

Katrin Hohmeyer: **Effectiveness of One-Euro-Jobs: Do programme characteristics matter?** (Korreferat)

Inhalt:

- Evaluation von 1-Euro-Jobs, die zwischen Februar und April 2005 begonnen wurden
- Beobachtungszeitraum: 28 Monate (Beschäftigungseffekte)
bzw. 32 Monate (andere Effekte)
- Einsatz von Matching (Propensity Score und Radius Matching)
- Ergebnisse:
 - Mittelfristig positive, aber geringe Effekte von 1-Euro-Jobs auf die Beschäftigungswahrscheinlichkeit von Frauen und Männern in West, Frauen in Ost
 - Negative Effekte auf kumulierte EWT-Dauer für Männer und Frauen in Ost, Männer in West (verschwindet auch nicht bei Betrachtung unterschiedlicher Maßnahmedauer und unterschiedlicher Arbeitsstundenanzahl)
 - Effekte auf ALGII-Bezug negativ

Allgemein Einschätzung:

- klare Struktur, interessantes Thema, einzelne Schritte informativ und gut nachvollziehbar beschrieben

Ziel: Evaluation 1-Euro-Jobs

Schwierigkeit: kein greifbares Erfolgskriterium (“soziale Integration schwer vermittelbarer Arbeitsloser” bzw. Test der Arbeitswilligkeit)

Lösung hier:

- Wahrscheinlichkeit und kumulierte Dauer einer Erwerbstätigkeit 28 Monate nach Programmbeginn,

(ok, aber sehr strenges Kriterium; trifft auch nur einen Teil des Ziels)
- Wahrscheinlichkeit der Teilnahme an weiteren ALMPs, anderen 1-Euro-Jobs,

(theoretisch ok, aber betrachtet man praktischen Einsatz von 1-Euro-Jobs, ist das oft nicht der erste Schritt, sondern eine Maßnahme unter vielen; dann kein Zusammenhang zwischen 1-Euro-Job und anderen Maßnahmen herstellbar)

- evtl. Veränderungen im Bezug von ALGII

(bestes Kriterium dafür, denn hier werden verschiedene Effekte „aufgefangen“; evtl. noch ausbauen: Höhe des ALGII einbeziehen, um vermuteten Bezug von „ergänzender Sozialhilfe“ zu bestätigen)

Einzelne Kritikpunkte:

Theorie und Ergebnisse früherer Studien

- theoretische Überlegungen und Ergebnisse früherer Studien stärker auf den Untersuchungsgegenstand bzw. die hier analysierten Gruppen fokussieren, Hypothesen klar formulieren

Evaluationsmethode

- Definition der Nichtteilnehmergruppe problematisch: nicht wirklich „Teilnahme gegenüber Warten“, sondern „Teilnahme gegenüber Warten oder Teilnahme an anderen Programmen“; Effekt des 1-Euro-Jobs verwischt, weil Alternativzustand nicht deutlich abgrenzbar
- Einhaltung SUTVA für Vergleich verschiedener 1-Euro-Jobs untereinander unproblematisch, aber evtl. Teilnahme vs. Nichtteilnahme, weil hier Möglichkeiten vielfältiger sind als Warten (kann ausgeräumt werden, wenn Nichtteilnehmergruppe strikter abgegrenzt wird)
- Behandlung unbeobachtbarer Heterogenität fehlt; essentielle Annahme von Matching: Heterogenitäten sind komplett beobachtbar und können berücksichtigt werden

(ist mit diesen detaillierten Daten sicher möglich, es sollte aber genauer drauf eingegangen werden)

Implementation

- Common Support Condition Voraussetzung für Matching, sollte vor der Beschreibung des Matchingprozesses behandelt werden

Einhaltung Common Support für Leser nicht nachprüfbar anhand der grafischen Darstellung; Kombination mit Minimum-Maximum-Vergleich (Caliendo/Kopeinig 2008) sinnvoll (keine Lücken in der Dichtefunktion, dann reicht das aus; sonst Trimming (Smith/Todd 2005) besser)

- Diskussion der verwendeten Variablen fehlt (Begründung, warum die wichtig sind für soziale Integration bzw. Arbeitsmarktintegration UND Teilnahme an 1-Euro-Jobs)
- Verwendung Propensity Score unproblematisch bei den vorhandenen Gruppengrößen,
- Test verschiedener Zuordnungsalgorithmen entspricht ebenfalls Standard,

- Qualität der Matchingergebnisse sollte an dieser Stelle behandelt werden;
- Qualitätskriterien überzeugen nicht!

Mean standardised absolute bias:

- aggregiertes Maß;
interessant ist aber Angleichung der Verteilung jeder einzelnen Variable;
- Maß für metrische Variablen,
in Tabelle 4 aber viele nominale und ordinale Variablen enthalten;
- Keine gesicherte Aussage über zulässige Größe des Bias nach Matching verfügbar

Pseudo-R²:

- ebenfalls aggregiert und statistisch gesicherte Aussage nicht möglich (Faustregeln gibt's auch nicht)

t-Tests:

- nur für metrische Variablen aussagefähig (Normalverteilung unterstellt)
- Besser: *skalenspezifische Tests* für Angleichung der Verteilung jeder einzelnen Variablen;
z.B. McNemartest (dichotome Vars) ,
 χ^2 -Homogenitätstest (nominale Vars),
Vorzeichen-Rangtest von Wilcoxon (metrische Vars)

(nicht in Stata verfügbar, aber einige nichtparametrische Tests in SPSS)

Ergebnisse

- viele Ergebnisse präsentiert, evtl. Beschränkung auf Kernaussagen besser