

Timely Microfinance Repayment: Survey Results from Puebla, Mexico.

Online Appendix

Data Base

Data is imported to R and then cleaned for missing values; several variables are re-defined:

```
# import with more data
ProvidentDF_complete<- read.csv("BaseDeDatos/Encuesta-
EducacionFinancieraDario.csv")

ProvidentDF_complete <- ProvidentDF_complete %>%
filter(Status=="Completa") %>%
  filter(!is.na(p9_batepelota)) # question 9 and 12 have missing values
# measure grit as the average of the points made in each item
ProvidentGRIT <- ProvidentDF_complete[,c(
  "X1..Los.proyectos.o.ideas.nuevas.me.distraen.de.proyectos.o.ideas.que.te
nía.desde.antes.",
  "X2..Los.óbstaculos.no.me.desaniman.",
  "X3..Estuve.concentrado.en.una.idea.o.proyecto.por.un.corto.tiempo..pero.
después.perdí.interés.",
  "X4..Trabajo.duro.",
  "X5..Con.frecuencia.me.propongo.un.objetivo..pero.luego.trato.de.cumplir.
un.objetivo.diferente",
  "X6..Me.resulta.dificil.mantener.mi.atención.en.proyectos.que.duran.más.a
llá.de.algunos.meses.en.terminar",
  "X7..Soy.chambeador.a.")

ProvidentGRIT <- apply(substr(apply(ProvidentGRIT, MARGIN = 2,FUN =
as.character),
                                start = 1,stop = 1),FUN = as.numeric,MARGIN
= 2)
ProvidentGRIT[,c(1,3,5,6)] <- -ProvidentGRIT[,c(1,3,5,6)] + 4
ProvidentDF_nna <- ProvidentDF_complete %>% # there is no missing values
in this DF
  mutate(grit=rowMeans(ProvidentGRIT))
ProvidentDF_nna <-
ProvidentDF_nna[!is.na(ProvidentDF_nna$p9_batepelota),]
```

```

# Education
escol <- as.character(ProvidentDF_nna$escol)
escol[escol=="No estudió"] <- 0 # revisar la configuración de acentos
escol[escol=="1ro Primaria"] <- 1
escol[escol=="2do Primaria"] <- 2
escol[escol=="3ro Primaria"] <- 3
escol[escol%in%c("4to primaria", "4to Primaria")] <- 4
escol[escol=="5to Primaria"] <- 5
escol[escol%in%c(" 6to Primaria", "6to Primaria")] <- 6
escol[escol=="1ro Secundaria"] <- 7
escol[escol%in%c("2do Secundaria", "2do secundaria")] <- 8
escol[escol%in%c("3ro secundaria", "3ro Secundaria ", "3ro Secundaria")] <-
9
escol[escol=="1ro Preparatoria"] <- 10
escol[escol=="2do Preparatoria"] <- 11
escol[escol=="3ro Preparatoria"] <- 12
escol[escol=="Licenciatura"] <- 13
escol[escol=="Posgrado"] <- 14
escol <- as.numeric(escol)

#original measure
escol1 <- escol
escol1[escol1%in%c(1:6)] <- 0 # Primaria
escol1[escol1%in%c(7:9)] <- 1 # Secundaria
escol1[escol1%in%c(10:12)] <- 2 # Preparatoria
escol1[escol1%in%c(13)] <- 3 # Licenciatura
escol1[escol1%in%c(14)] <- 4 # Posgrado

#some order for ordinal data
ProvidentDF_nna$grupo <- factor(ProvidentDF_nna$grupo,
                                levels=c("Current", "Low Arrear", "High
arrear"), ordered=TRUE)

attach(ProvidentDF_nna)

## The following object is masked _by_ .GlobalEnv:
##
##      escol

```

Re-definition of (some) variables:

Just married and not married

- Married = 1

Employment as an ordinal variable.

