 | 0.029 |
| $n$ | 2043 | 2043 | 1838 | 1838 | 1657 | 1527 |

Notes: 1. First-difference estimates of having three meals per day using rd 2 and 3 information. Standard errors are clustered at the group level.

2. large, large grace, cow, lost to flood, control, treated, credit are all time invariant and are interacted with a trend term. Regressions (1) - (4) include subjects who group-rejected or lost to flood as a reference group. Regressions (5) - (6) drop subjects who group-rejected or lost to flood and use the subjects who were initially assigned to the control as a reference group.

3. $*$, $**$, $***$ indicate significance levels at 10%, 5%, 1%, respectively.

We see no impacts of intervention when comparing two perids after the disbursement.

# IV Credit use

File names of rd 3 files are named for the page ordering. For example, ./2/section_21a.prn are the first 2 questions of Section 20, which is named as 21 as it is an unnumbered page that comes right after Section 20. ./2/section_22a.prn is Section 18.

In rd 2, Section 21 is stored under ./2/section_23.prn, in rd 3, ./3/section_21_use_of_credit_1.prn, ./3/section_21_use_of_credit_2.prn.

Merge rd 2 and 3.

Intended use of credit, mostly livestock (cows). It is interesting to note that the majority of our subjects choose livestock for an investment.

Work hours.

# V New loans

Loans in rds 1, 2, and 3.

```
Warning in `[.data.table`(bo1l, , `:=`(totalSum, sum(cashAmount, na.rm = T)), : Invalid .:
```

