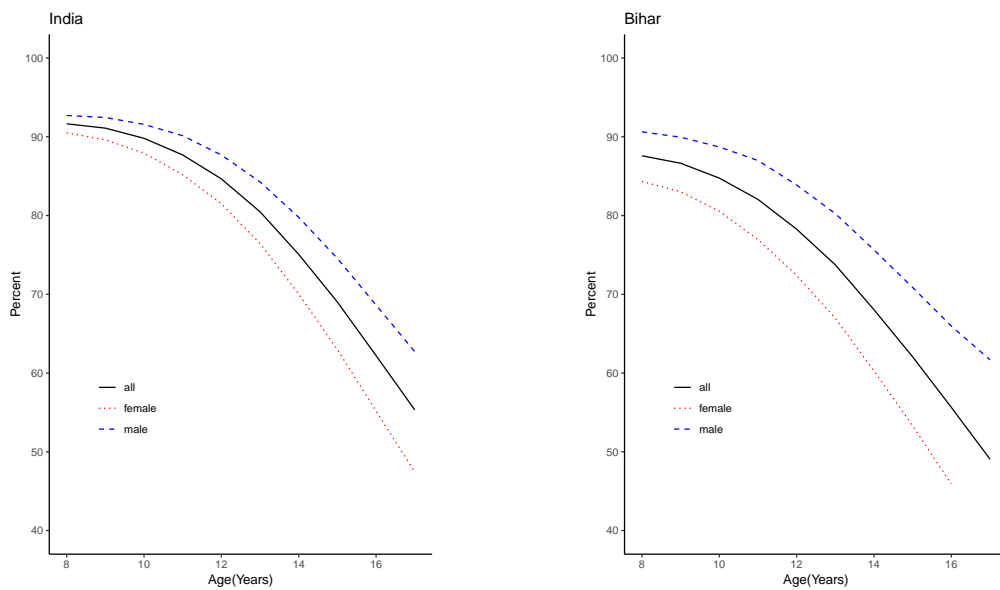# Replication of Muralidharan and Prakash (2017)

2024 年 8 月 31 日

1 論文の概要

インドでは secondary school への進学率における男女間格差が長年問題になっている。この論文では、女子学生に自転車を給付する政策がジェンダーギャップをどれほど解消したかを検証した。結果として、女子生徒入学率が 32 ％（16.3 ％→ 21.5 ％）上昇し、男女間の進学率格差が 40 ％（12.9 ％→ 7.7 ％）改善された。家計調査データに対し、Triple Differences（DDD）を用いて分析を行った。

Figure1 のレプリケーション結果: 年齢に応じたインドと Bihar 州における在学率

| Treatment group = age 14 and 15 | | | | |
Control group = age 16 and 17	(1)	(2)	(3)	(4)
Treat × female × Bihar	0.103	0.091	0.052	0.051
	(0.030)	(0.029)	(0.025)	(0.025)
Treat × female	0.020	0.024	0.038	0.039
	(0.026)	(0.026)	(0.021)	(0.021)
Treat × Bihar	−0.044	−0.042	−0.029	−0.028
	(0.018)	(0.018)	(0.016)	(0.016)
Female × Bihar	−0.094	−0.091	−0.067	−0.066
	(0.023)	(0.023)	(0.020)	(0.020)
Treat	−0.148	−0.143	−0.138	−0.138
	(0.014)	(0.014)	(0.013)	(0.013)
Female	−0.092	−0.088	−0.100	−0.101
	(0.020)	(0.020)	(0.017)	(0.017)
Bihar	0.011	−0.044	−0.032	−0.042
	(0.016)	(0.016)	(0.015)	(0.015)
Constant	0.464	0.771	0.593	0.557
	(0.013)	(0.024)	(0.027)	(0.039)
Demographic Controls	No	Yes	Yes	Yes
HH socioeconomic controls	No	No	Yes	Yes
Village level controls	No	No	No	Yes
Observation	30 295	30 295	30 147	30 112
R^2	0.035	0.088	0.207	0.208

Table 2 のレプリケーション結果

2 ソースコード

Listing 1: クリーニングのコード

```r
setwd("C:/Users/Owner/Desktop/Indian-cycling-A")
#install.packages("pacman")
pacman::p_load(tidyverse, haven)
df_longwd <- read_dta("analysis_cleaning/dlhs_long_wdist.dta")

#欠損値処理
df_longwd$school[is.na(df_longwd$school)] <-0
df_longwd$grade[is.na(df_longwd$grade)] <-0
df_longwd$relationship[is.na(df_longwd$relationship)] <- 0

#フィルター
df <- df_longwd |>
  filter(state == 10 | state == 20) |>
  mutate(inschool=if_else(school==1, 1, 0), #Indicator variable for being in school
         currgrade = if_else(inschool == 1, grade + 1, grade), #Current grade
         enrollment_secschool= if_else(currgrade==9|grade>=9, 1, 0),
         enrollment_middleschool= if_else(currgrade==8|grade>=8, 1, 0),
         female=sex-1,   #がfemale1
         bihar=if_else(state==10, 1, 0),
         child_sample=if_else(relationship == 3|relationship == 5|relationship == 8|relationship
             == 10, 1, 0),
         sc=if_else(hv116b == 1, 1, 0),
         )
```

