

## Kausale Effektschätzung - Propensity Score Methods II

### Aufgaben

Der Datensatz `training_dat.RData` (Objektname: `dat`) enthält simulierte Daten welche sich an die Studie von Pohl et al. (2009) anlehnen (Achtung: Effekte können abweichen). Für eine Englisch-Trainingsgruppe (`group == 1`) dient eine Mathe- Trainingsgruppe als Kontrollgruppe (`group == 0`). Die abhängige Variable ist das Ergebnis eines Englisch-Tests (`eng_post`).

1. Legen Sie zur Replizierbarkeit Ihrer Ergebnisse den Random Seed auf 12345 fest mit folgendem Befehl: `set.seed(12345)`.
2. Erstellen Sie mit `matchit()` ein PS-Matching-Modell mit additiven Effekten (keine Interaktionen oder Polynome) mit den Kovariaten Geschlecht (`sex`), Vorliebe für Mathe (`matVL`) und dem Englisch-Vortest (`eng_pre`). Verwenden Sie dabei den Nearest Neighbor Approach und ansonsten die Standardeinstellungen.
3. Extrahieren Sie den Datensatz der gematchten Personen, erstellen Sie eine PS-Logit Variable und zentrieren Sie diese am Gesamtmittelwert.
4. Berechnen Sie den ATE und ATT auf Basis des gematchten Datensatzes ohne zusätzliche Kontrolle von PS oder Kovariaten (*t*-Test).
5. Berechnen Sie den ATE und den ATT mithilfe einer generalisierten ANCOVA und dem PS-logit als Kovariate.
6. Berechnen Sie für den vollständigen Datensatz die Gewichte zur Schätzung des ATE durch Propensity-Score Gewichtung.  
Hinweis: Die PS für den vollständigen Datensatz erhalten Sie entweder über eine Stratifizierung (und dann Ignorieren der `subclass` Variable) oder über die Verwendung von `glm()` und `fitted()` (wie in Sitzung 3). Nutzen Sie dabei dasselbe Zuordnungsmodell wie in Aufgabe 1.
7. Ermitteln Sie den geschätzten Gesamt-ATE unter Berücksichtigung der unter 5) erstellten Gewichte.