

## Masterprüfung SS 2016

**Fach:** Panel- und Evaluationsverfahren

**Semester:** Sommersemester 2016

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

### Vorbemerkungen:

**Anzahl der Aufgaben:** Die Klausur besteht aus 4 Aufgaben, die alle bearbeitet werden müssen.  
**Es wird nur der Lösungsbogen eingesammelt.**

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

**Erlaubte Hilfsmittel:**

- Tabellen der statistischen Verteilungen (sind der Klausur beigelegt)
- Taschenrechner
- Fremdwörterbuch

**Wichtige Hinweise:**

- Sollte es vorkommen, dass die statistischen Tabellen, die dieser Klausur beiliegen, den gesuchten 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 (18 Punkte)

Sie interessieren sich für den kausalen Zusammenhang zwischen Bildungsniveau und Gesundheit. Ihnen liegt ein Querschnittsdatensatz über 646 Zwillingsspärchen vor.

Sie verwenden den Body-Mass-Index, der ein Maß für die Gesundheit ist, als abhängige Variable (*BMI*). Als erklärende Variablen nehmen Sie die Schulbildung (*educ*, in Jahren gemessen), das Geschlecht (*female*, =1 für Frau, 0 sonst) und einen Vektor mit weiteren Kontrollvariablen (*X*) auf. Sie stellen folgendes Modell auf:

$$BMI_i = \beta_0 + \beta_1 educ_i + \beta_2 female_i + \gamma X_i + u_i$$

- 1.1 Sie schätzen dieses Modell als gepoolten KQ-Schätzer. Welches Problem kann sich für die Effizienz der Schätzung unter der Annahme homoskedastischer Standardfehler ergeben? (2 Punkte)
- 1.2 Sie entscheiden sich dafür, den Zusammenhang mit einem Geschwister Fixed Effects Schätzer zu bestimmen. Stellen Sie die Schätzgleichungen für eine Least Squares Dummy Variablen (LSDV) Schätzung und eine Within Schätzung auf. Definieren Sie neue Variablen und Indizes. Wie unterscheiden sich die geschätzten Parameter zwischen den beiden Schätzverfahren? (7 Punkte)
- 1.3 Welche Effekte werden im vorliegenden Fall durch die Within-Transformation eliminiert und welche nicht? Erläutern Sie und geben Sie jeweils ein Beispiel an. (4 Punkte)
- 1.4 Alternativ zum Within Schätzer könnten Sie auch den Hausman-Taylor Schätzer verwenden. Beschreiben Sie das Vorgehen dieses Schätzers im allgemeinen Fall. (5 Punkte)

### Aufgabe 2 (16 Punkte)

Im australischen Bundesstaat New South Wales beginnt am 1. März die Sommerzeit, d.h. die Uhren werden eine Stunde vorgestellt. Im angrenzenden Bundesstaat Victoria wird die Uhrzeit nicht umgestellt. Sie verfügen über eine Stichprobe der Bevölkerung in beiden Bundesstaaten mit Befragungsdaten aus dem Jahr 2013. Die Daten enthalten Informationen zum genauen Interviewdatum, dem Wohnort (Bundesstaat), der Lebenszufriedenheit (auf einer Skala von 0-10, 0=sehr unzufrieden, 10=sehr zufrieden), und weitere sozio-ökonomische Merkmale.

- 2.1 Sie schätzen den kausalen Effekt der Zeitumstellung auf die Lebenszufriedenheit mit einem Difference-in-differences (DiD) Schätzer. Beschreiben Sie Ihr Vorgehen an diesem Beispiel. (3 Punkte)
- 2.2 Welche zentrale Annahme trifft das Difference-in-differences Verfahrens zur Identifikation des kausalen Effekts der Zeitumstellung? Geben Sie ein Beispiel, in dem die Annahme verletzt ist. (2 Punkte)
- 2.3 Sie schätzen den kausalen Effekt der Zeitumstellung auf die Lebenszufriedenheit mit einem Regression-Discontinuity Design (RD). Beschreiben Sie Ihr Vorgehen an diesem Beispiel. (3 Punkte)
- 2.4 Welche zentrale Annahme trifft das Regression-Discontinuity Verfahren zur Identifikation des kausalen Effekts der Zeitumstellung? Geben Sie ein Beispiel, in dem die Annahme verletzt ist. (3 Punkte)
- 2.5 Was versteht man im Zusammenhang mit Regression Discontinuity Design unter einem Sharp Design, was unter einem Fuzzy Design? Welches Design würden Sie im gegebenen Beispiel wählen? Begründen Sie Ihre Antwort kurz. (3 Punkte)
- 2.6 Welche Treatmenteffekte identifizieren Sie mit den DiD und RD Schätzverfahren? (2 Punkte)