- Full-time employment = 3
- Half-time employment = 2

- Self-employment = 1
- No earnings = 0

Tandas

- p10_tandacorto: Si obtuviera \$1,000 pesos de la participación en una tanda y tuviera que elegir entre dos alternativas de pago. ¿Cuál de las siguientes opciones elegiría?

* Opción 1: Que le entreguen los \$1,000 pesos en un mes = 0

* Opción 2: Que le entreguen \$1,100 en dos meses = 1

- p11_tandalargo: Si obtuvieras \$1,000 pesos de la participación en una tanda y tuviera que elegir entre dos alternativas de pago. ¿Cuál de las siguientes opciones elegiría?

* Opción 1: Que le entreguen los \$1,000 pesos en un año y un mes = 0

* Opción 2: Que le entreguen \$1,100 en un año y dos meses = 1

Inflation

- p13_preciosbajan: Cuando los precios bajan puedo comprar menos bienes y servicios con el mismo ingreso. ¿Verdadero o Falso? (la respuesta correcta es Falso)

* Verdadero = 1

* Falso = 0

```
p16_dependents<- p16_1menores+p16_2adulmay+p16_3otros # total number of dependents
```

```
casado<-{} # to capture later de degree of responsibility, we just need married or unmarried people
```

```
casado[p15_estadocivil=="Casado"]<- 1
```

```
casado[p15_estadocivil!="Casado"]<- 0
```

```
empleoRemunerado <- {} # consider non paid work together
```

```
empleoRemunerado[statusempleo%in%c("Desempleados","Ama de Casa","Retirado")]<- 0
```

```
empleoRemunerado[statusempleo=="Auto empleado"]<- 1
```

```
empleoRemunerado[statusempleo=="Empleado Medio tiempo"]<- 2
```

```
empleoRemunerado[statusempleo=="Empleado tiempo completo"]<- 3
```

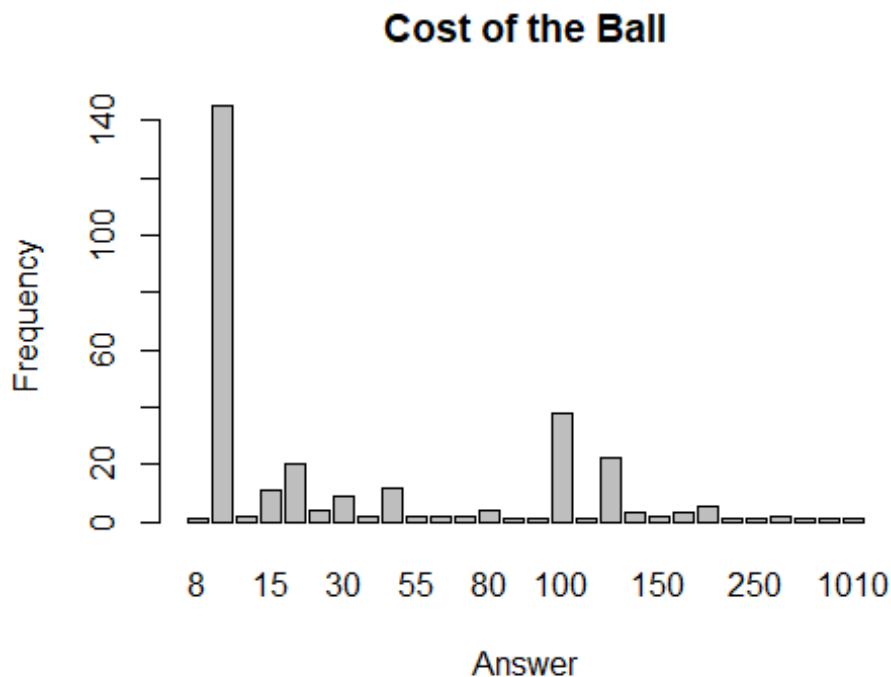
```
p10_tandacorto <- p10_tandacorto-1
```

```
p11_tandalargo <- p11_tandalargo-1
```

```
p13_preciosbajan <- as.numeric(p13_preciosbajan)-2
```

