Sjekke_oyt

Bernhard Teodor Thodesen

2025-11-13

Replisering av tabeller

I denne delen av oppgaven skal jeg replisere tabeller fra artikkel Drivers of change: Employment responses to the lifting of the Saudi female driving ban(Abou Daher et al, 2025).

Laste inn data

```
library(haven) # Lese inn data sett
library(dplyr) # Databehandling
library(fixest) # Regresjon
library(knitr) # Tabeller
library(tidytext) # kompilering av fil
library(GenericML) # CLAN
library(mlr3) # ML
library(mlr3tuning) # Tuning
library(data.table) # Brukes med Ml3r

df <- read_dta("data/Combined_allwaves_final.dta") # Leser data

df <- df |>
  mutate(miss_age_PAP = ifelse(is.na(age_med_BL), 1, 0)) |>
  mutate(age_med_BL_control= ifelse(is.na(age_med_BL), 0, 1))|>
  mutate(miss_household_size= ifelse(is.na(household_size), 1, 0))|>
  mutate(household_size_control= ifelse(is.na(household_size), 0, household_size)) |>
  mutate(miss_edu_category= ifelse(is.na(edu_nohs_BL), 1, 0))|>
  mutate(edu_nohs_BL_control= ifelse(is.na(edu_nohs_BL), 0, edu_nohs_BL))|>
  mutate(married_control= ifelse(is.na(married), 0, married))|>
  mutate(divorced_separated_control= ifelse(is.na(divorced_separated), 0,
  divorced_separated))|>
  mutate(single_control= ifelse(is.na(single), 0, single))|>
  mutate(widowed_control= ifelse(is.na(widowed), 0, widowed))|>
  mutate(miss_relationship= ifelse(is.na(rel_status_BL), 1, 0))|>
  mutate(miss_cars= ifelse(is.na(cars), 1, 0))|>
  mutate(one_car_control= ifelse(is.na(one_car), 0, one_car))|>
  mutate(mult_cars_control= ifelse(is.na(mult_cars), 0, mult_cars))|>
  mutate(miss_LF_BL= ifelse(is.na(LF_BL), 1, 0))|>
  mutate(LF_BL_control= ifelse(is.na(LF_BL), 0, LF_BL)) |>
  filter(endline_start_w3==1)
```

Tabell 1

Tabell 1 i artikkelen viser estimert behandlingseffekten på ulike utfallsvariabler. For å replisere tabell 1 i artikkelen støtter jeg meg på kildekoden. Jeg lager seks regresjoner, en for hver av utfallsvariablene: `s_train_bi_w3`(Started driver's training), `license_w3`(Received license), `employed_w3`(Employed), `unemployed_w3`(Unemployed), `G1_2_abovemed`(Allowed to leave house w/o permission), `G1_3_abovemed`(Allowed to make purchase w/o permission).

Regresjonsmodellen har følgene form:

$$Y_i = \beta_0 + \beta_1 Treatment + \beta' \mathbf{X}_i + \alpha + \varepsilon_i$$

Her er Y_i en av utfallsvariablene som nevnt ovenfor. "Treatment" er en binær variabel som indikerer om noen har mottatt behandling. Variabelen $\mathbf{X}$ skal indikere alle kontrollvariablene vi bruker i regresjonen. α_i betegner bruken av fixed effects på variabelen `randomization_cohort2`(Randomization strata). Til slutt benyttes "clustered standard errors" på variabelen `file_nbr`(Houshold ID).