Listing 2: Figure1 のコード

```r
setwd("C:/Users/Owner/Desktop/Indian-cycling-A/analysis_cleaning")
library(tidyverse)
library(ggplot2)
dist.data<-read_dta("schooldist.dta")
age.data<-read.dta("schoolage.dta")

#india figure A.1
figure_A1 = ggplot(age.data,aes(x=age))+
  geom_line(aes(y=100*inschool_india, colour = "all"))+
  geom_line(aes(y=100*inschool_india_male, colour ="male"), linetype="dashed")+
  geom_line(aes(y=100*inschool_india_female, colour = "female"), linetype="dotted")+
  labs(title="India", x="Age(Years)", y="Percent")+
  scale_x_continuous(breaks=seq(8, 16, by=2), labels=seq(8, 16, by=2))+
  scale_y_continuous(limits=c(40,100), breaks=seq(40, 100, by=10), labels=seq(40, 100, by=10))+
  scale_color_manual(name = "", values = c("all" = "black", "male" = "blue", "female" ="red"),
  labels = c("all" = "all","male" = "male","female" = "female")) +
  theme_classic()+
  theme(legend.position = c(0.2,0.3))

#india bihar figure A.2
figure_A2 = ggplot(age.data,aes(x=age))+
  geom_line(aes(y=100*inschool_bihar, colour = "all"))+
  geom_line(aes(y=100*inschool_bihar_male, colour ="male"), linetype="dashed")+
  geom_line(aes(y=100*inschool_bihar_female, colour = "female"), linetype="dotted")+
  labs(title="Bihar", x="Age(Years)", y="Percent")+
  scale_x_continuous(breaks=seq(8, 16, by=2), labels=seq(8, 16, by=2))+
  scale_y_continuous(limits=c(40,100),
  breaks=seq(40, 100, by=10), labels=seq(40, 100, by=10))+
  scale_color_manual(name = "",
  values = c("all" = "black", "male" = "blue", "female" ="red"),
  labels = c("all" = "all","male" = "male","female" = "female"))+
  theme_classic()+
  theme(legend.position = c(0.2,0.3))
```

Listing 3: Table2 のソースコード

```
1   setwd("C:/Users/Owner/Desktop/Indian-cycling-A/analysis_cleaning")
2   install.packages("fixest")
3   library("fixest")
4   pacman::p_load(modelsummary, stargazer)
5
6   df_reg <- read_dta("dlhs-reg-data.dta")
7
8   model1 <- feols(enrollment_secschool~ treat1_female_bihar +
9                   treat1_female + treat1_bihar + female_bihar +
10                  treat1 + female + bihar,
11                  data=df_reg,
12                  weights =~  hhwt,
13                  cluster =~ village)
14  model2 <- feols(enrollment_secschool~ treat1_female_bihar +
15                  treat1_female + treat1_bihar + female_bihar +
16                  treat1 + female + bihar
17                  + sc +   st+ obc + hindu +muslim,
18                  data=df_reg,
19                  weights =~  hhwt, cluster =~ village)
20
21
22  model3 <- feols(enrollment_secschool~ treat1_female_bihar +
23                  treat1_female + treat1_bihar + female_bihar +
24                  treat1 + female + bihar
25                  + sc +   st+ obc + hindu + muslim
26                  + hhheadschool + hhheadmale + land + bpl + media+ electricity,
27                  data=df_reg,
28                  weights =~  hhwt, cluster =~ village)
29
30
31  model4 <- feols(enrollment_secschool~ treat1_female_bihar +
32                  treat1_female + treat1_bihar + female_bihar +
33                  treat1 + female + bihar
34                  + sc +   st+ obc + hindu + muslim
35                  + hhheadschool + hhheadmale + land + bpl + media+ electricity
36                  + middle + bank + postoff + lcurrpop,
37                  data=df_reg,
38                  weights =~  hhwt, cluster =~ village)
39
40  models <- list(model1, model2, model3, model4)
41
42  msummary(models,
43              coef_map = c("treat1_female_bihar" = "Treat × female × Bihar",
44                          "treat1_female" = "Treat × female",
45                          "treat1_bihar" = "Treat × Bihar",
46                          "female_bihar" = "Female × Bihar",
47                          "treat1" = "Treat",
48                          "female" = "Female",
49                          "bihar" = "Bihar",
50                          "(Intercept)" = "Constant"),
51              gof_omit = 'R2 Adj.|AIC|BIC|RMSE|Std.Errors', "latex")
```

3 振り返り

- 現時点でチームメンバーが持っているスキルを合わせた結果、自分たちが再現したかった範囲を完成させることが出来た。
- 欠損値処理の方法について、もう少し議論が必要だと感じた。
- データから図を作成する際には、雑然データを整然データにする重要性を感じた。