

Empirische Ökonomie für Fortgeschrittene

Übungsblatt 7: Paneldaten

In dieser Übung schätzen wir eine *Lohngleichung*. Dabei ist der auf dem Arbeitsmarkt erzielte Lohn (typischer Weise der Stundenlohn) die abhängige Variable; sie wird in der Regel logarithmisch spezifiziert. Zu den erklärenden Variablen gehört vor allem das Humankapital, welches aber nicht direkt gemessen werden kann. Statt dessen nimmt man dafür z. B. die Dauer der Ausbildung und/oder die Berufserfahrung als erklärende Variablen auf.

Datensatzbeschreibung

Um die Determinanten der Lohnhöhe einer Person zu untersuchen, verwenden wir den Datensatz `wagepanel.dta`. Dieser Datensatz aus den Vereinigten Staaten enthält Angaben von 545 jungen Männern, die in den Jahren 1980–1987 durchgehend beschäftigt waren. Der Datensatz ist ein Auszug aus dem wesentlich umfangreicheren *National Longitudinal Survey*, der vom *U.S. Department of Labor* erhoben wird.

Variablenliste

<code>nr</code>	person identifier
<code>year</code>	year
<code>lwage</code>	log of the wage rate
<code>educ</code>	education (years of schooling)
<code>exper</code>	job market experience (years)
<code>black</code>	dummy = 1 if person is black
<code>hisp</code>	dummy = 1 if person is hispanic
<code>married</code>	dummy = 1 if person is married
<code>union</code>	dummy = 1 if person is a union member
<code>lwage_m</code>	individueller Mittelwert von <code>lwage</code>
<code>exper_m</code>	individueller Mittelwert von <code>exper</code>
<code>married_m</code>	individueller Mittelwert von <code>married</code>
<code>union_m</code>	individueller Mittelwert von <code>union</code>

Aufgaben

1. Um was für eine Art von Paneldatensatz handelt es sich?
2. Regressieren Sie zunächst die Lohn auf die Dummy Variable für den Familienstand (`married`). Wie ist das Ergebnis zu interpretieren. Was für ökonomische Erklärungen könnte es für dieses Ergebnis geben?
3. Schätzen Sie die Lohngleichung nun mit den zusätzlichen Variablen für Arbeitserfahrung (`exper`), der Mitgliedschaft in einer Gewerkschaft (`union`), des Ausbildungsnive-

aus (`educ`) und der beiden ethnischen Variablen (`black` und `hisp`). Interpretieren Sie die geschätzten Effekte.

4. Um den Einfluss des individuellen, nicht-beobachtbaren Talents auf den Lohn zu berücksichtigen, führen Sie nun zusätzlich „person fixed effects“ ein. Schätzen Sie die Regression mit Hilfe der Within-Transformation. Welche Rückschlüsse können Sie aus dieser Regression über die möglichen Erklärungen für den Zusammenhang zwischen Familienstand und Lohn ziehen?
5. Schätzen Sie nun das gleiche Modell wie in der vorherigen Aufgabe mit der Dummy-Variablen-Regression.
Hinweis: Mit `tabulate nr, gen(nr_)` erhalten Sie für jedes Individuum eine Dummy-Variable.
6. Führen sie nun zusätzliche noch Zeiteffekte „time effects“ in die Regression ein. Warum lässt STATA immer zwei der Zeit Dummies aus der Regression aus? Welchen Zweck können solche „time effects“ erfüllen?