For å gjennomføre regresjonene har jeg benyttet meg av pakken `fixest`. Dette gjør jeg fordi pakken har mange nyttige funksjoner som gjør at replisering av tabellene i artikkelen blir lettere. For det første er det enkelt å legge til "fixed effects" i regresjonen. Det kan man gjøre ved å legge til en pipe `|`. Ta f.eks. denne regresjonen: `feols(y ~ x_1 + x_2 | fe1, data)`, hvor `fe1` er fixed effects. For det andre kan man enkelt benytte "clustered standard errors" som de gjør i artikkelen. I tillegg kan man benytte seg av `etable()` som kan brukes til å lage fine regresjonstabeller.

```
s_train <- feols(s_train_bi_w3 ~ treatment + age_med_BL_control + miss_age_PAP
                 + edu_nohs_BL_control + miss_edu_category + married_control
                 + single_control + widowed_control+ miss_relationship + household_size_control
                 + miss_household_size + one_car_control+ miss_cars + LF_BL_control
                 + miss_LF_BL |randomization_cohort2, cluster = c("file_nbr"), data = df)


licence <- feols(license_w3~ treatment + age_med_BL_control + miss_age_PAP
                 + edu_nohs_BL_control+ miss_edu_category + married_control
                 + single_control + widowed_control+ miss_relationship
                 + household_size_control + miss_household_size + one_car_control
                 + miss_cars + LF_BL_control + miss_LF_BL |randomization_cohort2,
                 cluster = c("file_nbr"), data = df)


empl <- feols(employed_w3~ treatment +age_med_BL_control + miss_age_PAP
              + edu_nohs_BL_control+ miss_edu_category + married_control
              + single_control + widowed_control+ miss_relationship + household_size_control
              + miss_household_size + one_car_control + miss_cars + LF_BL_control
              + miss_LF_BL |randomization_cohort2, cluster = c("file_nbr"), data = df)


not_empl <- feols(unemployed_w3~  treatment + age_med_BL_control + miss_age_PAP
                  + edu_nohs_BL_control+ miss_edu_category + married_control
                  + single_control + widowed_control+ miss_relationship
                  + household_size_control + miss_household_size + one_car_control
                  + miss_cars + LF_BL_control + miss_LF_BL |randomization_cohort2,
                  cluster = c("file_nbr"), data = df)

leave_house <- feols(G1_2_abovemed ~ treatment + age_med_BL_control + miss_age_PAP
                     + edu_nohs_BL_control+ miss_edu_category + married_control
```

```
                    + single_control + widowed_control+ miss_relationship
                    + household_size_control + miss_household_size + one_car_control
                    + miss_cars + LF_BL_control + miss_LF_BL |randomization_cohort2,
                    cluster = c("file_nbr"), data = df)


make_purchase <- feols(G1_3_abovemed ~ treatment + age_med_BL_control + miss_age_PAP
                      + edu_nohs_BL_control+ miss_edu_category + married_control
                      + single_control + widowed_control + miss_relationship
                      + household_size_control + miss_household_size + one_car_control
                      + miss_cars + LF_BL_control + miss_LF_BL |randomization_cohort2,
                      cluster = c("file_nbr"), data = df)
```

Utforming av tabeller, her benytter jeg meg av funksjonen `fitstat_register()` som gjør det mulig å utforme egne funksjoner som man kan bruke når man lager regresjonstabellene.

```
fitstat_register("control_mean", function(x) mean(x), "Control mean")

fitstat_register("mean", function(x) mean(x, na.rm = T), "control_m")

fitstat_register("pval", function(x) pvalue(x), "p-value b = 0")

fitstat_register("mean_c",

                 function(x){
                   name <- x$fml[2] |> as.character()
                   xer <- df |>
                   filter(treatment == 0) |>
                   select(name) |> pull()
                   mean(xer,na.rm = T)
                   },
                 "Control mean"
)

fitstat_register("me",

                 function(x){
                   name <- x$fml[2] |> as.character()
                   xer <- df |>
                   filter(treatment == 0) |>
                   select(name) |> pull()

                   x$coefficients[1]/ mean(xer,na.rm = T)
                   },
                 "b/Control mean"
)

controls <- c("miss_age_PAP","age_med_BL_control", "miss_edu_category",
              "married_control", "widowed_control","miss_relationship",
              "household_size_control", "miss_household_size","one_car_control",
              "mult_cars_control", "miss_cars","LF_BL_control", "LF_BL_control",
              "edu_nohs_BL_control", "single_control", "miss_LF_BL", "Constant",
              "randomization_cohort2")
```

```
setFixest_etable(drop.section = "fixef")

t1pa <- etable(s_train,licence,
       se.below=T,
       drop = controls,
       title ="Tabell 1 Panel A for Started driver`s training og Received license ",
       digits = "r3",digits.stats = "r3",
       tex = T,
       signif.code = NA,
       dict=c(treatment = "Treatment",
              s_train_bi_w3 = "Started driver`s training",
              license_w3 = "Received license",
              randomization_cohort2 = ""),
       style.tex = style.tex("qje", ),
       fitstat = ~ n + mean_c +me + pval)


t1pb <- etable(empl,not_empl ,
       se.below=T,
       drop = controls,
       title ="Tabell 1 Panel B for Employed og Unemployed",
       digits = "r3",digits.stats = "r3",
       tex = T,
       signif.code = NA,
       dict=c(treatment = "Treatment",
              employed_w3 = "Employed",
              unemployed_w3 = "Unemployed",
              randomization_cohort2 = ""),
       style.tex = style.tex("qje", ),
       fitstat = ~ n + mean_c +me + pval)


t1pc <- etable(leave_house, make_purchase,
       se.below=T,
       tex = T,
       drop = controls,
       depvar = F,
       title = "Tabell 1 Panel C for Allowed to leave house w/o permission og
       Allowed to make purchase w/o permission",
       digits = "r3",digits.stats = "r3",
       headers = .(":_:" = .("Agreement with following statements"),
                   ":_:" = c("Allowed to leave house w/o permission",
                             "Allowed to make purchase w/o permission")),
       signif.code = NA,
       dict=c(treatment = "Treatment",
              randomization_cohort2 = ""),
       style.tex = style.tex("qje", ),
       fitstat = ~ n + mean_c +me + pval)



t1pa;t1pb;t1pc
```