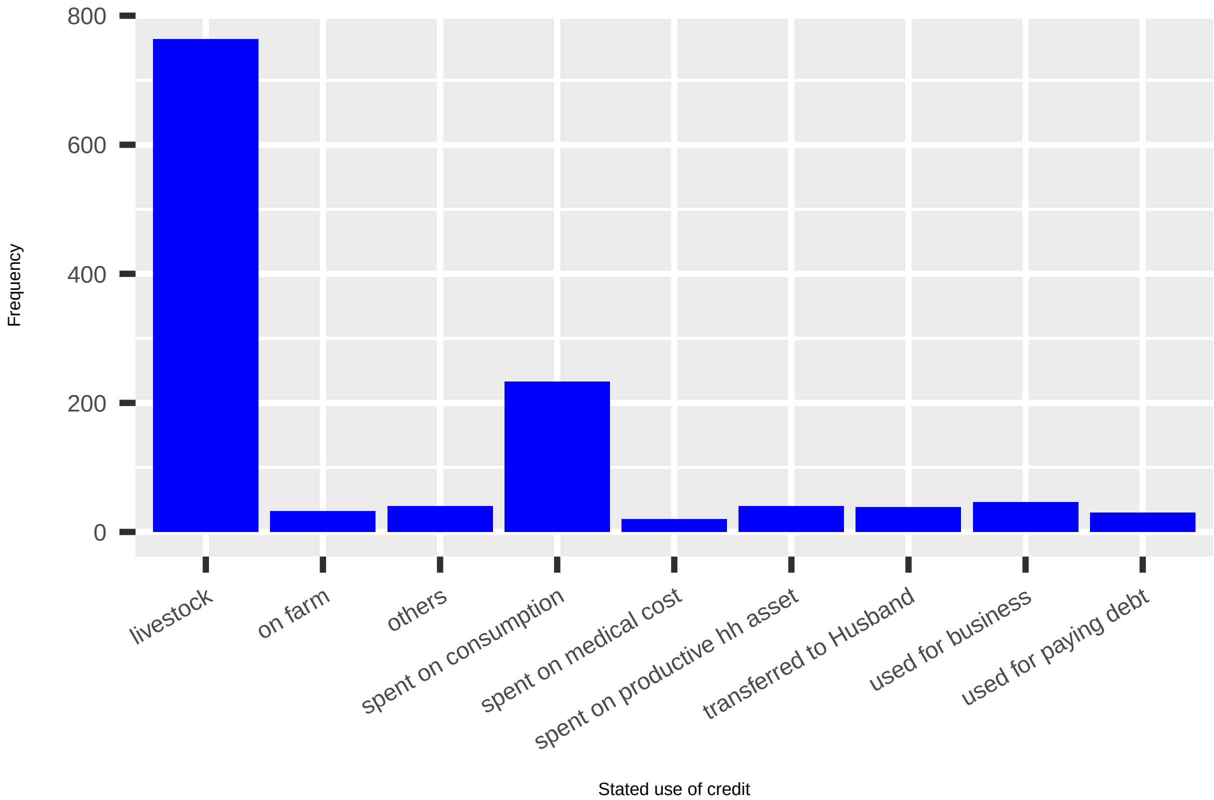


Figure 2 Stated use of credit in rd 2

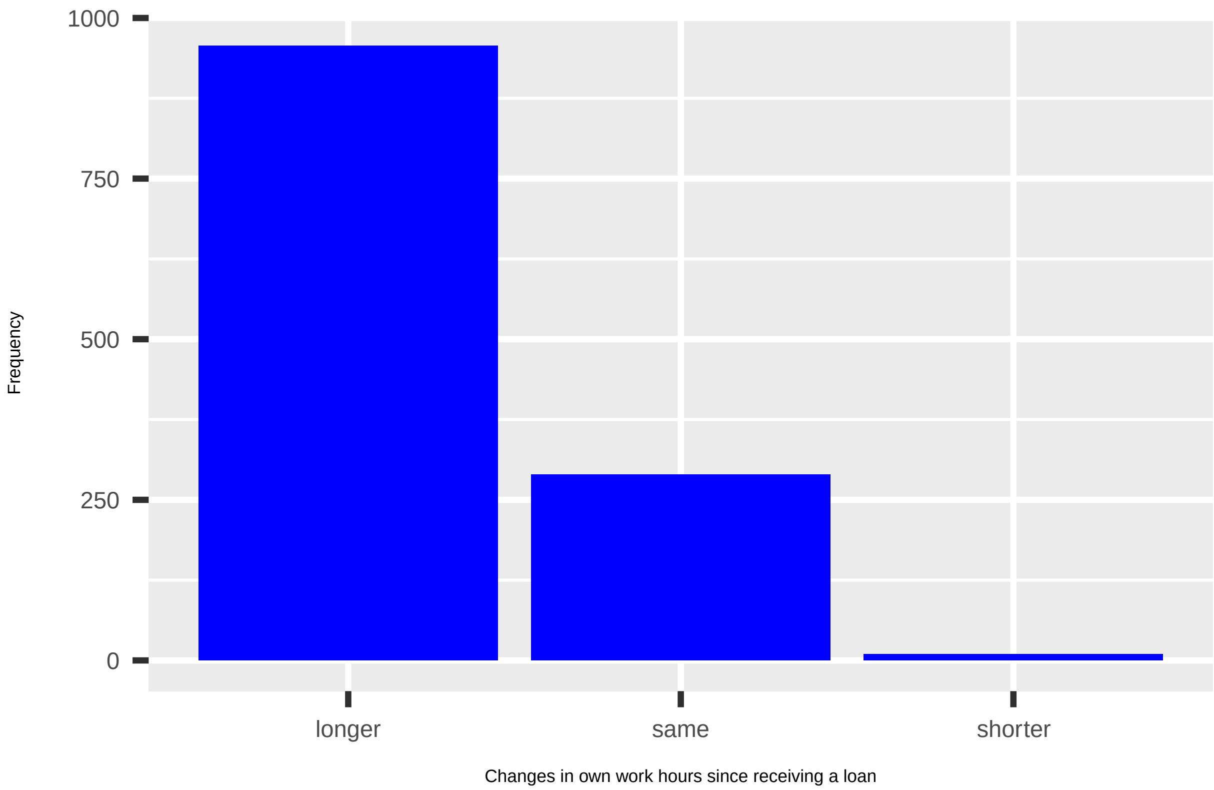


Figure 3 Work hours

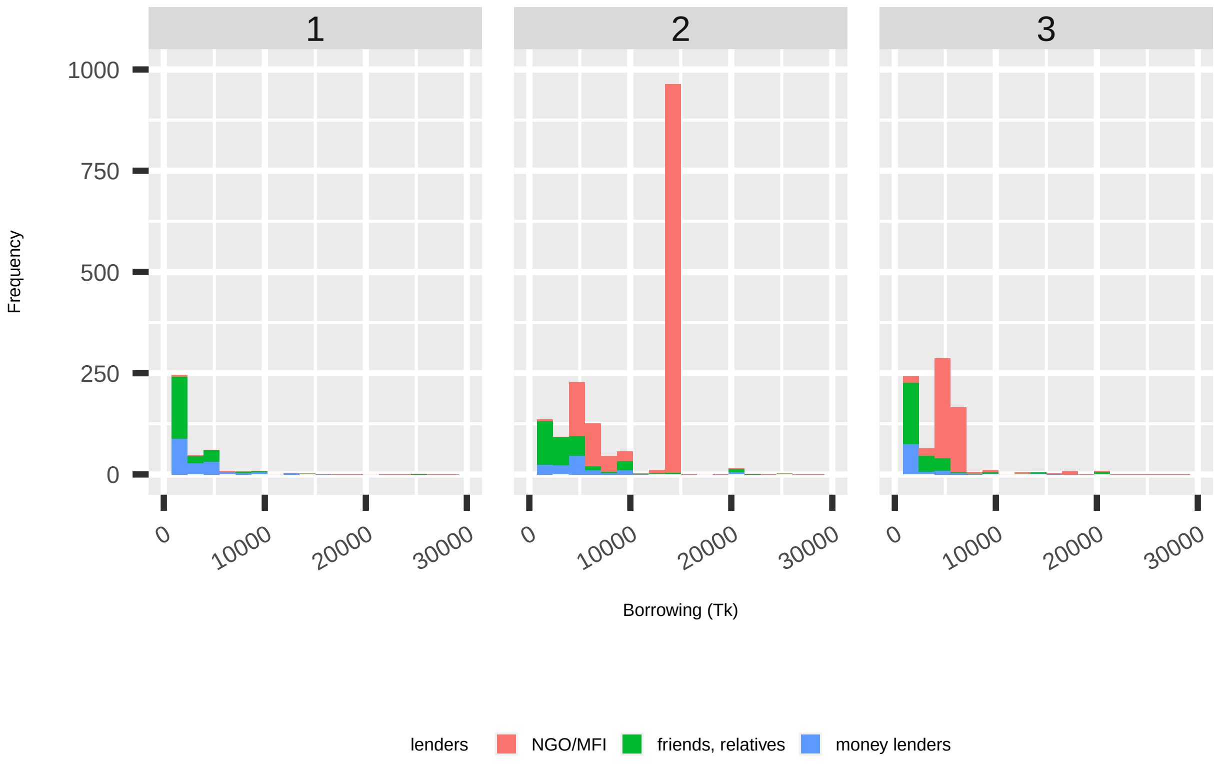


Figure 4  New loans

In rd 1, there are only 14 subjects who have borrowed from other NGO/MFI in the last 12 months. Most of the loans are taken from friends, relatives and money lenders, for about 9%, 13% of subjects, respectively.

Append rd 1.

Merge treatment info.

Plot new loans in each rd. I will combine shop owners/traders with money lenders. I will also combine GUK and other NGO/MFI to NGO/MFI. We also omit zero borrowing from the histogram for clarity.  One can see that, in rd 1, there is virtually no borrowing from NGO/MFI among our subjects. This indicates that our study areas are relatively free from other non-indiginous financial intermediaries which allows us to estimate the impacts of our loans without much concerns of treatment contamination. In rd 2, borrowing from NGO/MFI increased rapidly as a result of our intervention. In rds 2 and 3, some individuals report smaller amount, which correspond to our traditional loan arm. It is hard to say that the loans from friends, relatives or money lenders have decreased after our intervetion between rd 1 and rd 2.

# VI Assets

Read files.

Separate into rds for rd-specfic operations (to be merged back later).

Rd 1.

Rd 2.

Rd 3.

Bind all 3 rds together.

Drop rd 2 and 3 assets that were not bought in the lastYear to avoid double counting.  The

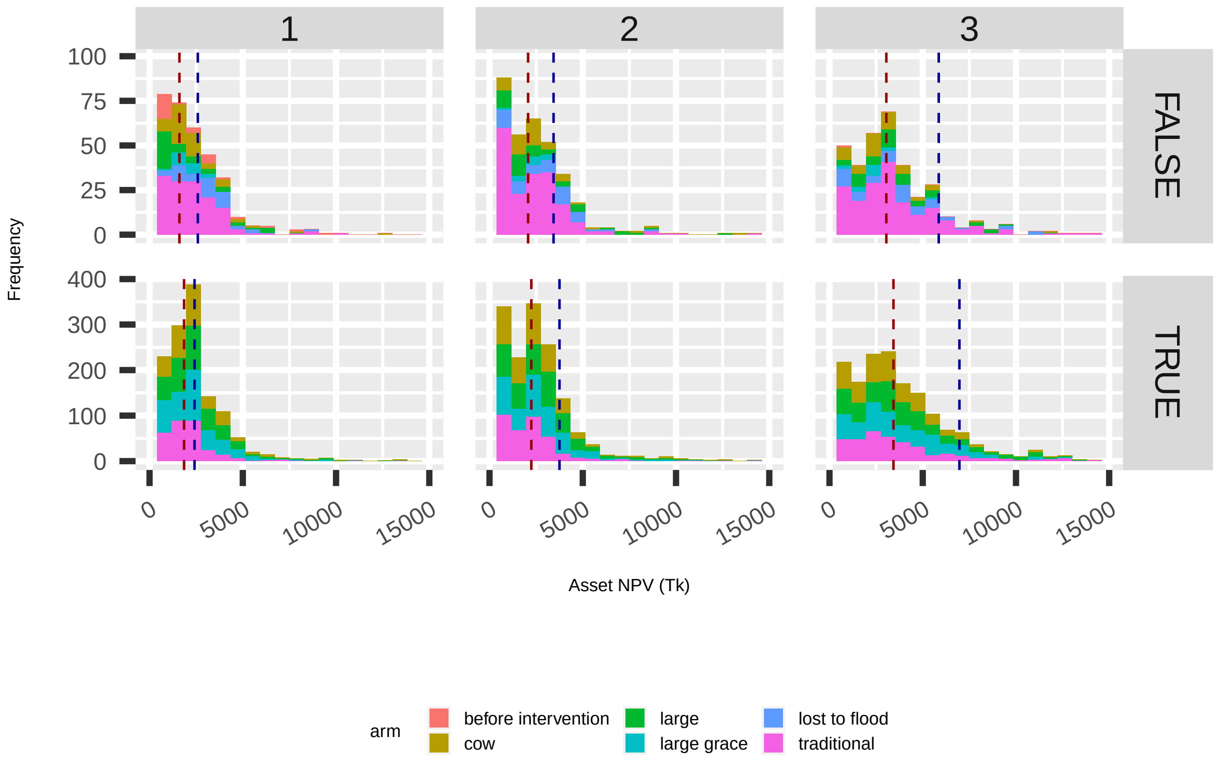


Figure 5  Assets by eventual treatmemt status

histogram is created by imputing the NPV of household assets by assuming an annual 5% depreciation rate. We see that, at rd 1, there is no difference in mean of asset holding, while the medians are different. Interestingly, the median difference is preserved in the later rounds. In the meantime, means are not different in rd 1 yet they come to differ in later rounds. The subjects who actually received credits have higher mean asset holding. Given that the median diffrences are unchanged, this indicates that the upper half of the treated asset holders are getting better than the control.

To align dates of receiving credits for the subjects who did not, we use the median daysFromStart.

In this figure, we dropped observations without gid and disbursed. When intDate is NA (not interviewed), we cannot define disbursement for that round. We know disbursement took place before rd 3, so all assignment = treated have disbursed = T in rd 3.

In this scatter and loess plots, we put asset values against the treatment exposure, facetted by treatment arms. This aims to mimic ATT under a continuous treatment. The treatment exposure is defined by the elapsed days since receiving a credit. Since the treatment exposure is randomised, this is a statisically valid procedure to observe the treatment response without major comfounding.