- Financial literacy is measured as a distance; the individual's answer minus the correct answer:

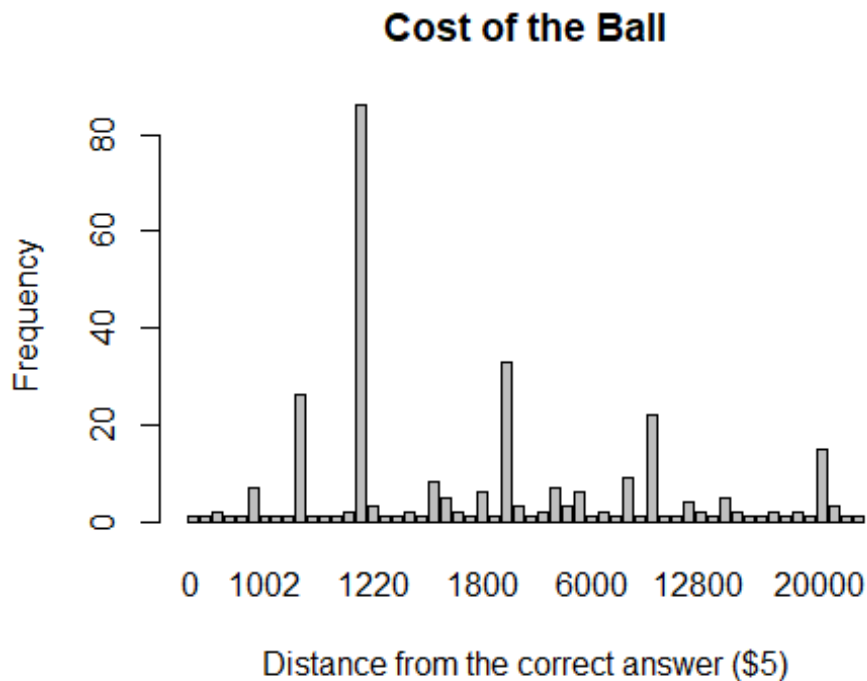
```
# pregunta del bat y la pelota:
# We can see that most people answered predictably bad: $ 10. But no one
answered correctly.
barplot(table(p9_batepelota),xlab = "Answer",ylab="Frequency",main =
"Cost of the Ball")
```



```
#9. Un bate y una pelota de b?isbol cuestan en total $110 pesos.
#El bate cuesta $100 m?s que la pelota,
#?Cu?nto cuesta la pelota? (Indique el costo en pesos de la pelota)
## Respuesta correcta: $5
```

```
p9_batepelota<- abs(p9_batepelota-5)
```

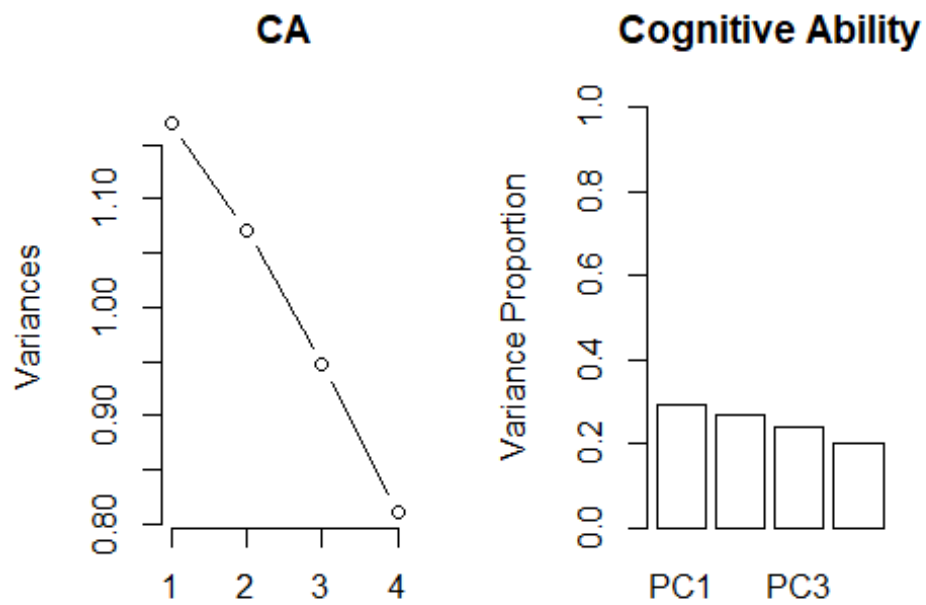
```
# pregunta sobre el Inter?s
barplot(table(p12_retornodeposito),xlab = "Distance from the correct
answer ($5)",ylab="Frequency",main = "Cost of the Ball")
```



