

Protokoll

**„Can statistical models increase the efficiency at the
employment office?“**

IFAU-Workshop in Uppsala, 3. Dezember 2004

Der Workshop wurde vom schwedischen Institut für Arbeitsmarktevaluierung (IFAU) veranstaltet, mit dem Ziel, einen Dialog zwischen WissenschaftlerInnen und arbeitsmarktpolitische PraktikerInnen anzuregen. Am Vormittag standen drei wissenschaftliche Vorträge am Programm, am Nachmittag eine innerschwedische Diskussion zum Stand der, offensichtlich noch nicht weit fortgeschrittenen Überlegungen zur Verwendung von Profiling-Methoden in Schweden.

Eingangs wurden von Erik Melander (IFAU) die zentralen Ideen in Bezug auf Profiling zusammengefasst: I) Diese Methode soll als „*decision support*“ dienen (durch effiziente Verwendung der vorhandenen Informationen und durch bessere Selektion); II) die Effizienz erhöhen, d.h. die Nettoerträge der Maßnahmen maximieren und III) die Chancengleichheit erhöhen. Die gesamte Profiling-Debatte steht dabei im Zusammenhang mit einer verwandten Frage, nämlich jener, ob die Nettoerträge maximiert werden (Effizienzkriterium) oder die bedürftigsten Personen („*equity goal*“) gefördert werden sollen. Nur wenn die Heterogenität der Maßnahmenwirkungen derart ist, dass beispielsweise die größten Wirkungen bei den Personengruppen auftreten, die die höchste Wahrscheinlichkeit aufweisen, langzeitarbeitslos zu werden, dann fallen die beiden Zielsetzungen zusammen. Bei homogenen Maßnahmenwirkungen und homogenen Grenzkosten können grundsätzlich keine Effizienzvorteile lukriert werden.¹ In Berger, Black, Smith (2001) ist der grundlegenden Sachverhalt folgendermaßen zusammengefasst: „... *the optimal profiling variable is the one that maximises the attainment of the goals of the allocation mechanism*“ (p.65)². Und weiter schreiben die Autoren: „*A profiling system (or any other system of assignment to services) might do a good job of allocating an ineffective service or it might do a bad job of allocating an effective service*“ (p.62). D.h. die beiden Fragen nach der Prognose der Langzeitarbeitslosigkeits-Wahrscheinlichkeit und nach den Maßnahmenwirkungen sind grundsätzlich voneinander verschieden, wenn auch alle Profiling-Anstrengungen darauf abzielen, die „*richtigen*“ Personen in Maßnahmen zu bekommen - also jene Personen, bei denen sich die Bezugsdauer von ALG am stärksten durch die Maßnahmenteilnahme verkürzt und/oder bei denen der Maßnahmeneffekt (etwa ein hoher Beschäftigungsanteil nach Teilnahme) am größten ist.

Diese grundsätzlichen Fragestellungen tauchten während des Workshops an mehreren Stellen immer wieder auf.

Die wichtigsten Diskussionsbeiträge des Workshops lassen sich wie folgt zusammenfassen:

¹ In diesem Fall können möglicherweise geringere Substitutions- und Verdrängungswirkungen von langzeitarbeitslosen Personen dennoch dafür sprechen, diese besonders zu fördern.

² Grundsätzlich lassen sich folgende Allokationsmechanismen unterscheiden: 1. Entscheidung des AMS-Beraters; 2. Deterministisch (etwa alle Jugendliche oder LZA erhalten eine Förderung); 3. „*First-come-first-serve*“; 4. Zufalls-Zuteilung; und 5. Profiling.