This plotting exercise leads one to consider the statisical model underlying the graphs. For an individual $i$'s outcome $y_i$, the treatment assignment $D_i = 0, 1$ may have an impact on the outcome. The standard Rubin causal model deals with a binary indicator variable for $D_i$. In our design, we vary the dates of intervention among the subjects. So what we randomly vary is the duration under treatment, or dose exposure, denoted with $D_i(t)$ where $t$ is the calendar date of intervention. On average, there is about 1 year difference in $t$ within a cluster of 20 subjects. Given that we randomise the calendar dates of starting the intevention, we can assume actual duration $t \in [t_0, t_1]$ is orthogonal to potential treatment response $y(t)$ for all $t$. Under the simplest setting, we follow Imbens (2000); Hirano and Imbens (2005); Imai and van Dyk (2004); Egger and von Ehrlich (2013) assume the following conditional orthogonality in the continuous case. Denoting $T$ as a random variable with

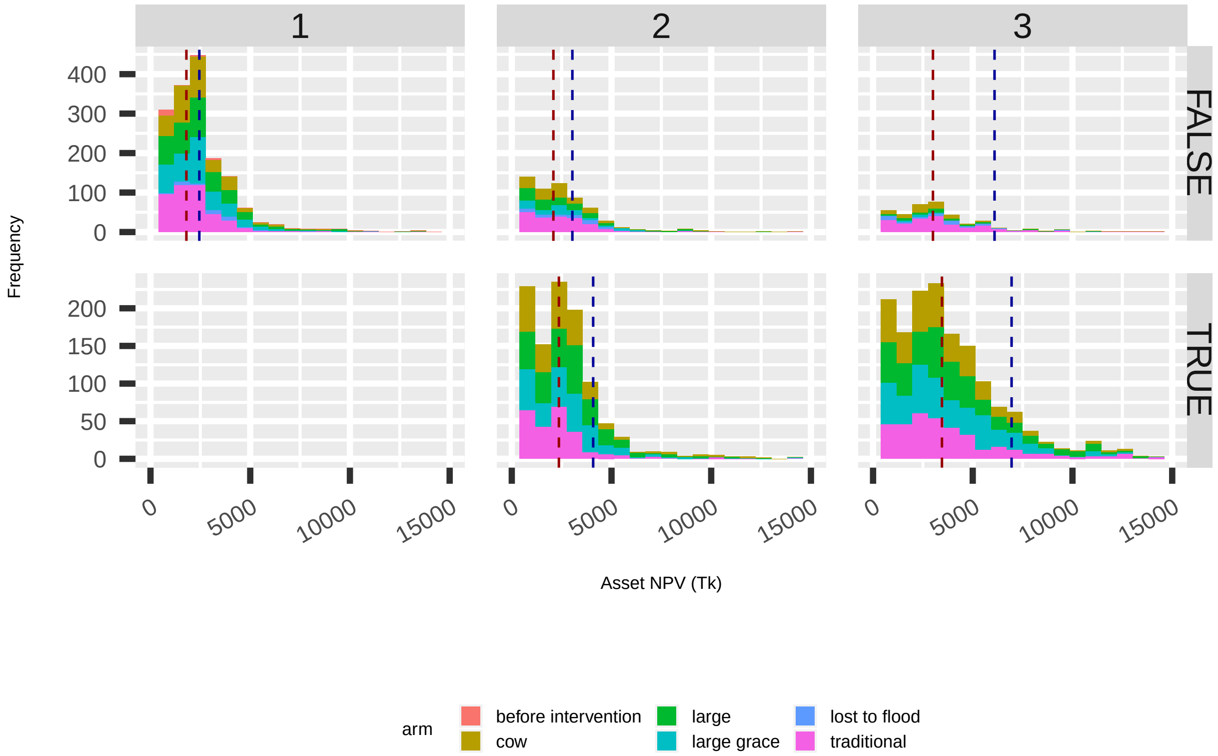


Figure 6 Assets by disbursement status

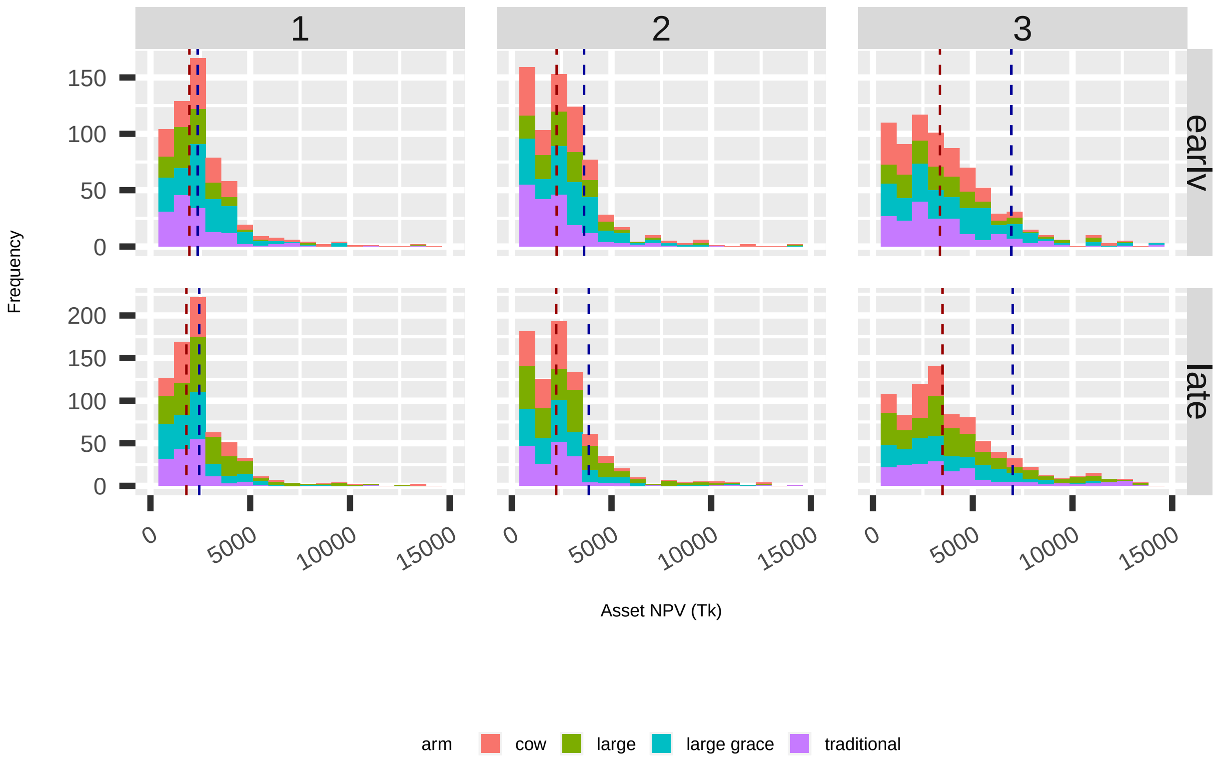


Figure 7 Assets by elapsed days from disbursement, ATT

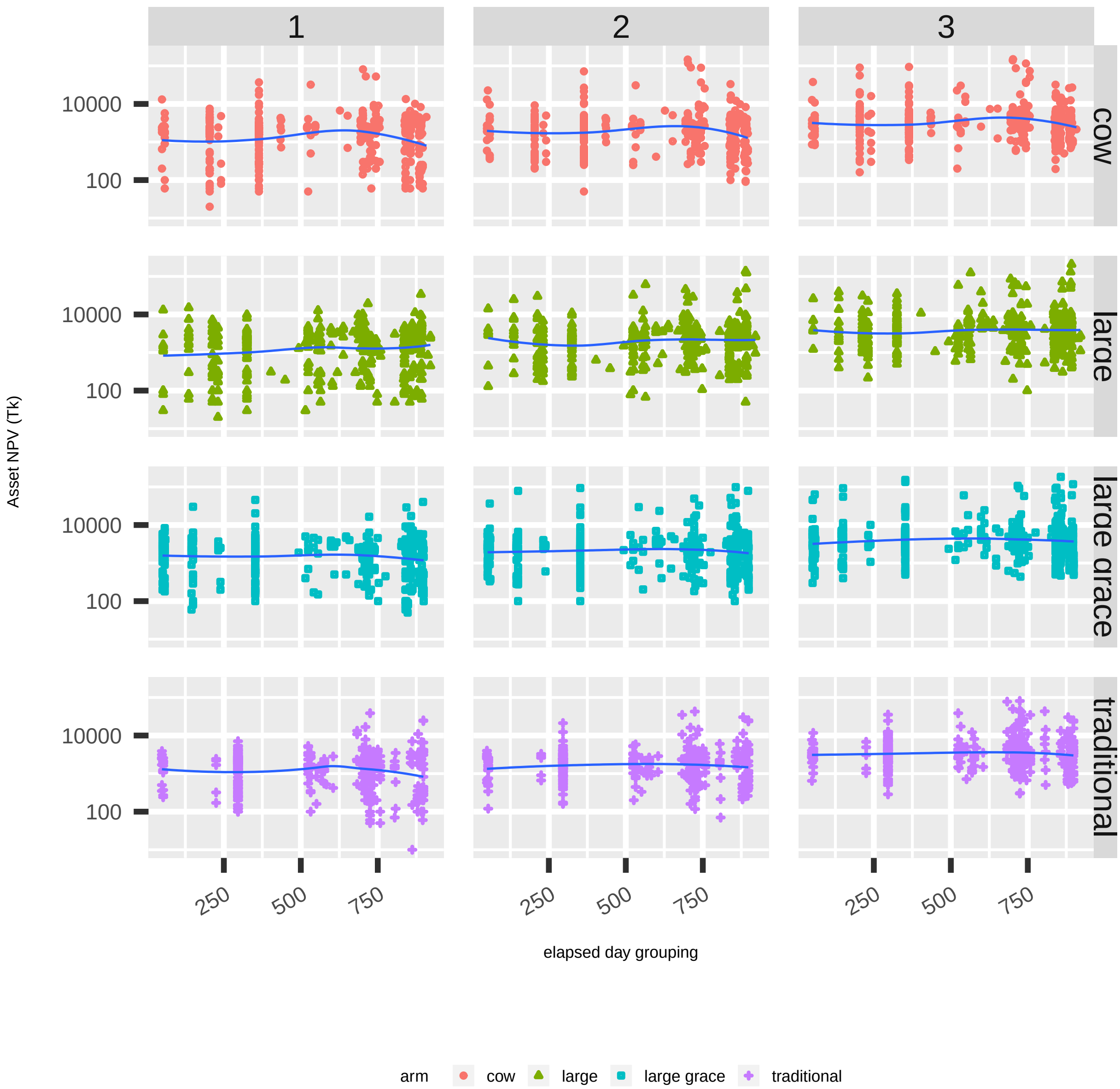