#12. Imagínese que usted deposita \$1,000 pesos al inicio del año en una cuenta de ahorro
 #con un interés garantizado del 2% al año y la cuenta no tiene ningún costo por mantenerla.
 #Además, suponga que usted no saca dinero de esa cuenta.
 #¿Cuánto dinero tendría en la cuenta después de un año incluyendo el pago de los intereses?
 ## Respuesta Correcta: 1020
 p12_retornodeposito <- abs(p12_retornodeposito-1020)

PCA analysis for Cognitive Ability

```
# cognitive ability with less levels for escol
CA <- prcomp(na.omit(cbind( p9_batepelota,
p12_retornodeposito,p13_preciosbajan,escol1)),center = T,scale. = T)
```



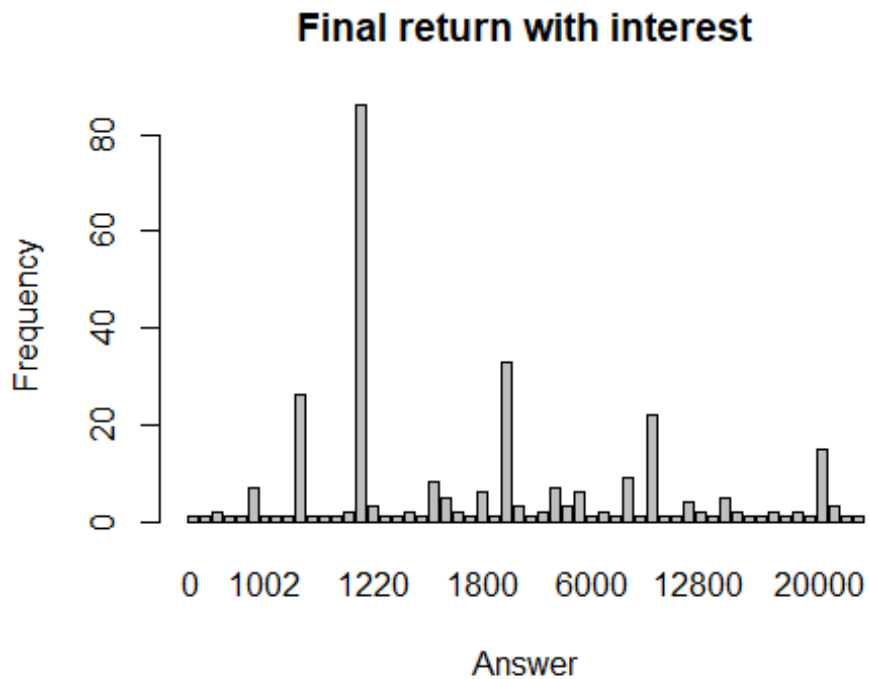
PCA weigths

CA\$rotation

##	PC1	PC2	PC3	PC4
## p9_batepelota	-0.14138706	0.71144797	-0.6054236	-0.32758770
## p12_retornodeposito	-0.08424451	0.64241756	0.7613148	0.02454236
## p13_preciosbajan	-0.67542768	-0.28069088	0.1832877	-0.65682242
## escol1	0.71882544	-0.04851902	0.1423644	-0.67872548

