

Exercises 13.27

Import data

First, Set the environment and import data. *Note that I wrote the code using R*

```
rm(list=ls())
library(dplyr) #dataframe management
library(haven) #library for importing data set
library(psych) #library for trace matrix
ajr <- read_dta("AJR2001.dta")
```

(a) Re-estimate the model estimated in part (j) by efficient GMM. Use the 2SLS estimates as the firststep for the weight matrix and then calculate the GMM estimator using this weight matrix without further iteration. Report the estimates and standard errors.

```
Y <- as.matrix(ajr$loggdgdp)
X <- as.matrix(ajr %>%
  select(risk) %>%
  bind_cols(data.frame(con=rep(1,nrow(ajr)))))
X.endo <- X[,1]
X.exo <- X[,2]
Z <- as.matrix(ajr %>%
  select(logmort0) %>%
  mutate(logmort0sq = logmort0^2) %>%
  bind_cols(data.frame(con=rep(1,nrow(ajr)))))

b2sls <- solve(t(X)%*%Z)%*%solve(t(Z)%*%Z)%*%t(Z)%*%X)%*%t(X)%*%Z)%*%
  solve(t(Z)%*%Z)%*%t(Z)%*%Y
e_hat2sls <- Y - X %*% b2sls

w <- Z*(e_hat2sls)%*%matrix(1,1,ncol(Z))
O <- t(w)%*%w

bgmm <- solve(t(X)%*%Z)%*%solve(O)%*%t(Z)%*%X)%*%(t(X)%*%Z)%*%solve(O)%*%t(Z)%*%Y)
vgmm <- solve((t(X)%*%Z)%*%solve(O)%*%(t(Z)%*%X))
segmm <- sqrt(diag(vgmm))
```

Table 1: Coefficient estimates and S.E.

	Coefficient	Standard errors
risk	0.73	0.10
con	3.34	0.65

(b) Calculate and report the J statistic for overidentification.

- Construct Hypothesis

$$H_0 : E[g_i(\beta_0)] = 0$$

with size of 0.05.

- Calculate J statistic.

```
J <- t(t(Z)%%Y-t(Z)%%X%%bgmm)%%solve(0)%%(t(Z)%%Y-t(Z)%%X%%bgmm)
J
```

```
      [,1]
[1,] 3.74788
```

Since $J(\hat{\beta}) \xrightarrow{d} \chi^2(1)$. We can calculate p-value with following code.

```
p <- 1 - pchisq(J,1)
p
```

```
      [,1]
[1,] 0.05287455
```

(c) Compare the GMM and 2SLS estimates. Discuss your findings.

- Comparing estimates

The estimates are not that different from 2SLS as shown in the tables. The coefficient on risk is differ by 0.04.

- Comparing Overidentification test

The null hypothesis can be rewritten, $H_0 : E[Ze] = 0$ which is same null hypothesis of problem 12.22 (k). In this problem's results, We cannot reject the null hypothesis at significance level of 0.05. But in the problem 12.22, we can reject the null hypothesis at significance level of 0.05. The results saying opposite results. The key difference in these two problems is the assumption on the variance. In the problem 12.22 (k), We assumed homoskedastic error. And in the problem of GMM we were not assuming homoskedastic which is strong assumption. Thus results from GMM is more reliable in my opinion.