Figure 8 Assets by elapsed days from disbursement, ATT scatter plot

its realisation written as $t$, we assume:

$$y(t) \perp T|\mathbf{x}.$$

Hirano and Imbens (2005) shows that this is equivalent to

$$\mathbf{x} \perp 1\{T = t\}|g(t, \mathbf{x})$$

where $g(t, \mathbf{x})$ is a generalised propesity score that gives the density of treatment at $t$ given $\mathbf{x}$. This shows that one can estimate continuous treatment effect by first, estimating GPS $g$, second, estimate the conditional expectation of outcome as a function of $g$ and $\mathbf{x}$:

$$\beta(t, g) = \mathcal{E}[y|T = t, G = g(t, \mathbf{x})],$$

and then average over $g$ for a given $t$ to obtain the dose-response function

$$\beta(t) = \mathcal{E}[\beta(t, g)|\mathbf{x}].$$

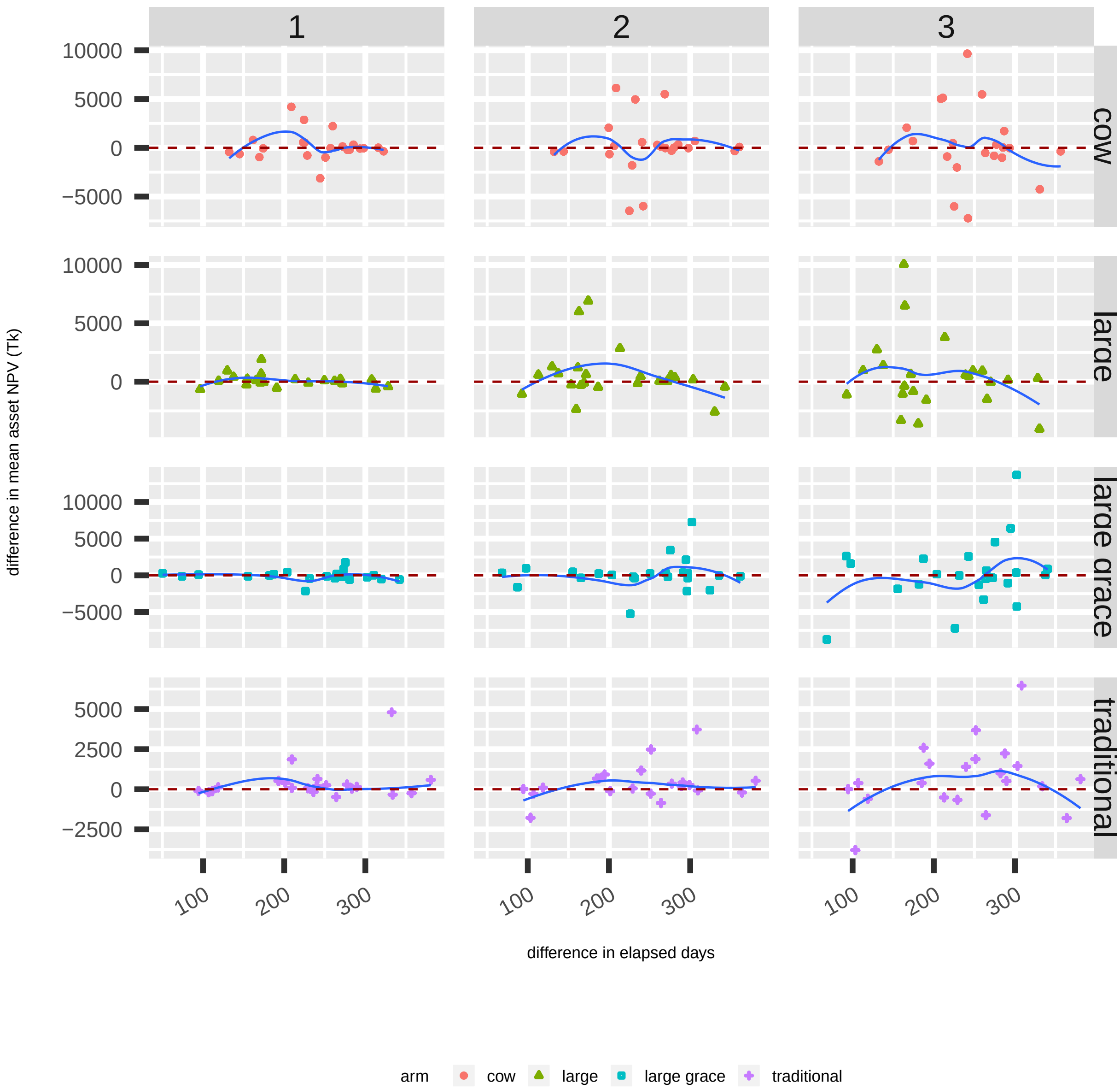


Figure 9 Assets by elapsed days from disbursement, within a group

The approach is preceded by applied works related job training duration (Kluve et al., 2012).

We compare the effects of treatment exposure differences within the same group.

# VII livestock

# VIII livestock

Rd 1.

Price: female calf, male calf, ox.

Let the price to be used as median price, and cow price is 15000. Lease share is 50%.

Goats: Take prices from late rounds. 1900.

Total livestock value.

Rd2.