Table 1: Tabell 1 Panel A for Started driver‘s training og Received license

	Started driver‘s training	Received license
	(1)	(2)
Treatment	0.619	0.430
	(0.040)	(0.039)
Observations	467	467
Control mean	0.192	0.102
b/Control mean	3.229	4.221
p-value b = 0	0.000	0.000

Table 2: Tabell 1 Panel B for Employed og Unemployed

	Employed	Unemployed
	(1)	(2)
Treatment	0.085	-0.105
	(0.043)	(0.049)
Observations	488	488
Control mean	0.210	0.569
b/Control mean	0.405	-0.185
p-value b = 0	0.049	0.032

Table 3: Tabell 1 Panel C for Allowed to leave house w/o permission og Allowed to make purchase w/o permission

	Agreement with following statements	
	Allowed to leave house w/o permission	Allowed to make purchase w/o permission
	(1)	(2)
Treatment	0.062	-0.090
	(0.046)	(0.048)
Observations	488	486
Control mean	0.344	0.484
b/Control mean	0.179	-0.186
p-value b = 0	0.184	0.059

Tabell 1 Panel A for Started driver's training og Received license

Å motta tilbud om kjøreopplæring økte sannsynligheten for å ha startet kjøreopplæring med 61,9 (se = 0.04) prosentpoeng, noe som er signifikant. Gjennomsnittet i kontrollgruppen var 0,192 (19,2%), så gitt at en person er i behandlingsgruppen så er det en forventet 81 prosent sannsynlighet for at vedkommende startet med kjøreopplæring.

Behandlingen ga også en gjennomsnittlig økning på 43 (se = 0.039) prosentpoeng på om vedkommende klarte å ta førerkort. Sammenlignet med kontrollgruppen sin andel på ti prosent ser vi at effekten av å motta opplæring er om lag 4 ganger så stor som kontrollgruppa. Gitt at en person er i behandlingsgruppen er den forventede sannsynligheten for å ha tatt førerkort om lag 53 prosent for gjennomsnittpersonen.

Tabell 1 Panel B for Employed og Unemployed

I denne tabellen ser vi på hvordan behandlingen påvirker arbiedsstatus. Behandlingen økte sannsynligheten for å være i arbeid med 8.5 prosentpoeng sllik at behandlingsgruppa har en forventet sannsynlighet på 29,5 prosent for å være i arbeid. Resultat er akkurat signifikant under fem prosent med en p-verdi på 0,049.

Samtidig som at behandlingen øker sannsynligheten for å bli ansatt, synker sannsynligheten for å være uten arbeid ved behandling. Ved å være i behandlingsgruppa reduserte dette sannsynligheten for å være arbeidsledig med 10,5 prosentpoeng, estimatet er signifikant på 5 prosent med tilhørende p-verdi på 0,032. Dette gir at behandlingsgruppa har en forventet sannsynlighet på 0,464(46,4%) for å være arbeidsledig.