I) „Profiling Benefit Claimants in Britain“, Alex Bryson (und Diana Kasparova), Policy Study Institute, London

Alex Bryson stellte eine Machbarkeitsstudie für GB vor, bei der es nicht darum ging, die „beste“ Profiling-Methode zu ermitteln, sondern im Wesentlichen die folgenden zwei Fragen zu beantworten: Wie genau können die Profile erstellt werden? Welche Anforderungen an die Daten müssen beim Profiling erfüllt sein? Illustriert werden sollte mit der Studie auch die Sensitivität der Ergebnisse hinsichtlich einzelner Modellspezifikationen und anderer Faktoren³. Es wurden drei Ergebnis-Variablen⁴ für die drei Hauptgruppen von LeistungsbezieherInnen „Kranke/Behinderte“, „Allein erziehende Eltern“ und „AIG-BezieherInnen“ prognostiziert. 12 Pilot-Gebiete wurden 12 zufällig ausgewählten Vergleichsregionen gegenübergestellt. Die zentrale Frage der ausreichenden Genauigkeit der Prognosen - die Beantwortung dieser Frage kann lt. Autoren letztlich nur von der Politik vorgenommen werden - wurde über zwei Verfahrensschritte angegangen: a) den Vergleich der Durchschnitte einzelner Quintile im Vergleichssampel (siehe nachfolgende Tabelle 1) und b) durch die Festlegung einer Entscheidungsregel (30%, 50% oder 70% Maßnahmenteilnehmer) und die anschließende Ermittlung von „Korrektur Teilnahme“ (das Modell prognostiziert im Vergleichssampel - out of sample prediction - eine Maßnahmenteilnahme für eine Person, die tatsächlich an einer Maßnahme teilgenommen hat), „korrekter Nichtteilnahme“, „nicht korrekter Teilnahme“ und „nicht korrekter Nichtteilnahme“ (siehe nachfolgende Tabelle 2).

Ergebnisse der Machbarkeitsstudie:

- Profiling ist besser als eine Zufallszuteilung zu Maßnahmen - eine vergleichsweise niedrige Hürde! Ein Vergleich mit der Maßnahmenzuteilung von Beratern in GB konnte aus Datengründen nicht erfolgen. „*Wrong denial*“ und „*wrong treatment*“ bleiben jedoch ein gravierendes Problem.
- Die Genauigkeit der Prognose hängt von der Varianz der Outcome-Variable (höher ist besser) ab; zudem vom Anteil der Personen, die in irgendeine Maßnahme kommen; und von der Verfügbarkeit von erklärenden Variablen.
- Die verschiedenen Schätzmethode zeigen keinen großen Performance-Unterschiede.
- Modelle mit mehr erklärenden Variablen prognostizieren meist etwas genauer.
- Zur nachfolgenden Tabelle 1: Die Unterschiede in verschiedenen Bereichen der Verteilung „*mean predicted outcome*“ liegen etwa zwischen 35%-50%, die Prognosegüte ist jedoch offensichtlich nicht abhängig davon, welcher Bereich der Verteilung betrachtet wird.
- Zur nachfolgenden Tabelle 2: Wenn die Gruppe der Anspruchsberechtigten groß ist (70% treated), dann können etwa 75% richtige vorhergesagt werden. Wenn die Maßnahmenteilnahme jedoch stärker fokussiert angeboten wird, dann fällt die Prognosegüte auf etwa 58%; in diesem Fall fällt auch die Anzahl der „*wrongly*

³ Wie unterschiedliche Funktionsformen bzw. Schätzmethode (Logit, Probit, Tobit und OLS); unterschiedliche Outcome-Variablen und „*cut off points*“, d.h. angenommene Entscheidungsregel, dass 30%, 50% bzw. 70% der nach geschätzten Wahrscheinlichkeiten geordneten Personen eine Maßnahme erhalten; Modelle mit unterschiedlicher Anzahl an erklärenden Variablen (von „*parsimonious*“ bis „*full model*“, d.h. etwa 220 Variablen); Modelle für einzelnen Untergruppen von Personen.

⁴ Die individuelle Wahrscheinlichkeit, keine Beschäftigung 1 Jahr nach der Beobachtungsperiode im Juni/Juli 2000 zu haben; die Wahrscheinlichkeit des Bezuges von Transferleistungen 1 Jahr nach Beobachtungsperiode; die Wahrscheinlichkeit des Prozent-Anteils an Leistungsbezug in einer Zeitperiode von 30 Monaten nach Beobachtungsperiode.

treated⁵, aber die Anzahl der fälschlicherweise Abgelehnten steigt jedoch deutlich an.

