

Masterprüfung

Fach: Panel- und Evaluationsverfahren

Prüfer: Prof. Regina T. Riphahn, Ph.D.

Name, Vorname	
Matrikelnr.	
Studiengang	
Semester	
Datum	
Raum	
E-Mail	
Unterschrift	

Vorbemerkungen:

Anzahl der Aufgaben: Die Klausur besteht aus 4 Aufgaben, die alle bearbeitet werden müssen.

Bewertung: Es können maximal 60 Punkte erworben werden. Die Punktzahl ist für jede Aufgabe in Klammern angegeben. Sie entspricht der für die Aufgabe empfohlenen Bearbeitungszeit in Minuten.

Erlaubte Hilfsmittel:

- Taschenrechner
- Fremdwörterbuch

Wichtige Hinweise:

- Sollte es vorkommen, dass die statistischen Tabellen, die dieser Klausur beiliegen, den gesuchten exakten Wert der Freiheitsgrade nicht ausweisen, machen Sie dies kenntlich und verwenden Sie den nächstgelegenen Wert.
- Sollte es vorkommen, dass bei einer Berechnung eine erforderliche Information fehlt, machen Sie dies kenntlich und treffen Sie für den fehlenden Wert eine plausible Annahme.

Aufgabe 1:**[20 Punkte]**

Sie möchten für einen Datensatz mit 10084 arbeitenden deutschen Männern aus dem Jahr 2007 eine Arbeitsangebotsfunktion schätzen:

$$hours_i = \beta_0 + \beta_1 wage_i + \beta_2 wealth_i + u_i$$

mit i als Personenindikator.

hours = Zahl der durchschnittlich gearbeiteten Stunden pro Woche

wage = Nettostundenlohn, gemessen in €

wealth = Vermögen, gemessen in €

- a) Erklären Sie anhand eines Beispiels, warum die Variable *wealth* endogen sein könnte und in welche Richtung der Koeffizient β_2 in Ihrem Beispiel verzerrt wäre. (3 Punkte)
- b) Sie möchten *wealth* mit den folgenden beiden Variablen instrumentieren:
 - erbschaft* = 1, wenn Person im letzten Jahr eine Erbschaft > 5000 € gemacht hat, sonst 0.
 - part_ek* = Einkommen des Ehepartners in € für unverheiratete Männer 0.
 - i. Diskutieren Sie, ob die Variable *part_ek* ein geeignetes Instrument für *wealth* ist. (4 Punkte)
 - ii. Sie verwenden beide Instrumente. Stellen Sie die für die Instrumentierung relevanten Momentenbedingungen auf. (2 Punkte)
- c) Sie möchten die gemeinsame Gültigkeit von *erbschaft* und *part_ek* als Instrumente mit Hilfe eines Overidentifying Restriction Tests (Sargan Test) beurteilen.
 - i. Welche Hypothese wird dabei getestet? (1 Punkt)
 - ii. Stellen Sie die zur Ermittlung der Teststatistik nötige Hilfsregression auf. (3 Punkte)
- d) Nach dem Test entschließen Sie sich, nur noch die Variable *erbschaft* als Instrument für *wealth* zu verwenden. Erklären Sie allgemein, sowie am Beispiel dieser IV-Schätzung, was man unter einem *Local Average Treatment Effect* versteht. (3 Punkte)
- e) Ihr Kollege meint, dass Sie es bei *erbschaft* mit einem schwachen Instrument zu tun haben.
 - i. Was bedeutet diese Aussage? (1 Punkt)
 - ii. Was wäre die Konsequenz? (1 Punkt)
 - iii. Wie ließe sich die Stärke von *erbschaft* als Instrument für *wealth* überprüfen? (2 Punkte)

Aufgabe 2:**[15 Punkte]**

Ihnen stehen Daten zu 5889 Personen zur Verfügung, die 2007 arbeitslos waren. Einige dieser Arbeitslosen haben während ihrer Arbeitslosigkeit an IT-Fortbildungskursen teilgenommen. Sie möchten nun prüfen, ob sich die Teilnahme an diesen Kursen positiv auf die Beschäftigung der Teilnehmer im Jahr 2008 ausgewirkt hat. Sie führen ein Radius-Caliper-Matching mit Zurücklegen durch ($\text{caliper} = 0,001$).