Rd3.
Merge.
We compare the effects of treatment exposure differences within the same group.

```
asset ← Asset01t[assetNumber == 1, ]
asset.ss ← subset(asset, assetNPV > 0 & !is.na(gid) & !is.na(elapsed))
setkey(asset.ss, rd, gid, asssignment)
```

```
Error in setkeyv(x, cols, verbose = verbose, physical = physical): some columns are not in
```

```
asset.ss[, avgElapsed := mean(elapsed, na.rm = T), by = c("rd", "gid", "assignment")]
asset.ss[, avgElapsed0 := avgElapsed[1], by = c("rd", "gid")]
asset.ss[, avgElapsed1 := avgElapsed[.N], by = c("rd", "gid")]
asset.ss[gid == 70650 & grepl("co", assignment), ]
asset.ss[gid == 70204& rd == 1, .(rd, gid, assignment, avgElapsed, avgElapsed0, avgElapsed
asset.ss[, avgNPV := mean(assetNPV/1000, na.rm = T),
        by = c("rd", "gid", "assignment")]
asset.ss[, avgNPV0 := avgNPV[1], by = c("rd", "gid")]
asset.ss[, avgNPV1 := avgNPV[.N], by = c("rd", "gid")]
asset.ss[, avgDiffElapsed := avgElapsed1 - avgElapsed0]
asset.ss[, avgDiffNPV := avgNPV1 - avgNPV0]
setkey(asset.ss, rd, gid, assignment)
dim(asset.sss ← asset.ss[!duplicated(asset.ss[, .(rd, gid, assignment)]), ])
library(ggplot2)
ggplot(data = asset.sss, aes(x = avgDiffElapsed, y = avgDiffNPV)) +
        geom_point(aes(colour = arm, shape = arm), size = .05) +
        scale_shape(solid = F) +
        xlab("difference in elapsed days") + ylab("difference in mean asset NPV (Tk '000)"
        labs(fill = "arm") + facet_grid( ~ rd) +
#       stat_smooth(method = "loess", size = .2, n = 150) +
        geom_smooth(method = "loess", size = .2) +
        geom_hline(aes(yintercept = 0), colour="#990000", linetype="dashed", size = .2) +
        theme(axis.title.y = element_text(size = rel(.25), angle = 90),
                axis.title.x = element_text(size = rel(.25), angle = 0),
                axis.text.x = element_text(size = rel(.5), angle = 30, hjust = 1),
                axis.text.y = element_text(size = rel(.5), angle = 0),
                legend.text = element_text(size=rel(.25)),
                legend.position = "bottom",
                legend.title = element_text(size = rel(.25)),
                legend.key = element_rect(size = rel(.25)),
                legend.key.size = unit(.15, "cm"),
                strip.text = element_text(size=rel(.5)),
                strip.text.x = element_text(margin = margin(.05, 0, .05, 0, "cm")),
                strip.text.y = element_text(margin = margin(.05, 0, .05, 0, "cm")))
```

```
library(ggplot2)
ggplot(data = asset.sss, aes(x = avgDiffElapsed, y = avgDiffNPV)) +
        geom_point(aes(colour = arm, shape = arm), size = .05) +
        scale_shape(solid = F) +
        xlab("difference in elapsed days") + ylab("difference in mean asset NPV (Tk '000)"
        labs(fill = "arm") + facet_grid(arm ~ rd, scale = "free_y") +
#       stat_smooth(method = "loess", size = .2, n = 150) +
        geom_smooth(method = "loess", size = .2) +
        geom_hline(aes(yintercept = 0), colour="#990000", linetype="dashed", size = .2) +
        theme(axis.title.y = element_text(size = rel(.25), angle = 90),
                axis.title.x = element_text(size = rel(.25), angle = 0),
```

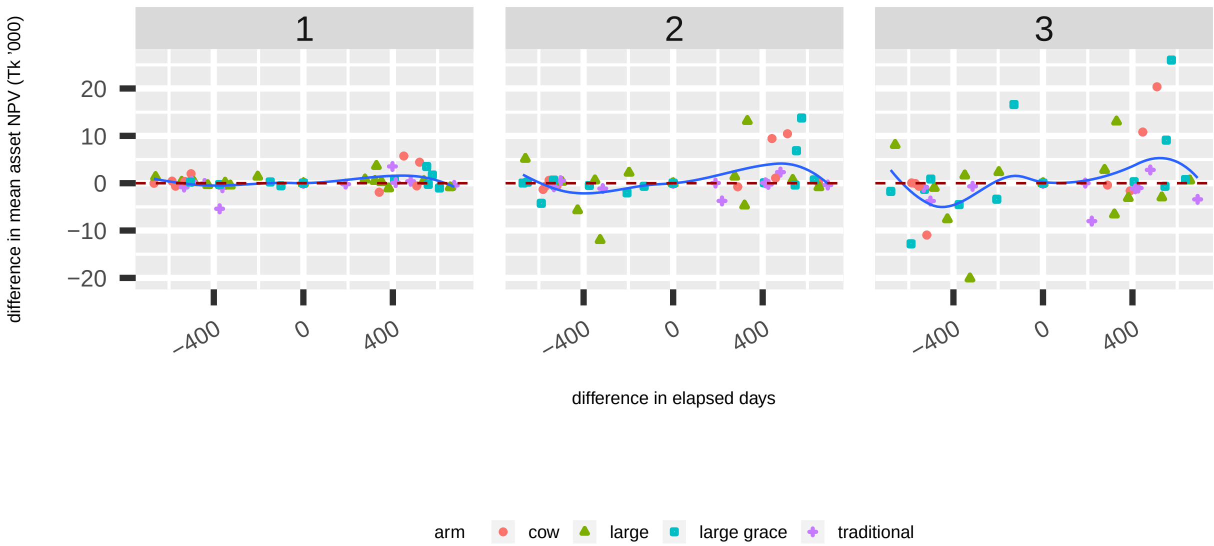


Figure 10 Assets by elapsed days from disbursement within a group

```
                    axis.text.x = element_text(size = rel(.5), angle = 30, hjust = 1),
                    axis.text.y = element_text(size = rel(.5), angle = 0),
                    legend.text = element_text(size=rel(.25)),
                    legend.position = "bottom",
                    legend.title = element_text(size = rel(.25)),
                    legend.key = element_rect(size = rel(.25)),
                    legend.key.size = unit(.15, "cm"),
                    strip.text = element_text(size=rel(.5)),
                    strip.text.x = element_text(margin = margin(.05, 0, .05, 0, "cm")),
                    strip.text.y = element_text(margin = margin(.05, 0, .05, 0, "cm")))
```

# IX livestock

Rd 1.
Price: female calf, male calf, ox.
Let the price to be used as median price, and cow price is 15000. Lease share is 50%.
Goats: Take prices from late rounds. 1900.
Total livestock value.
Rd2.
Rd3.
Merge.
We compare the effects of treatment exposure differences within the same group.

```
library(ggplot2)
ggplot(data = lstk.sss, aes(x = avgDiffElapsed, y = avgDiffLstkValue)) +
          geom_point(aes(colour = arm, shape = arm), size = .05) +
          scale_shape(solid = F) + scale_y_continuous() +
          xlab("difference in elapsed days") + ylab("difference in mean livestock value (Tk
          labs(fill = "arm") + facet_grid( ~ rd) +
#         stat_smooth(method = "loess", size = .2, n = 150) +
          geom_smooth(method = "loess", size = .2) +
          geom_hline(aes(yintercept = 0), colour="#990000", linetype="dashed", size = .2) +
```