Final PCA's

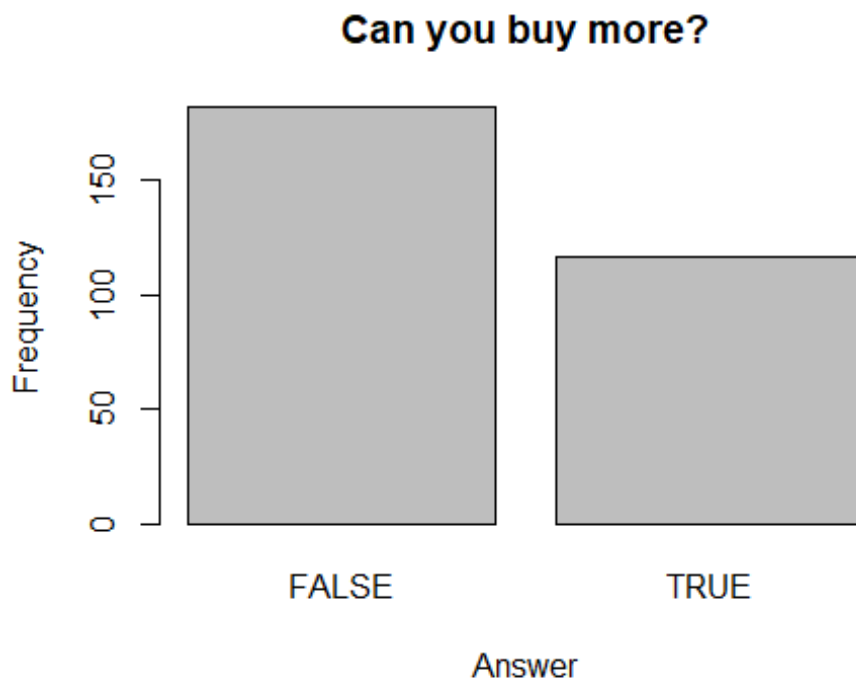
```
#ggplot(data =
ProvidentDF_nna,aes(x=factor(p12_retornodeposito)))+geom_bar()
barplot(table(ProvidentDF_nna$p12_retornodeposito),xlab =
"Answer",ylab="Frequency",main = "Final return with interest")
```



```
ProvidentDF_nna <- ProvidentDF_nna %>%
mutate(p13_preciosbajan=revalue(p13_preciosbajan,c("FALSO"="FALSE", "VERDA
DERO"="TRUE")))
#ggplot(data =
ProvidentDF_nna,aes(x=factor(p13_preciosbajan)))+geom_bar()
table(p13_preciosbajan)

## p13_preciosbajan
##    0    1
## 182 117

barplot(table(ProvidentDF_nna$p13_preciosbajan)[2:3],xlab =
"Answer",ylab="Frequency",main = "Can you buy more?")
```



Regression analysis

Specification without controlling of the number of credits

```
#Dependent Variables
DV <- cbind.data.frame(
  grupo = grupo,
  Grit = grit,
  Short_Delay = p10_tandacorto,
  Long_Delay = p11_tandalargo,
  CA1 = CA$x[,1], CA2 = CA$x[,2],
  Age = edad, hijos, casado, p16_dependents,
  Active_Loans= Prestamos.activos,
  Genero)

#weights
weights_groups <- read_excel("BaseDeDatos/weights.xlsx")
wt <- weights_groups$Weight[match(DV$grupo, weights_groups$Category)]
wt <- round(wt, 0)
weights_groups$Weight <- round(weights_groups$Weight, 0)
stargazer(weights_groups, header = F, title = "Weights for each
Group", summary = F)
```



```
weighted_regression <- polr(
  grupo ~ Grit+Short_Delay +Long_Delay+CA1+Age,
  data = DV,Hess=TRUE,na.action = na.omit, weights = wt) # model
weighted_regression
```

Call: polr(formula = grupo ~ Grit + Short_Delay + Long_Delay + CA1 + Age, data = DV, weights = wt, na.action = na.omit, Hess = TRUE)

