

Aufgabenblatt 2

Überblick

Heteroskedastie liegt vor, wenn die Fehlervarianz von den Regressoren abhängt. Das hat keinen Einfluss auf die Unverzerrtheit und Konsistenz des OLS Schätzers, jedoch sind die t - und F -Tests sowie Konfidenzintervalle der OLS Schätzung und das Gauss-Markov Theorem dann nicht mehr gültig. Mit dem so genannten White-Test kann die Alternativhypothese getestet werden, dass Heteroskedastie vorliegt. Wenn das der Fall ist, haben wir drei verschiedene Möglichkeiten, damit umzugehen: Variablentransformation (zum Beispiel Logarithmierung), Nutzung von Heteroskedastie-robusten Schätzern oder der (F)GLS Schätzung.

Paneldaten haben zusätzlich zur Querschnittsdimension eine Zeitdimension und erfassen Daten der selben Individuen über mehrere Zeitpunkte. Die *gepoolten Querschnittsdaten* sind eine Variante davon, wobei wir dort nicht davon ausgehen, die gleichen Individuen erneut zu befragen. Der *FD-Schätzer* (FD steht für First Difference) ist der einfachste Paneldatenschätzer.

Aufgaben

1. Bitte schätzen Sie das Modell

$$\text{cigs} = \beta_0 + \beta_1 \log(\text{income}) + \beta_2 \log(\text{cigpric}) + \beta_3 \text{educ} + \beta_4 \text{age} + \beta_5 \text{age}^2 + \beta_6 \text{restaurn} + u \quad (1)$$

mit dem Datensatz `wooldridge::smoke` und interpretieren Sie die Koeffizienten im Sachzusammenhang. Welche Koeffizienten sind signifikant zu dem Niveau $\alpha = 5\%$?

2. Liegt in Modell (1) Heteroskedastie vor?
3. Verwenden Sie statt OLS den FGLS-Schätzer um (1) zu schätzen. Haben sich die Ergebnisse verändert?
4. In der Vorlesung haben Sie den FD-Schätzer anhand von Kriminalitätsdaten aus den USA kennengelernt. Bitte beschreiben Sie dessen Vorteile gegenüber der normalen OLS Schätzung.
5. Angenommen, wir möchten die Auswirkungen mehrerer Variablen auf das jährliche Sparen schätzen und wir verfügen über Paneldaten von Einzelpersonen, die heute und vor genau zwei Jahr erhoben wurden. Wenn wir eine Jahres-Dummy-Variable für 2023 einbeziehen und die Methode der ersten Differenzen verwenden, können wir dann auch das Alter der Person im ursprünglichen Modell aufnehmen?

6. Schätzen Sie mit dem FD-Schätzer das Modell

$$\log \mathbf{rent}_{it} = \beta_0 + \delta_0 \mathbf{y90}_t + \beta_1 \log \mathbf{pop}_{it} + \beta_2 \log \mathbf{avginc}_{it} + \beta_3 \mathbf{pctstu}_{it} + a_i + u_{it} \quad (2)$$

mit dem Datensatz `wooldridge:rental` und interpretieren Sie die zu dem Niveau $\alpha = 5\%$ signifikanten Koeffizienten im Sachzusammenhang.