### Aufgabe 3 (14 Punkte)

Eine Studie untersucht, wie sich kindliche Entwicklungsverzögerungen auf die mütterliche Erwerbstätigkeit auswirken. Die Studie verwendet Informationen zu 10,510 Müttern im Alter von 25-35 Jahren nach der ersten Geburt. Die abhängige Variable ist eine Indikator-Variable, die den Wert 1 annimmt, wenn die Frau erwerbstätig ist. Die zentrale erklärende Variable ist eine Indikator-Variable, die den Wert 1 annimmt, wenn das Kind der Frau Entwicklungsverzögerungen aufweist. Als weitere erklärende Variablen werden das Alter, die Schulbildung und die Arbeitsmarkterfahrung der Frau berücksichtigt.

- 3.1 Welche der Gauss-Markov-Annahmen könnte bei einer KQ-Schätzung verletzt sein und Inkonsistenz des KQ-Schätzers hervorrufen? Erläutern Sie die Annahme und einen möglichen Grund für die inhaltliche Verletzung an einem Beispiel. (3 Punkte)
- 3.2 Sie sollen nun eine Instrumentalvariablen (IV)-Schätzung durchführen. Als Instrument für die Entwicklungsverzögerung des Kindes verwenden Sie eine Indikatorvariable, die den Wert 1 annimmt, wenn die Mutter während der Schwangerschaft stark geraucht (mehr als 5 Zigaretten am Tag) hat. Nennen Sie die Kriterien für ein gutes Instrument und diskutieren Sie, ob Sie das verwendete Instrument für geeignet halten. (4 Punkte)
- 3.3 Wer sind im gegebenen Beispiel für die IV Schätzung “Complier”, “Always-Taker”, und “Never-Taker”? Kann es alle drei Gruppen im Datensatz geben? Erläutern Sie Ihre Antworten kurz. (4 Punkte)
- 3.4 Sie sorgen sich um die Güte des Instruments. Nennen Sie eine Konsequenz von schwachen Instrumenten. Auf der ersten Stufe schätzen Sie einen Koeffizienten von 0,300 für den Parameter der Instrumentalvariable. Unterstellen Sie für den relevanten Hypothesentest einen p-Wert von 0,01. Haben Sie in diesem Fall ein schwaches Instrument? Begründen Sie Ihre Antwort. (3 Punkte)

### Aufgabe 4 (12 Punkte)

In einem Forschungsprojekt untersuchen Sie, ob das Geburtsgewicht eines Kindes mit der Teilnahme an der Vorsorgeuntersuchung der Mutter korreliert. Sie führen dazu eine Umfrage durch. In einem Pretest erhalten Sie Informationen zu acht Neugeborenen.

Geburtsgewicht (in kg)	3,32	3,61	2,97	3,80	4,27	3,49	3,95	3,64
Vorsorge (1=regelmäßige Teilnahme, 0=sonst)	0	1	0	0	1	1	1	1

Sie schätzen folgendes Modell mit einer Quantilsregression:

$$\text{Geburtsgewicht}_i = \beta_0 + \beta_1 \text{Vorsorge}_i + \varepsilon_i$$

- 4.1 Berechnen Sie die geschätzten Koeffizienten für  $\beta_0$  und  $\beta_1$  für das 75. Perzentil. (4 Punkte)
- 4.2 In der Hauptbefragung erheben Sie Informationen zu 1200 Neugeborenen. Sie schätzen das Modell erneut und nehmen noch einige Kontrollvariablen auf. Sie erhalten folgende Regressionskoeffizienten für die Variable *Vorsorge*, die alle statistisch signifikant sind:

q=0,05	q=0,10	q=0,50	q=0,90	q=0,95
0,34	0,29	0,21	0,09	0,02

Interpretieren Sie den geschätzten Koeffizienten für q=0,10. (2 Punkte)

- 4.3 Statt einer 10%-Quantilsregression hätten Sie auch eine KQ-Schätzung durchführen können. Beschreiben Sie verbal das Minimierungsproblem der beiden Schätzer und wie Vorhersagefehler gewichtet werden. (4 Punkte)
- 4.4 Ist eine Vorsorgeuntersuchung effektiver für Kinder mit hohem Geburtsgewicht (siehe Tabelle in 4.2)? Begründen Sie. (2 Punkte)