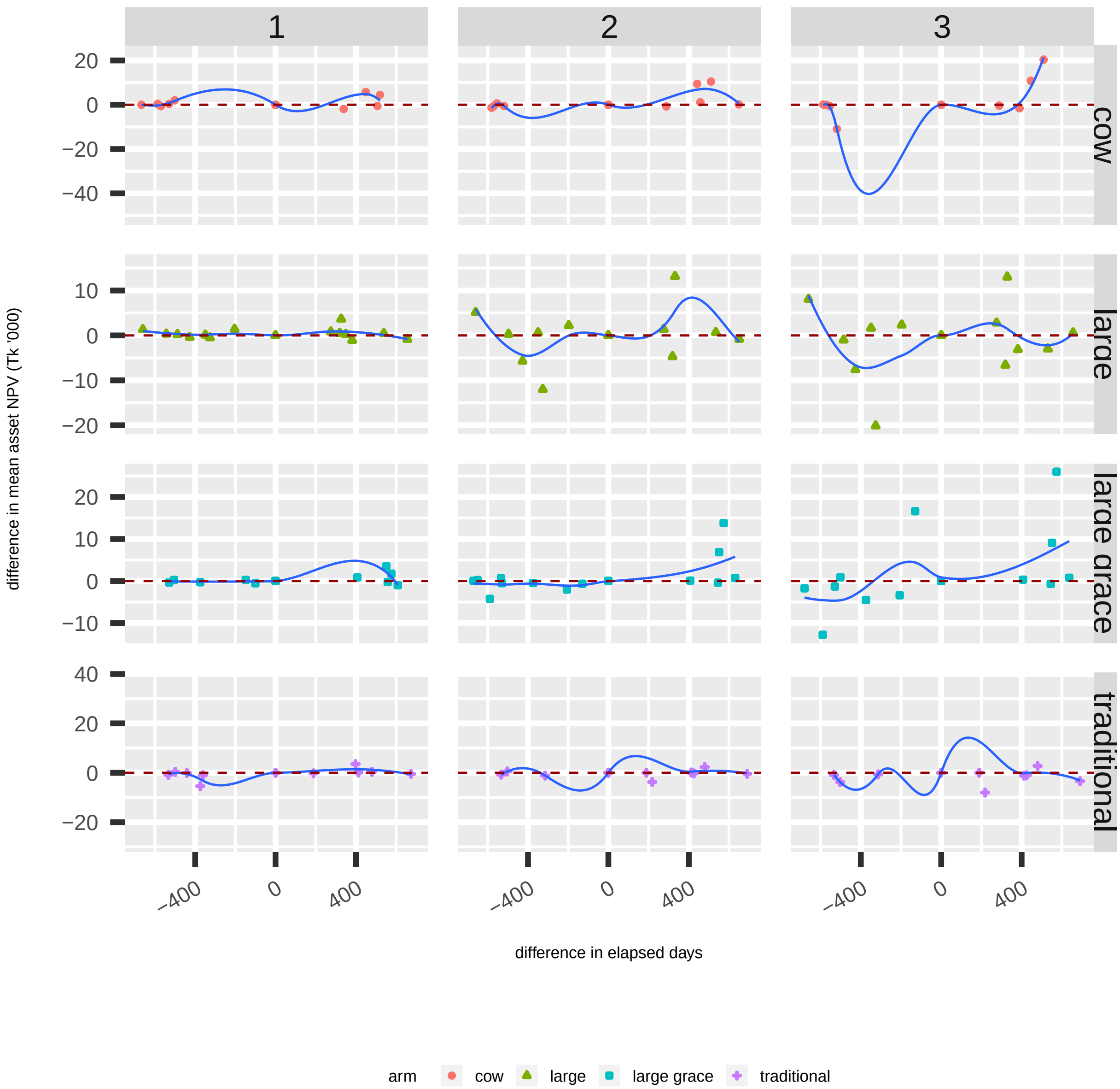


Figure 11 Assets by elapsed days from disbursement, within a group

```
        theme(axis.title.y = element_text(size = rel(.25), angle = 90),
                axis.title.x = element_text(size = rel(.25), angle = 0),
                axis.text.x = element_text(size = rel(.5), angle = 30, hjust = 1),
                axis.text.y = element_text(size = rel(.5), angle = 0),
                legend.text = element_text(size=rel(.25)),
                legend.position = "bottom",
                legend.title = element_text(size = rel(.25)),
                legend.key = element_rect(size = rel(.25)),
                legend.key.size = unit(.15, "cm"),
                strip.text = element_text(size=rel(.5)),
                strip.text.x = element_text(margin = margin(.05, 0, .05, 0, "cm")),
                strip.text.y = element_text(margin = margin(.05, 0, .05, 0, "cm")))
```

```
library(ggplot2)
ggplot(data = lstk.sss, aes(x = avgDiffElapsed, y = avgDiffLstkValue)) +
        geom_point(aes(colour = arm, shape = arm), size = .05) +
```

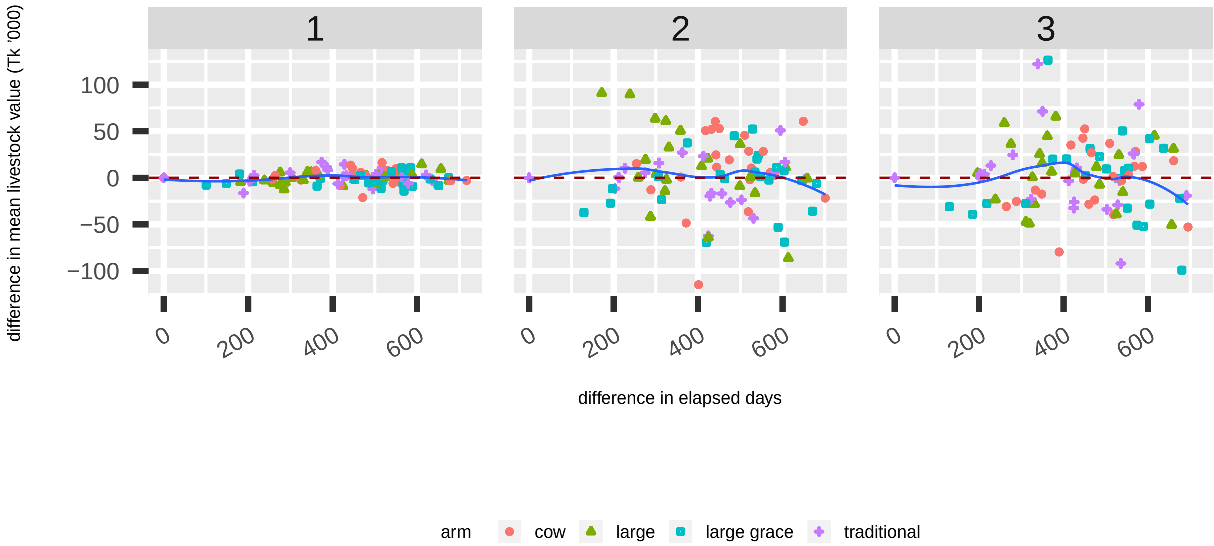


Figure 12 Livestock by elapsed days from disbursement within a group

```
        scale_shape(solid = F) +
        xlab("difference in elapsed days") + ylab("difference in mean livestock value (Tk
        labs(fill = "arm") + facet_grid(arm ~ rd) +
#       stat_smooth(method = "loess", size = .2, n = 150) +
        geom_smooth(method = "loess", size = .2) +
        geom_hline(aes(yintercept = 0), colour="#990000", linetype="dashed", size = .2) +
        theme(axis.title.y = element_text(size = rel(.25), angle = 90),
                axis.title.x = element_text(size = rel(.25), angle = 0),
                axis.text.x = element_text(size = rel(.5), angle = 30, hjust = 1),
                axis.text.y = element_text(size = rel(.5), angle = 0),
                legend.text = element_text(size=rel(.25)),
                legend.position = "bottom",
                legend.title = element_text(size = rel(.25)),
                legend.key = element_rect(size = rel(.25)),
                legend.key.size = unit(.15, "cm"),
                strip.text = element_text(size=rel(.5)),
                strip.text.x = element_text(margin = margin(.05, 0, .05, 0, "cm")),
                strip.text.y = element_text(margin = margin(.05, 0, .05, 0, "cm")))
```

Add assets and livestock.

```
library(ggplot2)
ggplot(data = al.sss, aes(x = avgDiffElapsed, y = avgDiffVal)) +
        geom_point(aes(colour = arm, shape = arm), size = .05) +
        scale_shape(solid = F) +
        xlab("difference in elapsed days") + ylab("difference in mean value (Tk '000)") +
        labs(fill = "arm") + facet_grid(. ~ rd) +
#       stat_smooth(method = "loess", size = .2, n = 150) +
        geom_smooth(method = "loess", size = .2) +
        geom_hline(aes(yintercept = 0), colour="#990000", linetype="dashed", size = .2) +
        theme(axis.title.y = element_text(size = rel(.25), angle = 90),
                axis.title.x = element_text(size = rel(.25), angle = 0),
                axis.text.x = element_text(size = rel(.5), angle = 30, hjust = 1),
                axis.text.y = element_text(size = rel(.5), angle = 0),
                legend.text = element_text(size=rel(.25)),
```