Tabell 1 Panel C for Allowed to leave house w/o permission og Allowed to make purchase w/o permission

I denne tabellen viser man hvordan behandlingen påvirker kvinnes selvstendighet i hverdagen. Ved behandling økte sannsynligheten for at kvinner forlot huset uten tillatelse med 6,2 prosentpoeng. Den forventede sannsynligheten for at kvinner gjør det er 0,406(40,6 %). Effekten derimot er ikke signifikant med en tilhørende p-verdi på 0,184.

Samtidig reduserer behandlingen tilbøyeligheten for å foreta kjøp uten tillatelse med 9 prosentpoeng, noe som gir en forventet sannsynlighet på 0,394(39.4 %). Dette resultatet er ikke signifikant på fem prosent med en p-verdi på 0,059, men kan tyde på at det har en viss innvirkning.

Sammenlignet med tabell 1 i artikkelen ser vi at tallene er relativt like. Alle estimatene peker i samme retning. Det er noen avvik her og der, men da dreier det seg om hundredeler. Resultatene samsvarer godt med artikkelen(Abou Daher et al, 2025, s. 3258-3259) som viser at analysen er reproduserbar.

Tabell 2

Tabell 2 viser hetrogeniteter i behandlingseffeketer. For å replisere tabell 2 i artikkelen støtter jeg meg på kildekoden(KILDE). Regresjonene prøver å estimere hetrogenitet i behandlingseffekter med hensyn på fem variabler: `age_med_BL`(Age above median indicator), `edu_nohs_BL`(Less than HIgh school education), `single`, `married` og `widowed` for tre ulike utfallsvariabler: `license_w3`(Received license), `employed_w3`(Employed) og `G1_3_abovemed`(Allowed to make purchase w/o permission). Regresjonene har følgende form:

$$Y_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Z_i + \beta_3 (Treatment_i \times Z_i) + \beta' \mathbf{X}_i + \alpha + \varepsilon_i$$

Regresjonen er nesten helt lik som for tabell 1, men leddet $\beta_2 Z_i + \beta_3 (Treatment_i \times Z_i)$ blir inkludert for å estimere hetrogenitet i behandlinseffekter. Her representerer Z_i en av variablene som skal undersøkes.

```
##### Panel a

rl2a <- feols(license_w3 ~ treatment + age_med_BL + treatment:age_med_BL + miss_age_PAP
              + miss_edu_category + married_control + single_control + widowed_control
              + miss_relationship + household_size_control + miss_household_size
              + one_car_control+ miss_cars  + miss_LF_BL |randomization_cohort2,
```

```
              cluster = c("file_nbr"),data = df)


emp2a <- feols(employed_w3 ~ treatment +age_med_BL+ treatment:age_med_BL+ miss_age_PAP
               + miss_edu_category + married_control + single_control + widowed_control
               + miss_relationship + household_size_control + miss_household_size
               + one_car_control+ miss_cars  + miss_LF_BL |randomization_cohort2,
               cluster = c("file_nbr"),data = df)


make_purchase2a <- feols(G1_3_abovemed ~ treatment +age_med_BL
                         + treatment:age_med_BL+ miss_age_PAP  + miss_edu_category
                         + married_control + single_control + widowed_control
                         + miss_relationship + household_size_control
                         + miss_household_size + one_car_control + miss_cars
                         + miss_LF_BL |randomization_cohort2, cluster = c("file_nbr")
                         ,data = df)


##### Panel b

rl2b <- feols(license_w3 ~ treatment + edu_nohs_BL  + treatment:edu_nohs_BL
              + miss_age_PAP + miss_edu_category + married_control + single_control
              + widowed_control+ miss_relationship + household_size_control
              + miss_household_size + one_car_control + miss_cars  + miss_LF_BL
              |randomization_cohort2, cluster = c("file_nbr"),data = df)


emp2b <- feols(employed_w3 ~ treatment + edu_nohs_BL + treatment:edu_nohs_BL
               + miss_age_PAP  + miss_edu_category + married_control + single_control
               + widowed_control + miss_relationship + household_size_control
               + miss_household_size + one_car_control+ miss_cars  + miss_LF_BL
               |randomization_cohort2, cluster = c("file_nbr"),data = df)


make_purchase2b <- feols(G1_3_abovemed ~ treatment + edu_nohs_BL
                         + treatment:edu_nohs_BL + miss_age_PAP + miss_edu_category
                         + married_control + single_control + widowed_control
                         + miss_relationship + household_size_control + miss_household_size
                         + one_car_control + miss_cars  + miss_LF_BL
                         |randomization_cohort2, cluster = c("file_nbr")
                         ,data = df)

##### Panel c


rl2c <- feols(license_w3 ~ treatment + single_control + married_control+ widowed_control
              + treatment:married + treatment:single + treatment:widowed
              + miss_age_PAP + miss_edu_category + miss_relationship
              + household_size_control + miss_household_size + one_car_control
              + miss_cars  + miss_LF_BL
              |randomization_cohort2, cluster = c("file_nbr")
              ,data = df)
```

```
emp2c <- feols(employed_w3 ~ treatment + single_control + married_control+ widowed_control
               + treatment:married + treatment:single + treatment:widowed
               + miss_age_PAP + miss_edu_category + miss_relationship
               + household_size_control + miss_household_size + one_car_control
               + miss_cars  + miss_LF_BL
               |randomization_cohort2, cluster = c("file_nbr")
               ,data = df)

make_purchase2c <- feols(G1_3_abovemed~ treatment + single_control + married_control
                         + widowed_control+ treatment:married + treatment:single
                         + treatment:widowed + miss_age_PAP + miss_edu_category
                         + miss_relationship + household_size_control + miss_household_size
                         + one_car_control+ miss_cars  + miss_LF_BL |randomization_cohort2,
                         cluster = c("file_nbr"),data = df)
```

