Report for statisticians

Becky Yuen

2024-03-24

Introduction

This is the overall report for the analysis on the European Value Study (EVS) from 2017 which is a survey research program on how Europeans think about family, work, religion, politics, and society. We are mainly interested in Enropeans thoughts on two questions:

1. When a mother works for pay, do Europeans think the children suffer?
2. When jobs are scarce, do Europeans think employers should give priority to local people over immigrants?

```
library(haven)
EVS = read_sav("../data/EVS_data_cleaned.sav")
```

Descriptives of variables

In the following table, the variables are:

1. `v72` represents the first question of interest (1-strongly agree, 2-agree, 3-disagree, or 4-strongly disagree)
2. `v80` represents the second question of interest (1-strongly agree, 2-agree, 3-neither agree nor disagree, 4-disagree, or 5-strongly disagree)
3. `sex` (1-male or 2-female)
4. `age` (years)
5. `education` (1-lower, 2-medium, or 3-higher)

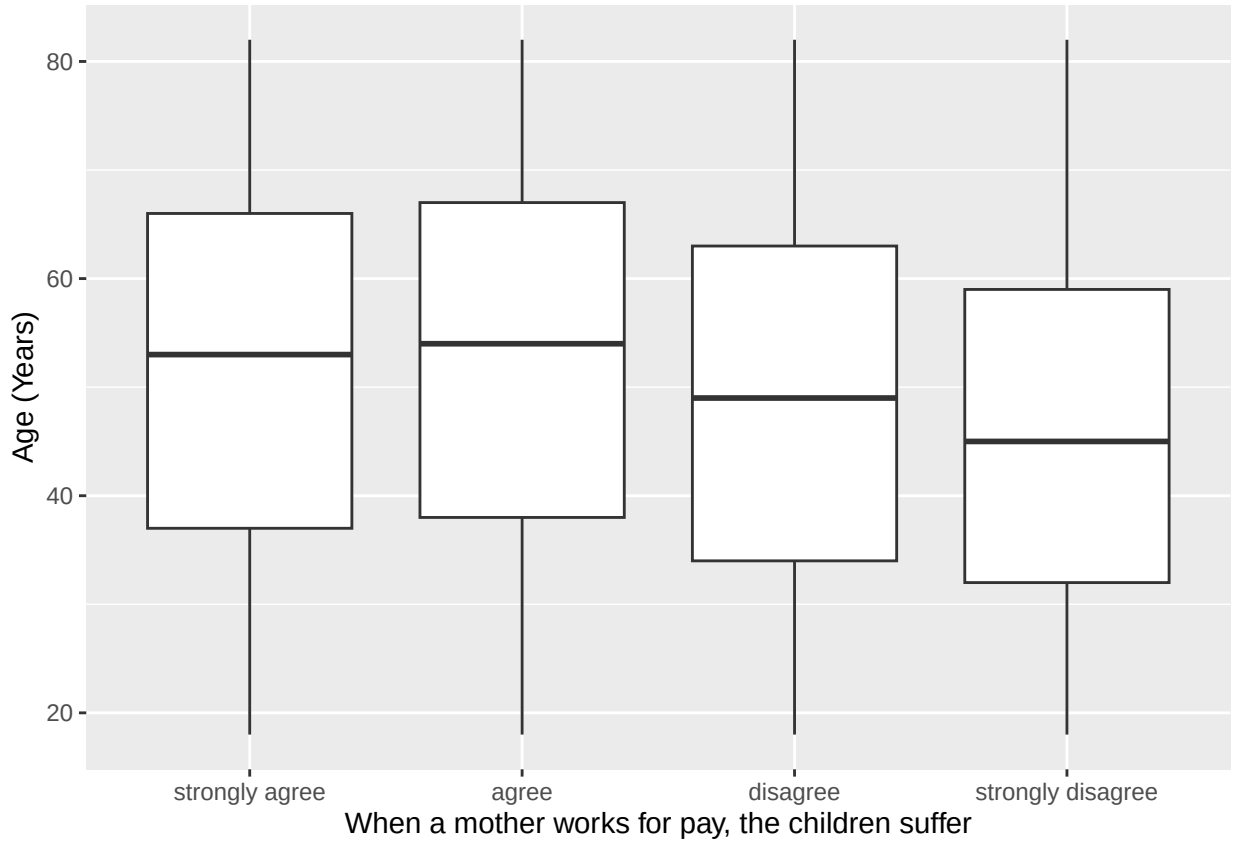
```
library(memisc)
library(pander)
pander(summary(EVS[,-which(names(EVS)=="country")]))
```

v72	v80	sex	age	education
Min. :1.000	Min. :1.000	Min. :1.000	Min. :18.00	Min. :1.000
1st Qu.:2.000	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:35.00	1st Qu.:2.000
Median :3.000	Median :2.000	Median :2.000	Median :50.00	Median :2.000
Mean :2.713	Mean :2.313	Mean :1.558	Mean :49.57	Mean :2.139
3rd Qu.:3.000	3rd Qu.:3.000	3rd Qu.:2.000	3rd Qu.:64.00	3rd Qu.:3.000
Max. :4.000	Max. :5.000	Max. :2.000	Max. :82.00	Max. :3.000