Coefficients: Grit Short_Delay Long_Delay CA1 Age -0.44583048 0.91883668 -
0.44785813 0.09983325 -0.01785037

Intercepts: Current|Low Arrear Low Arrear|High arrear -2.6084638 0.1008318

Residual Deviance: 18959.63 AIC: 18973.63

```
stargazer(weighted_regression,header = F,title = "Final Regression with
Weights")
```

Marginal effects. There are three tables with the inference in terms of standard deviations. Each table corresponds to a category, and shows the effect of independent variables. Finally, there is a table that summarise the effects.

```
#Marginal Effects
me_weighted_final <- erer::ocME(weighted_regression,rev.dum = T)

## Warning in z[1:J] - xb: Recycling array of length 1 in vector-array
arithmetic is deprecated.
## Use c() or as.vector() instead.

## Warning in z[2:(J + 1)] - xb: Recycling array of length 1 in vector-
array arithmetic is deprecated.
## Use c() or as.vector() instead.

stargazer(me_weighted_final$out,header = F)
```

Specification considering the No of credits borrowed by each individual

```
prestamos_regression <- polr(
  grupo ~ Grit+Short_Delay +Long_Delay+CA1+Age+Active_Loans,
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)

stargazer(prestamos_regression,
  add.lines = list( # AIC and BIC
    c("AIC", round(AIC(
      prestamos_regression),digits = 2)),
    c("BIC", round(BIC(
      prestamos_regression),digits = 2))
```

```
),header = F,title = "Regressions Considering Debt and  
Weights")
```

For regression this we calculate the marginal effects:

```
#Marginal Effects  
me_weighted_final2 <- erer::ocME(prestamos_regression,rev.dum = T)  
  
## Warning in z[1:J] - xb: Recycling array of length 1 in vector-array  
arithmetic is deprecated.  
## Use c() or as.vector() instead.  
  
## Warning in z[2:(J + 1)] - xb: Recycling array of length 1 in vector-  
array arithmetic is deprecated.  
## Use c() or as.vector() instead.  
  
stargazer(me_weighted_final2$out,header = F)
```

Robustness

Without Debt

```
final_regressions3 <- polr(  
  grupo ~ Grit+Short_Delay +Long_Delay+CA1,  
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)  
  
final_regressions4 <- polr(  
  grupo ~ Grit+Short_Delay +Long_Delay+Age,  
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)  
  
final_regressions5 <- polr(  
  grupo ~ Grit+Short_Delay +CA1+Age,  
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)  
  
final_regressions6 <- polr(  
  grupo ~ Grit+Long_Delay+CA1+Age,  
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)  
  
final_regressions7 <- polr(  
  grupo ~ Short_Delay +Long_Delay+CA1+Age,  
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)
```

With Debt

```
final_regressions3 <- polr(  
  grupo ~ Grit+Short_Delay +Long_Delay+CA1+Active_Loans,  
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)  
  
final_regressions4 <- polr(  
  grupo ~ Grit+Short_Delay +Long_Delay+Age+Active_Loans,  
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)
```

```

final_regressions5 <- polr(
  grupo ~ Grit+Short_Delay +CA1+Age+Active_Loans,
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)

final_regressions6 <- polr(
  grupo ~ Grit+Long_Delay+CA1+Age+Active_Loans,
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)

final_regressions7 <- polr(
  grupo ~ Short_Delay +Long_Delay+CA1+Age+Active_Loans,
  data = DV,Hess=TRUE,na.action = na.omit,weights = wt)

```

Final Regression Without Weights

```

final_regression <- polr(
  grupo ~ Grit+Short_Delay +Long_Delay+CA1+Age,
  data = DV,Hess=TRUE,na.action = na.omit)

debt_regression <- polr(
  grupo ~ Grit+Short_Delay +Long_Delay+CA1+Age+Active_Loans,
  data = DV,Hess=TRUE,na.action = na.omit) # model

```