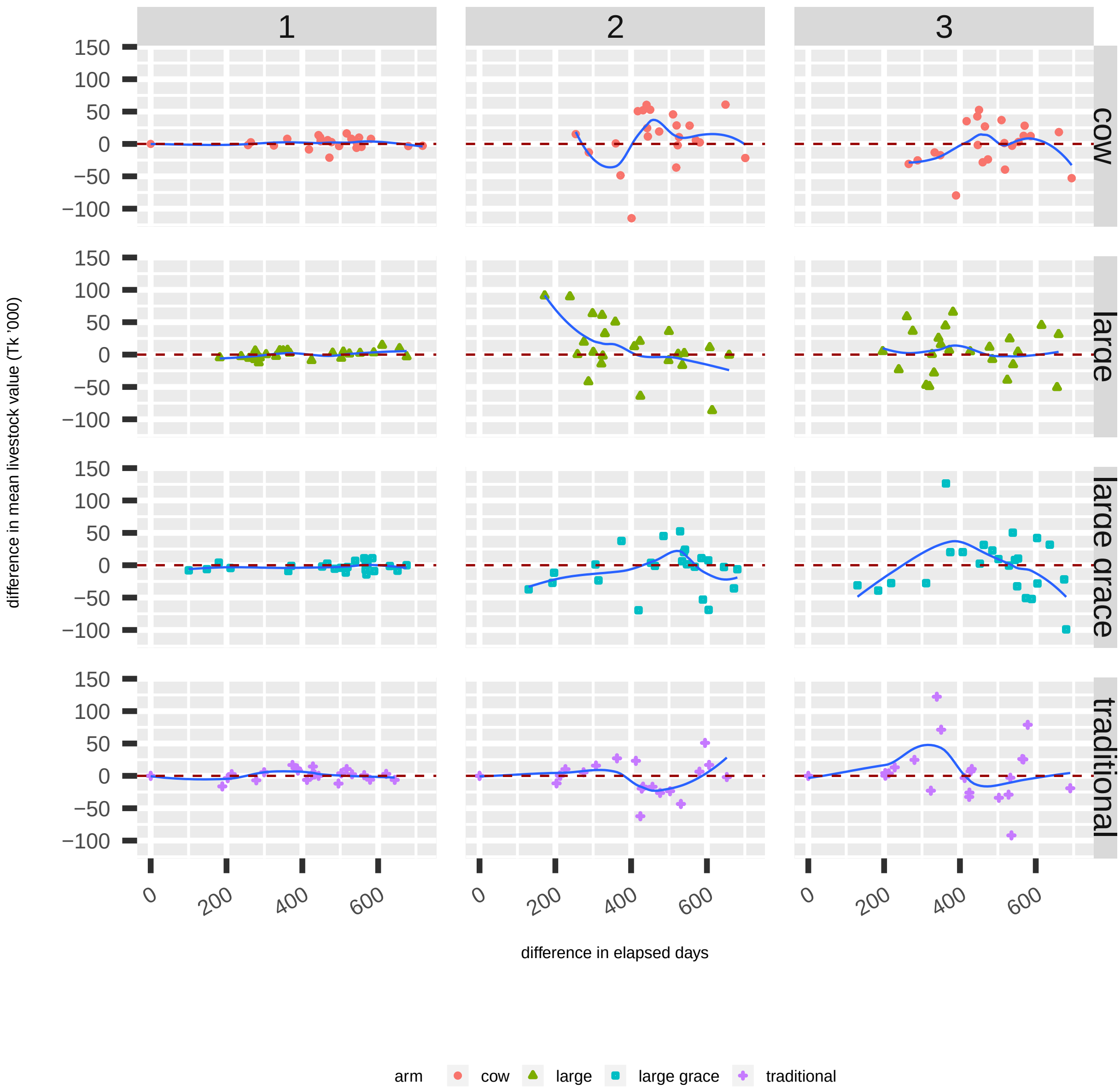


Figure 13 Livestock by elapsed days from disbursement, within a group

```
                 legend.position = "bottom",
                 legend.title = element_text(size = rel(.25)),
                 legend.key = element_rect(size = rel(.25)),
                 legend.key.size = unit(.15, "cm"),
                 strip.text = element_text(size=rel(.5)),
                 strip.text.x = element_text(margin = margin(.05, 0, .05, 0, "cm")),
                 strip.text.y = element_text(margin = margin(.05, 0, .05, 0, "cm")))

library(ggplot2)
ggplot(data = al.sss, aes(x = avgDiffElapsed, y = avgDiffVal)) +
        geom_point(aes(colour = arm, shape = arm), size = .05) +
        scale_shape(solid = F) +
        xlab("difference in elapsed days") + ylab("difference in mean value (Tk 1000)") +
        labs(fill = "arm") + facet_grid(arm ~ rd) +
#       stat_smooth(method = "loess", size = .2, n = 150) +
        geom_smooth(method = "loess", size = .2) +
```

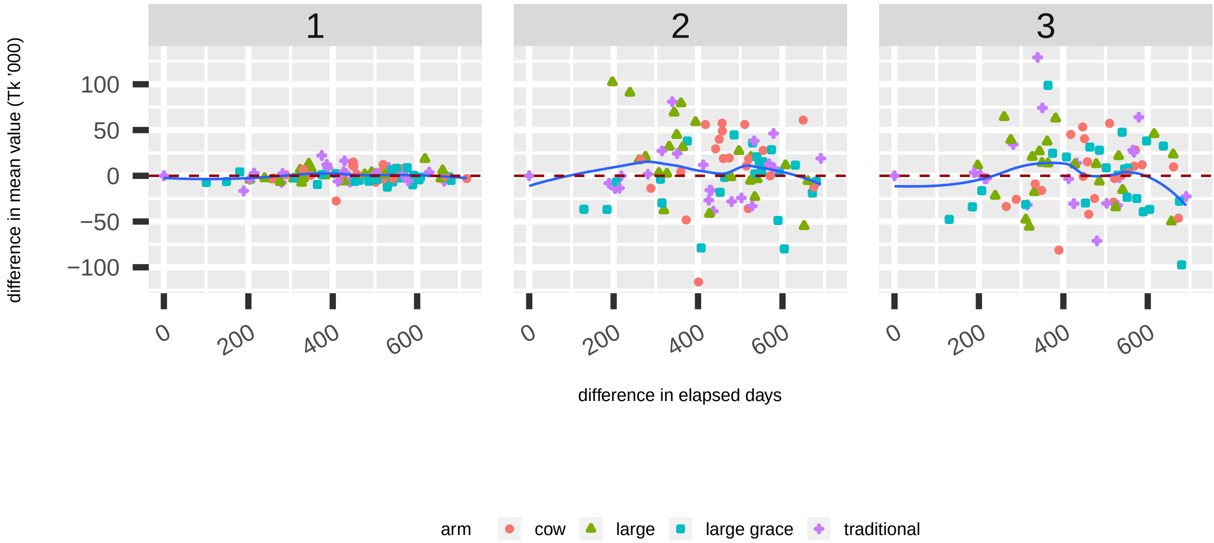


Figure 14 Total assets by elapsed days from disbursement within a group

```
        geom_hline(aes(yintercept = 0), colour="#990000", linetype="dashed", size = .2) +
        theme(axis.title.y = element_text(size = rel(.25), angle = 90),
                axis.title.x = element_text(size = rel(.25), angle = 0),
                axis.text.x = element_text(size = rel(.5), angle = 30, hjust = 1),
                axis.text.y = element_text(size = rel(.5), angle = 0),
                legend.text = element_text(size=rel(.25)),
                legend.position = "bottom",
                legend.title = element_text(size = rel(.25)),
                legend.key = element_rect(size = rel(.25)),
                legend.key.size = unit(.15, "cm"),
                strip.text = element_text(size=rel(.5)),
                strip.text.x = element_text(margin = margin(.05, 0, .05, 0, "cm")),
                strip.text.y = element_text(margin = margin(.05, 0, .05, 0, "cm")))
```

Regressions. First, get roster files to obtain hh background.

Summarise at cluster level.

Merge with asset data.

# References

**Egger, Peter H. and Maximilian von Ehrlich**, "Generalized propensity scores for multiple continuous treatment variables," *Economics Letters*, 2013, *119* (1), 32 – 34.

**Hirano, Keisuke and Guido W. Imbens**, *The Propensity Score with Continuous Treatments*, John Wiley & Sons, Ltd,

**Imai, Kosuke and David A van Dyk**, "Causal Inference With General Treatment Regimes," *Journal of the American Statistical Association*, 2004, *99* (467), 854–866.

**Imbens, Guido W.**, "The role of the propensity score in estimating dose-response functions," *Biometrika*, 2000, *87* (3), 706.

**Kluve, Jochen, Hilmar Schneider, Arne Uhlendorff, and Zhong Zhao**, "Evaluating continuous training programmes by using the generalized propensity score," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2012, *175* (2), 587–617.