Folgende Variablen stehen Ihnen dabei zur Verfügung:

teilnahme = 1, wenn Person an der Fortbildungsmaßnahme teilgenommen hat, sonst 0.

job = 1, wenn Person 2008 nicht mehr arbeitslos ist, sonst 0.

ausbildung = Schulausbildung, gemessen in Jahren.

bau = 1, wenn Person vor 2007 zuletzt in der Baubranche gearbeitet hat, sonst 0.

- a) Was versteht man im Rahmen eines Matchingverfahrens unter einem Propensity Score und was ist seine Funktion? (2 Punkte)
- b) Bei Durchführung des Matchings erhalten Sie bezüglich der Mittelwerte von *job* folgende Ergebnisse:

	Treatmentgruppe (teilnahme = 1)	Kontrollgruppe (teilnahme = 0)
Unmatched Sample	0.7	0.44
Matched Sample	0.7	0.58

Berechnen und interpretieren Sie den durch das Matching-Verfahren vorhergesagten ATT.

(2 Punkte)

- c) Sie stellen fest, dass wegen des gewählten Matching-Verfahrens 23 Personen, die an der IT-Fortbildung teilgenommen haben, aus der Stichprobe fallen. Erklären Sie, welche Rolle dabei der gewählte Caliper spielt. (3 Punkte)
- d) Erklären Sie, warum sich der Mittelwert von *job* in der Kontrollgruppe im gematchten und ungematchten Sample unterscheiden. (2 Punkte)
- e) Erhalten Sie durch das in dieser Aufgabe verwendete Matchingverfahren tatsächlich den ATT? Nennen Sie zwei Annahmen, die hier verletzt sein könnten und begründen Sie warum. (4 Punkte)
- f) Nennen Sie einen Vorteil von Matchingverfahren gegenüber einem linearen Regressionsmodell. (2 Punkte)

Aufgabe 3:**[10 Punkte]**

Sie möchten mit Daten von 250 verschiedenen Unternehmen aus den Jahren 2001-2010 (balanced panel) folgendes Modell schätzen:

$$m_ausg_{it} = \beta_0 + \beta_1 m_ausg_{it-1} + \beta_2 umsatz_{it-1} + a_i + \varepsilon_{it}$$

i ist ein Unternehmensindikator t ist ein Zeitindikator.

m_ausg = Ausgaben für Marketingmaßnahmen, gemessen in €

$umsatz$ = Umsatz des Unternehmens, gemessen in €

- Erklären Sie, warum es nicht möglich ist, das gegebene Modell mit einem Fixed Effects Schätzer konsistent zu schätzen. (2 Punkte)
- Erklären Sie kurz die bei einer Anderson-Hsiao Schätzung nötige Datentransformation und Instrumentierung. (2 Punkte)
- Wie viele Beobachtungen könnte eine Anderson-Hsiao Schätzung im vorliegenden Beispiel höchstens verwenden? Wie viele Beobachtungen könnte eine Arellano-Bond Schätzung höchstens verwenden? Begründen Sie. (4 Punkte)
- Welche Konsequenz hat das Vorliegen von Autokorrelation in ε_{it} für die Anderson-Hsiao Schätzung (kurze Begründung)? (2 Punkte)

Aufgabe 4:**[15 Punkte]**

Sie möchten den Effekt der Schulbildung auf die Löhne bestimmen und nutzen dazu Daten aus Rheinland-Pfalz. In Rheinland-Pfalz wurde zum Schuljahr 1967/68 die Schulpflicht von 8 auf 9 Jahre erhöht. Schülerinnen und Schüler, die im Sommer 1967 oder später das 8. Schuljahr vollendeten, mussten 9 statt 8 Jahre zur Schule gehen.

In Ihrem Datensatz befinden sich Personen, die zwischen 1950 und 1956 geboren wurden und in Rheinland-Pfalz ihre Schulzeit absolviert haben.

- Wie können Sie im gegebenen Beispiel den kausalen Effekt der Schulbildung auf die Löhne identifizieren? Benennen Sie das gewählte Schätzverfahren und erläutern Sie Ihre Vorgehensweise. (4 Punkte)
- Was versteht man im Zusammenhang mit Regression Discontinuity Design unter einem Sharp Design, was unter einem Fuzzy Design? Welches Design liegt im gegebenen Beispiel vor (kurze Begründung)? (3 Punkte)
- Stellen Sie, nach geeigneter Definition der nötigen Variablen, den RDD Schätzer dar. (5 Punkte)
- Nennen Sie drei mögliche Nachteile des RDD Verfahrens. (3 Punkte)