Graphs

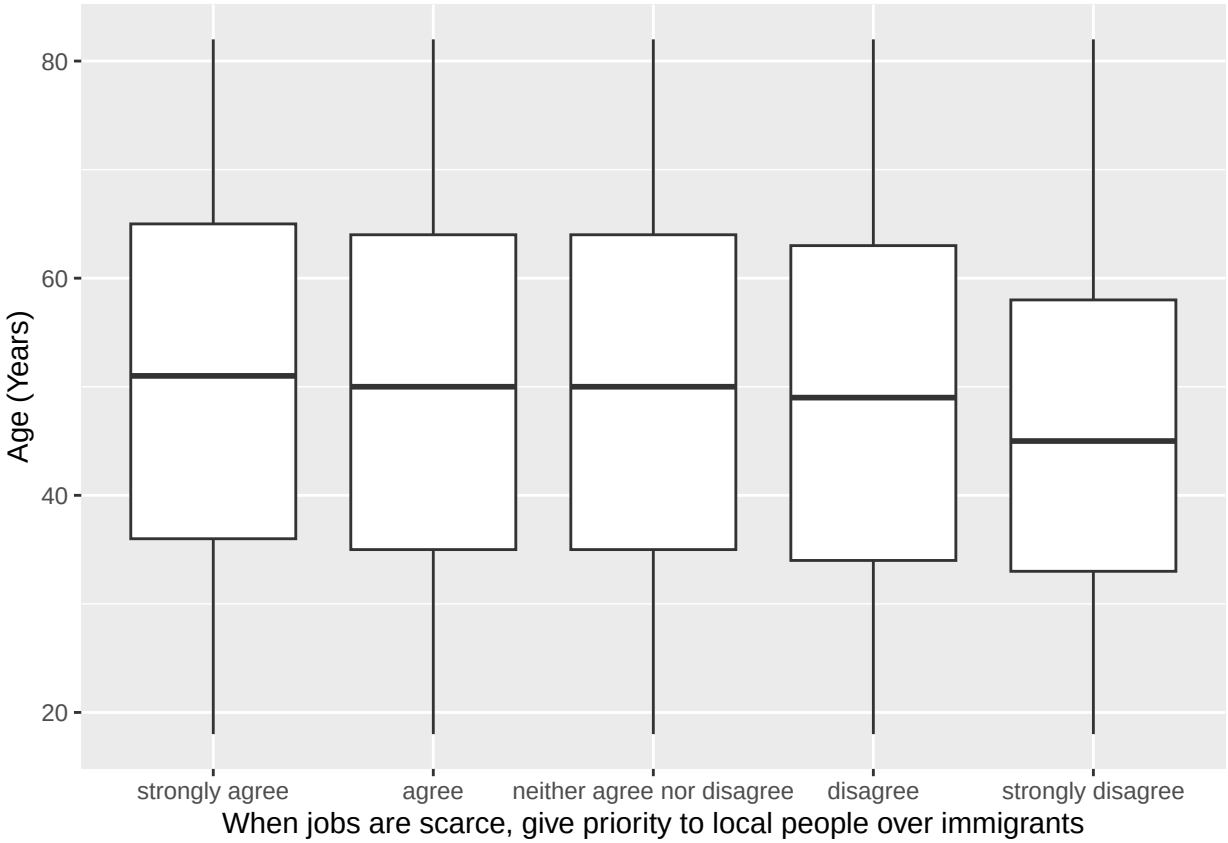
```
library(ggplot2)

ggplot(EVS, aes(as.factor(v72), age)) +
  geom_boxplot() +
  labs(x = "When a mother works for pay, the children suffer", y = "Age (Years)") +
  scale_x_discrete(labels = c("strongly agree", "agree", "disagree", "strongly disagree"))
```



We can see that the distributions of age among categories of opinion are quite similar.

```
ggplot(EVS, aes(as.factor(v80), age)) +
  geom_boxplot() +
  labs(x = "When jobs are scarce, give priority to local people over immigrants",
       y = "Age (Years)") +
  scale_x_discrete(labels = c("strongly agree", "agree", "neither agree nor disagree",
                              "disagree", "strongly disagree"))
```



Same as the previous plot, we can see that the distributions of age among categories of opinion are quite similar.

Regression models

```
model_v72 = lm(v72 ~ age + sqrt(age) + sex + as.factor(education), data = EVS)
model_v80 = lm(v80 ~ age + sqrt(age) + sex + as.factor(education), data = EVS)

pander(summary(model_v72))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2.664	0.09775	27.25	1.897e-162
age	-0.004774	0.002203	-2.167	0.03023
sqrt(age)	-0.001149	0.02976	-0.03861	0.9692
sex	0.06448	0.007257	8.886	6.537e-19
as.factor(education)2	0.1233	0.009852	12.51	7.325e-36
as.factor(education)3	0.4012	0.01046	38.36	7.574e-318

Table 3: Fitting linear model: v72 ~ age + sqrt(age) + sex + as.factor(education)

Observations	Residual Std. Error	R^2	Adjusted R^2
56755	0.8576	0.04769	0.04761

```
pander(summary(model_v80))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2.375	0.1434	16.56	1.878e-61
age	-0.003823	0.003232	-1.183	0.2369
sqrt(age)	0.006788	0.04367	0.1554	0.8765
sex	-0.03151	0.01065	-2.959	0.003084
as.factor(education)2	-0.03504	0.01446	-2.424	0.01536
as.factor(education)3	0.4238	0.01535	27.61	9.812e-167

Table 5: Fitting linear model: v80 ~ age + sqrt(age) + sex + as.factor(education)

Observations	Residual Std. Error	R^2	Adjusted R^2
56755	1.258	0.03124	0.03115