

Übungen – Einheit 5:

Aufgabe 5.1:

Seit langer Zeit wird diskutiert, dass die Mieten auf dem Wohnungsmarkt zu hoch sind. Gerade Haushalte mit geringem Einkommen zahlen zu hohe Mieten. Im SOEP geben rund 2.360 Personen mit validen Informationen zum Haushaltseinkommen (und Miete) an in Miete zu wohnen. Verwenden Sie die Variablen *hhinc* (Einkommen des Haushaltes) und *rent* (Miete) im SOEP und evaluieren Sie den Zusammenhang zwischen monatlicher Mietzahlung und Haushaltseinkommen.

- a) Stellen Sie ein lineares Regressionsmodell auf. Benennen Sie alle Variablen in Ihrem Modell
Unser einfaches lineares Regressionsmodell kann wie folgt dargestellt werden

$$rent_i = \alpha + \beta hhinc_i + u_i$$

für Haushalte $i = 1, 2, \dots, n$. Der Parameter α stellt dabei die Konstante in unserer Regression dar. Der Parameter β misst den Einfluss von Haushaltseinkommen auf die Mietzahlungen dar. Schlussendlich, haben wir noch den (unbeobachtet) Störterm u_t . Der Störterm fäßt alle Informationen zusammen, die nicht durch die Variable *hhinc* abgebildet werden.

- b) Stellen Sie den Zusammenhang zwischen Mietzahlungen und Haushaltseinkommen anhand eines Streudiagrammes dar. Diskutieren Sie (auch unter Verwenden Ihres Diagramms) ob die Annahmen von Linearität und konstanten Parametern in Ihrem Modell erfüllt sind.

Im Allgemeinen lässt unser Streudiagramm nicht darauf schließen, dass die Annahme der Linearität verletzt ist. Es gibt auch keine Anzeichen darauf, dass ein möglicher Strukturbruch in den Daten vorliegt. In anderen Worten, unsere Annahme der konstanten Parameter ist sehr wahrscheinlich erfüllt.

Allerdings lässt unser Streudiagramm auf einige sogenannte Ausreiser schließen. Manche Haushalte mit relativ geringem Einkommen zahlen eine relativ hohe Miete. In manchen Fällen kann dies dazu führen, dass unsere Ergebnisse von diesen einzelnen Ausreiser getrieben werden. Um dies zu vermeiden, ist es oft ratsam, Variablen zu transformieren, z.B. Verwendung von Logs anstatt von Levels

Aufgabe 5.2:

Verwenden Sie Ihr Modell aus Aufgabe 5.1

- a) Schätzen Sie diese mit den Methoden der kleinsten Quadrate. (Hinweis: Sie müssen alle Zellen mit fehlenden Informationen löschen, damit Sie die Regression durchführen können)
- b) Interpretieren Sie Ihr Ergebnis. Können Sie anhand Ihrer Ergebnisse die These unterstützen, dass Haushalte mit geringerem Einkommen tatsächlich mehr Miete zahlen?

Hinweis: Für eine einfachere Interpretierbarkeit der Ergebnisse, drücken Sie das Haushaltseinkommen in 1'000 aus (d.h. teilen Sie die Variable *hhinc* durch 1.000 bevor Sie Ihr Modell schätzen).

Laut unserem linearen Regressionmodell finden wir einen positiven Zusammenhang zwischen Haushaltseinkommen und Mietzahlungen. Unser geschätzter Koeffizient für Haushaltseinkommen impliziert, dass Haushalte mit 1.000 Euro mehr an jährlichem Einkommen rund 6.90 Euro/Monat mehr an monatlicher Miete zahlen. Wir können also nicht die aufgestellte These unterstützen.

Allerdings kann man von den Ergebnissen auch den Schluss ziehen, dass der Anteil am Einkommen, der für Miete ausgegeben wird, mit zunehmenden Einkommen sinkt (siehe Excel-Dokument).

Anmerkung: Unser geschätzter Koeffizient lässt auf den ersten Blick auf keinen großen Effekt schließen. Wir vergleichen hier aber monatliche Mietzahlung und jährliches Einkommen. Wenn wir annehmen, dass der Haushalt über das Jahr hinweg eine konstant Miete zahlt, so geben Haushalte mit 1000 Euro mehr Einkommen im Jahr rund 83 Euro mehr an Miete aus – oder knapp 8% des zusätzlichen Einkommens ($690 \cdot 12 / 1000$).

Aufgabe 5.3:

Damit wir unsere Ergebnisse für Politikentscheidungen verwenden können, sollte unser Modell eine klare Ursachen-Wirkung Beziehung darstellen. Diskutieren Sie, ob solch eine Beziehung in Ihrem Modell vorliegt. Anhand dieser Diskussion, wie verändert sich Ihre Antwort zu Aufgabe 5.2.b?

Es ist anzuzweifeln, dass unser Modell eine klare Ursachen-Wirkung Beziehung darstellt. Es ist möglich, dass Haushalte mit einem höheren Einkommen in dynamischen und urbanen Regionen wohnen. Zur gleichen Zeit sind aber Mieten in solchen Regionen auch höher. Denken Sie hier, zum Beispiel, an Silicon Valley und San Francisco.

Aufgabe 5.4:

Neue Ergebnisse im Bereich der Umfrageforschung zeigen, dass ärmere Haushalte dazu tendieren, ein zu hohes Haushaltseinkommen bei Befragungen anzugeben. Im Vergleich dazu tendieren reichere Haushalte dazu, ein niedrigeres Einkommen anzugeben.

- a) Diskutieren Sie, welche Auswirkung solche Messfehler in den Daten auf Ihre Ergebnisse hat. Wenn wir nun systematische Fehlangaben in unseren Daten haben, führt dies dazu, dass die Annahme das der Störterm einen Erwartungswert von 0 hat verletzt ist. In anderen Worten, unser geschätztes Modell ist falsch. Auf der einen Seite führt dies dazu, dass wir für ärmere Haushalte den Einfluss von Haushaltseinkommen auf Miete überschätzen (kleine Einkommen werden überschätzt, daher ist der Zusammenhang zu steil). Auf der anderen Seite überschätzen wir den Einfluss von Haushaltseinkommen auf Mietzahlungen für reichere Haushalte (auch hier wäre der Anstieg durch den Messfehler zu steil).
- b) Nehmen Sie an, die Forschung zeigt auch, dass Messfehler nur bei Haushalten mit einem monatlichen Einkommen unter 10.000 Euro oder über 60.000 Euro auftreten. Können Sie ein (intuitives) Verfahren vorschlagen, welches es erlauben würde, die Wichtigkeit solch eines Messfehlers zu evaluieren?

Ein ad-hoc Verfahren wäre es, alle Haushalte mit weniger als 10.000 Euro oder mehr als 60.000 Euro aus unseren Daten zu entfernen und unser Modell nochmals erneut zu schätzen. Wir können dann die geschätzten Koeffizienten mit denen unseres „vollen“ Modells vergleichen

Wenn wir solch ein simples Verfahren durchführen, finden wir einen etwas stärkeren Effekt von Einkommen auf Mietzahlungen. Man kann daraus schließen, dass das Überschätzen kleiner bzw. das Unterschätzen großer Einkommen für unsere Analyse kein sehr großes Problem darstellen sollte.