Table 2: Descriptive statistics of asset regression data

| | min | 25% | median | 75% | max | mean | std | 0s | NAs | n |
|---|---|---|---|---|---|---|---|---|---|---|
| elapsed | 49 | 296 | 352 | 556 | 892 | 400.8 | 204.6 | 0 | 0 | 276 |
| size | 65 | 82.8 | 89 | 97 | 171 | 91.8 | 17 | 0 | 0 | 276 |
| ratioChildren | 0.3 | 0.4 | 0.4 | 0.5 | 0.6 | 0.4 | 0.1 | 0 | 0 | 276 |
| ratioAdults | 0.4 | 0.5 | 0.6 | 0.6 | 0.7 | 0.6 | 0.1 | 0 | 0 | 276 |
| ratioDisabled | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 168 | 0 | 276 |
| ratioMale | 0.4 | 0.5 | 0.5 | 0.5 | 0.6 | 0.5 | 0 | 0 | 0 | 276 |
| ratioLiterate | 0 | 0.2 | 0.3 | 0.4 | 0.5 | 0.3 | 0.1 | 0 | 0 | 276 |
| ratioLiterateMale | 0 | 0.1 | 0.2 | 0.2 | 0.3 | 0.2 | 0.1 | 0 | 0 | 276 |
| ratioHeadLiterate | 0 | 0 | 0 | 0 | 0.1 | 0 | 0 | 51 | 0 | 276 |
| avgDiffElapsed | 0 | 331.6 | 452.3 | 543.5 | 717.4 | 431.7 | 156.3 | 9 | 0 | 276 |
| avgDiffVal | −116 | −7.1 | 1 | 14.3 | 129.4 | 3 | 30.7 | 9 | 0 | 276 |
| avgVal | 1.2 | 15.2 | 43.2 | 87.6 | 192.4 | 56.4 | 48.3 | 0 | 0 | 276 |
| avgVal1 | 0.7 | 15.3 | 54.8 | 88.8 | 230.4 | 59.4 | 48.4 | 0 | 0 | 276 |
| avgVal0 | 1.2 | 15.2 | 43.2 | 87.6 | 192.4 | 56.4 | 48.3 | 0 | 0 | 276 |
| avgElapsed | 122.6 | 267.9 | 346.3 | 477.5 | 892 | 374.5 | 145.7 | 0 | 0 | 276 |
| avgElapsed1 | 544.3 | 729 | 834.8 | 857.3 | 899 | 806.2 | 71.8 | 0 | 0 | 276 |
| avgElapsed0 | 122.6 | 267.9 | 346.3 | 477.5 | 892 | 374.5 | 145.7 | 0 | 0 | 276 |

Table 3: DID estimates of asset impacts

| rn | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| (Intercept) | 3.511<br>(4.794) | 3.127<br>(4.744) | −177.737<br>(162.074) | −156.113<br>(162.405) |
| avgDiffElapsed | −0.006<br>(0.010) | | | |
| avgDiffElapsed * cow | | −0.008<br>(0.011) | −0.006<br>(0.010) | −0.003<br>(0.011) |
| avgDiffElapsed * large | | −0.003<br>(0.012) | −0.003<br>(0.011) | −0.003<br>(0.011) |
| avgDiffElapsed * large grace | | −0.006<br>(0.013) | −0.009<br>(0.013) | −0.005<br>(0.013) |
| avgDiffElapsed * traditional | | −0.004<br>(0.014) | −0.010<br>(0.014) | −0.008<br>(0.015) |
| size | | | −0.027<br>(0.107) | −0.007<br>(0.104) |
| ratioChildren | | | 150.652<br>(165.090) | 144.698<br>(161.753) |
| ratioAdults | | | 202.275<br>(165.389) | 191.880<br>(160.061) |
| ratioDisabled | | | −374.066** <br>(173.656) | −364.441** <br>(171.045) |
| ratioMale | | | 17.408<br>(33.939) | −28.147<br>(50.547) |
| ratioLiterate | | | | −33.128<br>(36.090) |
| ratioLiterateMale | | | | 105.570<br>(77.780) |
| ratioHeadLiterate | | | | −6.597<br>(110.909) |
| $R^2$<br>$n$ | 0.001<br>184 | 0.001<br>184 | 0.022<br>184 | 0.028<br>184 |

Notes: 1. Difference-in-differences estimates of asset accumulation against elapsed days.

2. large, large grace, cow are all time invariant and are interacted with a trend term.

3. $*$, $**$, $***$ indicate significance levels at 10%, 5%, 1%, respectively.

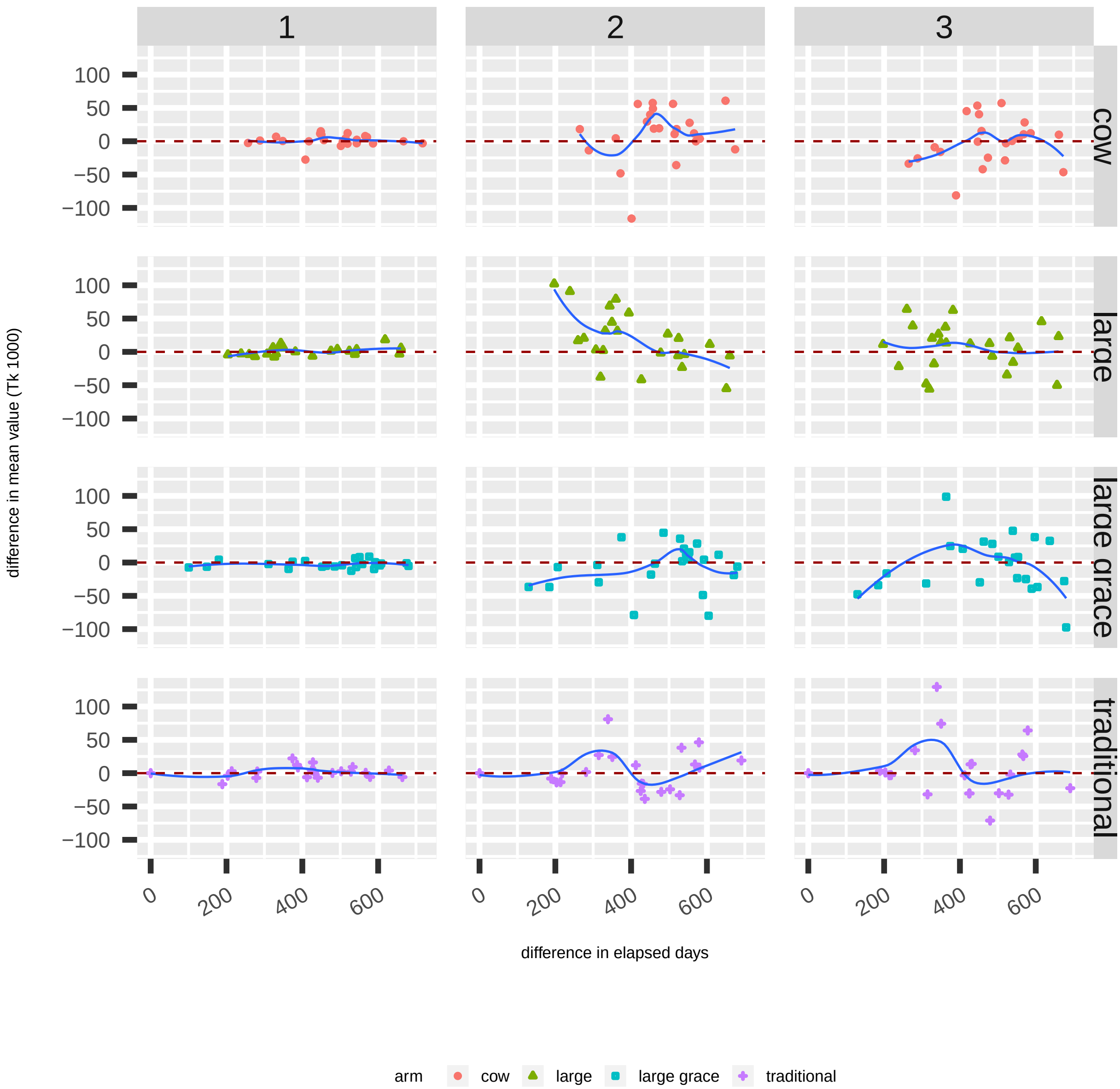


Figure 15 Total assets by elapsed days from disbursement, within a group