Tabelle 1: Differences in mean predicted outcome (oow at 12m) for quintiles of predicted out of work distribution

	Logit		Probit		OLS	
	Pars	Full	Pars	Full	Pars	Full
T80%-B20%	0.41	0.54	0.40	0.53	0.35	0.44
T60%-B40%	0.39	0.48	0.38	0.47	0.34	0.41
T40%-B60%	0.37	0.42	0.37	0.42	0.36	0.42
T20%-B80%	0.34	0.36	0.34	0.37	0.39	0.45
Average difference	0.38	0.45	0.37	0.45	0.36	0.43

Tabelle 2: Success in targeting treatment, logits out of Work

	Correctly treated	Correctly denied	Wrongly treated	Wrongly denied
70% treated				
Pars	55	16	15	14
Full	57	17	14	12
50% treated				
Pars	43	22	9	26
Full	45	24	7	24
30% treated				
Pars	29	37	3	41
Full	30	29	2	39

II) “The Value of Profiling: Evidence from the JTPA?”, Jeff Smith, University of Maryland, USA

Jeff Smith wies in seinem Vortrag, der im Kern immer wieder um die oben in der Einleitung angesprochenen grundsätzlichen Themenfelder kreiste, thesenhaft auf einige zentrale Probleme beim Profiling bzw. Targeting hin⁵:

⁵ In der Diskussion scheint sich mehr und mehr der Begriff Targeting durch zu setzen, weil in den USA der Begriff Profiling auch mit dem Erstellen von sogenannten „Täter-Profilen“ durch die Polizei assoziiert wird. Dem Vortrag von Smith war nicht immer einfach zu folgen, weil er - typisch für einen amerikanischen Professor - ständige Anspielungen und Querverweise, häufig nur für Insider verständlich, verwendete, ständig witzelte und das alles ziemlich zynisch daherkam!

- In der wissenschaftlichen Literatur wird viel gearbeitet zur Frage der Wirkungen alternativer Maßnahmen, wenig jedoch zu alternativen Allokationsmechanismen (in Maßnahmen). Als Benchmark wird entweder Zufallszuteilung oder „*caseworker's discretion*“ - wenn dazu Daten vorliegen - herangezogen.
- In Bezug auf die Arbeit der AMS-Betreuer ist Smith, mit Hinweis auf die wissenschaftliche Literatur, ziemlich skeptisch: Er meint, Betreuer könnten nicht exakt genug vorhersagen, wer von einem Programm in welchem Ausmaß profitieren wird und sie hätten wenig Einfluss auf die Maßnahmenwirkungen. Weiters führte er aber auch aus: „*We still know relatively little about what caseworkers do, and what value they add. This ignorance matters as caseworkers are the primary alternative allocation mechanism within social programs*“.
- Profiling von Outcome Levels (z.B. Verbleibswahrscheinlichkeiten in AL, Beschäftigungswahrscheinlichkeiten) funktioniert lt. Smith einigermaßen zufrieden stellend, die Verbindung zu den Programm-Wirkungen lässt jedoch noch zu wünschen übrig. Dies hängt auch mit dem Umstand der unsicheren Korrelation zwischen Performanz-Maß und Programm-Wirkung zusammen. Können die Wirkungen vorhergesagt werden, dann lässt sich die Effizienz verbessern. Die genaue Vorhersage von Outcomes hingegen erleichtert meist die Erreichung von Zielen der Chancengleichheit.
- Für Smith kristallisieren sich folgende „*Key issues*“ heraus: Was ist die Absicht, die mit dem Targeting/Profiling verfolgt wird? Was ist die optimale Targeting-Variable? Ist die Varianz in der Targeting-Variable hoch? Kann die Varianz mit den beobachtbaren „*covariates*“ erklärt werden? Welche Qualität besitzen die Daten? Ist der Zusammenhang zwischen Outcomes/Wirkungen und erklärenden Variablen stabil/strukturell? Welcher Anteil an der Zielpopulation des Targeting nimmt tatsächlich an einer Maßnahme teil? Wieviel kostet das Targeting-System im Vergleich zu anderen Allokationsmechanismen?
- Profiling über den Konjunkturzyklus: die Vorhersagegenauigkeit von Profiling-Modellen nimmt im Abschwung zu, weil die Heterogenität der von AL Betroffenen (d.h. die Varianz ihrer Eigenschaften) zunimmt; daher: Profiling-Modelle sollten von Zeit zu Zeit erneut geschätzt werden!
- Die Verwendung von OLS bzw. von stetigen Outcome-Variable ist gegenüber binären Variablen (wie „*ALG-Anspruch ausgeschöpft*“) häufig von Vorteil, weil beispielsweise die Ergebnisse leichter interpretiert werden können. Zudem vernachlässigen binäre Variable Informationen, die in den Abgangsquoten von AL vor 6 Monaten (im Falle von USA) enthalten sind.
- Nach der Euphorie in Bezug auf das Targeting in der 2. Hälfte der 90er Jahre kommt Smith zu folgendem, leicht skeptischen Resümee: „*Caseworkers do a bad job, profiling helps only a little*“. Er sagt aber auch: „*Statistical treatment rules are a potentially very useful tool*“.