Utforming av tabeller.

```
hte_var <- c("age_med_BL_control", "edu_nohs_BL_control","LF_BL_control" )

panelb_controls <- c("miss_age_PAP", "miss_edu_category", "married_control",
                     "widowed_control","miss_relationship", "household_size_control",
                     "miss_household_size","one_car_control", "mult_cars_control",
                     "miss_cars", "LF_BL_control", "single_control", "miss_LF_BL",
                     "Constant", "randomization_cohort2")


fitstat_register("samlet",
        function(x){
          antall_koeff <- length(x$coefficients)
          b_1 <- x$coefficients[1]
          b_3 <- x$coefficients[antall_koeff]

          round(b_1 + b_3,3)
        },
        "B1 + B3"
        )


fitstat_register("se",

                 function(x){
          antall_koeff <- length(x$coefficients)

          v_b_1 <- x$se[1]**2
          v_b_3 <- x$se[antall_koeff]**2

          cov13 <- x$cov.unscaled[antall_koeff]

          se <- sqrt(v_b_1 + v_b_3 + 2*cov13)

          paste0("(",round(se,3),")")
        },
        " "
```

```
        )

fitstat_register("median_age",
                 function(x){

                   name <- formula(x)[2] |> as.character()

                   df |>
                     filter(treatment == 0 & age_med_BL == 0) |>
                     pull(name) |> mean(na.rm = T) |> round(3)

                 }, "Mean: control, below median age")


fitstat_register("hs",
                 function(x){

                   name <- formula(x)[2] |> as.character()

                   df |>
                     filter(treatment == 0 & edu_nohs_BL == 0) |>
                     pull(name) |> mean(na.rm = T) |> round(3)

                 },
                 "Mean: control, completed HS")
fitstat_register("mcd",
                 function(x){

                   name <- formula(x)[2] |> as.character()
                   df |> filter(treatment == 0 & divorced_separated == 1) |>
                     pull(name) |> mean(na.rm = T) |> round(3)


                 }, "Mean: control, divorced")

fitstat_register("mcm",
                 function(x){

                   name <- formula(x)[2] |> as.character()
                   df |> filter(treatment == 0 & married == 1) |>
                     pull(name) |> mean(na.rm = T) |> round(3)


                 }, "Mean: control, married")
fitstat_register("mcnm",
                 function(x){

                   name <- formula(x)[2] |> as.character()
                   df |> filter(treatment == 0 & single == 1) |>
                     pull(name) |> mean(na.rm = T) |> round(3)


                 }, "Mean: control, never married")
```

```
fitstat_register("mcw",
                 function(x){

                   name <- formula(x)[2] |> as.character()
                   df |> filter(treatment == 0 & widowed == 1) |>
                     pull(name) |> mean(na.rm = T) |> round(3)


                 }, "Mean: control, widowed")



panela <- etable(rl2a,emp2a,make_purchase2a, se.below = T,
       tex = T, title = "Panel A",
       drop = panelb_controls,
       digits = "r3",digits.stats = "r3",
       signif.code = NA,
       dict=c(treatment = "Treatment",
              license_w3 = "Received license",
              G1_3_abovemed = "Allowed to make purchase w/o permission",
              employed_w3 = "Employed",
              age_med_BL = "Above median age",
              randomization_cohort2 = ""),
       style.tex = style.tex("qje", ),
       fitstat = ~samlet+ se+ n +  median_age);

panelb <- etable(rl2b,emp2b,make_purchase2b, se.below = T,
       tex = T, title = "Panel B",
       drop = panelb_controls,
       digits = "r3",digits.stats = "r3",
       signif.code = NA,
       dict=c(treatment = "Treatment",
              license_w3 = "Received license",
              G1_3_abovemed = "Allowed to make purchase w/o permission",
              employed_w3 = "Employed",
              edu_nohs_BL = "Less than HS",
              randomization_cohort2 = ""),
       style.tex = style.tex("qje", ),
       fitstat = ~samlet+ se+ n + hs )

panelc <- etable(rl2c, emp2c, make_purchase2c, se.below = T,
       tex = T, title = "Panel C",
       drop = panelb_controls,
       digits = "r3",digits.stats = "r3",
       signif.code = NA,
       dict=c(treatment = "Treatment",
              license_w3 = "Received license",
              G1_3_abovemed = "Allowed to make purchase w/o permission",
              employed_w3 = "Employed",
              single = "never married",
              randomization_cohort2 = ""),
       style.tex = style.tex("qje", ),
       fitstat = ~n+mcd + mcm + mcnm + mcw)
```

```
panela;panelb;panelc
```

