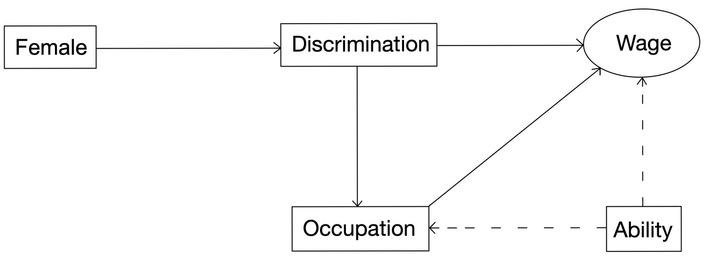
Universität Duisburg-Essen
Lehrstuhl für Ökonometrie
Prof. Dr. Christoph Hanck
M.Sc. Karolina Gliszczynska
Wintersemester 2022/2023

Methoden der Ökonometrie - Übung 5

Aufgabe 1



a) Diskutieren Sie die obige DAG Abbildung. Gehen Sie insbesondere auf die Auswirkungen der Diskriminierung (`Discrimination`) auf den Lohn (`Wage`) ein. Ist dieser Graph sinnvoll zu interpretieren?

b) Führen Sie die folgenden Schritte in `R` aus:

1. Erzeugen Sie 10000 binäre Zufallszahlen und weisen Sie diese der Variable `Female` zu.
2. Erzeugen Sie 10000 $N(0,1)$-verteilte Zufallszahlen und weisen Sie diese der Variable `Ability` zu.
3. Erzeugen Sie die Vektoren `Discrimination`, `Occupation`, `Wage` wie folgt:
```
Discrimination <- Female
Occupation <- 1 + 2*ability + 0*female - 2*discrimination + rnorm(10000)
Wage <- 1 - 1*discrimination + 1*occupation + 2*ability + rnorm(10000)
```
4. Führen Sie insgesamt drei Regressionen durch. Regressieren Sie zunächst den Lohn (`Wage`) auf das Geschlecht (`Female`). Im zweiten Schritt den Lohn (`Wage`) auf Geschlecht (`Female`) und Beruf (`Occupation`). Und zum Schluss den Lohn (`Wage`) auf Geschlecht (`Female`), Beruf (`Occupation`) und Leistung (`Ability`), jeweils mit Konstante. Was fällt Ihnen auf? Wie passt dieses Ergebnis zu den Resultaten aus a)?

Aufgabe 2

Gehen Sie von folgenden DGP (Daten generierenden Prozess) aus

$$y_t = 3 + 4X_{t1} + 3X_{t2} + u_t,$$

für $t = 1, \ldots, 100$, $u_t \overset{i.i.d.}{\sim} N(0,1)$ und $X_{t1},\ X_{t2} \sim \mathcal{U}(0,10)$.

In unserer Analyse nehmen wir fäschlicherweise das Modell

$$y_t = c + \beta_1 X_{t1} + \epsilon_t$$

an.

a) Ist der OLS-Schätzer für β_1 durch das falsch spezifizierte Modell verzerrt?
   Führen Sie eine Simulationsstudie durch.

#Lösung 1. Teil
beta1_hat <- numeric(length = 100) #Leerer Vektor um geschätzte Parameter zu speichern

for(i in 1:100){ #Daten Simulieren
x1 <- runif(100, min = 0, max = 10) x2 <- runif(100, min = 0, max = 10) eps <- rnorm(100) y <- 3 + 4 * x1 + 3 * x2 + eps

#Model Schätzen mod <- lm(y ~ x1) #Beta1 aus lm Objekt abgreifen und im Ergebnisvektor speichern beta1_hat[i] <- mod$coefficients[2] }

#Mittelwert über alle geschätzte Betas bilden mean(beta1_hat) #Der Wert ist sehr nahe am wahren Erwartungswert. #Es ist also plausibel anzunehmen, dass der Schätzer #in diesem Fall unverzerrt ist.

```

b) Wiederholen Sie die Simulation mit dem Unterschied, dass $X_{t2}$ nicht mehr $\mathcal{U}(0,10)$-verteilt ist, sondern wie folgt erzeugt wird:

$$X_{t2} = 0.2X_{t1} + \epsilon_t \text{ , mit } \epsilon_t \sim \mathcal{U}(0,10).$$

   Ändert sich das Ergebnis?

## Aufgabe 3

1. Zeigen Sie, dass $\boldsymbol{PX} = \boldsymbol{X}$.
2. Zeigen Sie, dass $\boldsymbol{MX} = \boldsymbol{O}$, wobei $\boldsymbol{O}$ die Nullmatrix ist.
3. Zeigen Sie, dass $\hat{\boldsymbol{u}} = \boldsymbol{My} = \boldsymbol{Mu}$.
4. Zeigen Sie, dass $SSR = \boldsymbol{u}^T\boldsymbol{Mu}$.
5. Sind $\boldsymbol{P}$ und $\boldsymbol{M}$ invertierbar?
6. Zeigen Sie, dass $\boldsymbol{PM} = \boldsymbol{O}$, wobei $\boldsymbol{O}$ die Nullmatrix ist.
7. Zeigen Sie, dass die Vektoren $\boldsymbol{Pv}$ und $\boldsymbol{Mw}$ orthogonal zueinander sind, für beliebige Vektoren $\boldsymbol{v}, \boldsymbol{w}$.
```