III) “Statistically Assisted Program Selection: Targeting of Swiss Active Labour Market Policies”, Markus Fröhlich (und Micheal Lechner, Heidi Steiger), Schweiz

Markus Fröhlich stellte in seinem Vortrag eine Pilotstudie für die Schweiz vor, die von April-Oktober 2005 umgesetzt werden soll. Einleitend wies er darauf hin, dass eine Situation, in der die Berater die Auswahl zu den aktiven Maßnahmen vornehmen, Vor- und Nachteile habe. Zu letzteren gehört sicherlich der Umstand, dass AMS-Berater in der Regel nur über eingeschränkte Informationen zu den Erwerbs-

karrieren ihrer Klienten nach Beendigung der Arbeitslosigkeitsepisode haben und dass ihre Erwartungen bezüglich der „*potential outcomes*“ aufgrund einer niedrigen Anzahl von „*Beobachtungen*“ häufig ziemlich volatil sind (daher sollten die Berater-Entscheidungen durch statistische Hilfsmittel unterstützt werden).

Im Pilotversuch wird folgendermaßen vorgegangen: 16 regionale AMS-Geschäftsstellen (von 120 RAVs insgesamt) nehmen teil, 50% der zufällig ausgewählten Berater innerhalb eines RAVs werden durch statistische Prognosen der zukünftigen Erwerbskarrieren ihrer KundInnen in ihrer Arbeit unterstützt. Am Ende soll überprüft werden, ob sich die Arbeitsergebnisse der von SAPS-Instrumenten unterstützten Berater signifikant von den nicht unterstützten unterscheiden.

Neben methodischen Fragen ging Fröhlich in seinem Vortrag noch auf eine Simulation von alternativen Maßnahmenallokationen, wie sie in Fröhlich, Lechner, Staiger (2003) und Lechner, Smith (2003) publiziert sind, ein: Wie hätten sich die Beschäftigungschancen gegenüber der von den Beratern praktizierten Allokation zu den Maßnahmen verändert, wenn die Arbeitslosen zu jenen Maßnahmen zugeteilt worden wären, für die die individuellen Wirkungen maximal gewesen wären - unter der Annahme, dass die ursprüngliche Verteilung auf die einzelnen Maßnahmen auf Grund von Angebotsbeschränkungen konstant gehalten wird. In der nachfolgenden Tabelle 3 (siehe Fröhlich, Lechner, Staiger 2003, p.18) werden die Ergebnisse der Re-allokation mit der aktuellen Allokation und mit einer Zufallszuteilung verglichen. Beispielsweise steigt der Beschäftigungsanteil nach 12 Monaten von 49,8% (bzw. 49,5% bei der Zufallszuteilung) auf 57,7% bei optimaler Allokation an. Dieses Ergebnis gibt einen ersten Eindruck, welche Wirkungen durch die Verwendung von SAPS im Prinzip zu erwarten sind.