Table 4: Panel A

	Received license (1)	Employed (2)	Allowed to make purchase w/o permission (3)
Treatment	0.526	0.139	0.038
	(0.056)	(0.063)	(0.070)
Above median age	0.150	-0.041	0.274
	(0.062)	(0.075)	(0.089)
Treatment × Above median age	-0.189	-0.105	-0.246
	(0.076)	(0.082)	(0.092)
B1 + B3	0.338	0.034	-0.208
	(0.033)	(0.057)	(0.067)
Observations	467	488	486
Mean: control, below median age	0.092	0.247	0.329

Table 5: Panel B

	Received license (1)	Employed (2)	Allowed to make purchase w/o permission (3)
Treatment	0.507	0.078	-0.023
	(0.052)	(0.055)	(0.059)
Less than HS	-0.001	-0.125	0.080
	(0.057)	(0.060)	(0.079)
Treatment × Less than HS	-0.230	0.029	-0.186
	(0.079)	(0.077)	(0.093)
B1 + B3	0.277	0.107	-0.209
	(0.058)	(0.059)	(0.074)
Observations	459	479	477
Mean: control, completed HS	0.129	0.265	0.451

Table 6: Panel C

	Received license (1)	Employed (2)	Allowed to make purchase w/o permission (3)
Treatment	0.433	-0.029	-0.221
	(0.059)	(0.072)	(0.072)
Treatment × married	-0.156	0.105	0.058
	(0.101)	(0.111)	(0.131)
Treatment × never married	0.161	0.175	0.344
	(0.087)	(0.107)	(0.108)
Treatment × widowed	-0.234	0.295	0.031
	(0.145)	(0.135)	(0.159)
Observations	463	484	482
Mean: control, divorced	0.083	0.250	0.597
Mean: control, married	0.091	0.171	0.472
Mean: control, never married	0.080	0.246	0.293
Mean: control, widowed	0.208	0.080	0.654

TOlking av REsultater

Resultate