Tabelle 3: Outcomes after simulated reallocation

	(1) Actual allocation	Allocation by potential outcome			(3) Random allocation
		(2a)	(2b)	2(c)	
Outcomes					
Employed after 7 months (%)	38,9	46,6	43,0	40,0	38,3
Employed after 12 months (%)	49,8	55,1	57,7	53,0	49,5
Employed after 17 months (%)	62,4	65,2	66,4	67,9	61,9
Earnings after 7 months (CHF)	1280	1510	1410	1310	1250
Earnings after 12 months (CHF)	1580	1730	1800	1680	1560
Earnings after 17 months (CHF)	1960	2060	2110	2150	1950
Earnings after 7 months if employed	3290	3240	3280	3280	3260
Earnings after 12 months if employed	3170	3140	3120	3170	3150
Earnings after 17 months if employed	3140	3160	3180	3170	3150
Per capita costs for active labour market programmes	4500	828	1160	1770	1500

Simulation of potential outcomes for all 28'130 persons. (1) Actual allocation: composition of participants in programmes actually found in the data. (2a-c) Allocation by potential outcome: optimal programme is determined by multiple comparison procedure for the computed potential outcome (2a) 'employment after 7 months', (2b) 'employment after 12 months', and (2c) 'employment after 17 months'; significance level 50 %; persons are allocated to unique best programme if available, remaining persons are assigned such as to keep original composition

IV) Resümee

Auch wenn die anfängliche Euphorie in Bezug auf die Wirkungen von Profiling bzw. Targeting etwas abgeebbt ist, scheint es doch sinnvoll, die Arbeit von Beratern durch die Verwertung zusätzlicher nützlicher Informationen zu unterstützen. Der Pilot-Versuch in der Schweiz scheint in diesem Zusammenhang besonders interessant. Wie ein Blick auf Dornmayer (1996) zeigt, könnte auch die österreichische Diskussion zum Thema von einer Aktualisierung profitieren.

Va) Behandelte Literatur

- Bryson,A., Kasparova,D. (2003), Profiling benefit claimants in Britain: A feasibility study, Report of research carried out by the National Centre for Social Research on behalf of the Home Office and the Department for Work and Pensions, Department of Work and Pensions, Research Report No 196, London.
- Fröhlich,M., Lechner, M.,Steiger,H.(2003), Statistically Assisted Programme Selection - International Experience and Potential Benefits for Switzerland, in: Swiss Journal of Economics and Statistics, 139, 311-331, Schweiz.

Vb) Weiterführende Literatur

- Berger,M.C., Black,D., Smith,J.A. (2000), Evaluating profiling as a means of allocating government services, in: Lechner, M., Pfeifer,F. (2001), Econometric Evaluation of Labour Market Policies, ZEW Economic Studies 13, Physika Verlag Heidelberg.
- Black,D., Smith,J.A., Plesca,M., Shannon,S. (2003), Profiling UI Claimants to Allocate Reemployment Services: Evidence and Recommendations for States, Final Report, March 2003.
- Dornmayr,H. (1996), Projektive Analyse der individuellen Dauer von Arbeitslosigkeit, Teil I: Textband, Teil II: Datenband, IBE, Studie im Auftrag des AMS, Linz, März 1996.
- Eberts,R.W., O'Leary,Ch.J., Wandner,St.A. (2002), Targeting Employment Services, W.E.Upjohn Institute for Employment, 2002.
- Lechner,M., Smith,J.A. (2003), What is the value added by caseworkers? in: IZA Discussion Papers 728, Februar 2003.
- OECD (2002), The ins and outs of long-term unemployment, Chapter "*profiling*" p. 215, in: Employment Outlook, Chapter IV, Paris July 2002.
- PLS RAMBOL (2001), Study on Early Identification of High Risk Unemployed, Draft Final Report, Study commissioned by DG EMPL, Brussels, February 2001.
- Rudolf,H., Müntnich,M. (2001), Profiling zur Vermeidung von Langzeitarbeitslosigkeit, in: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, 34. Jg/2001.