	WIA Adult		
	Overall	No Training	Training
Sample size			
Unique individuals	72,934	61,141	11,793
WIA entries, or quarters of comparison program participation	74,858	62,579	12,279
Demographic	Mean	Mean	Mean
Male	0.431	0.448	0.346
Black	0.522	0.541	0.426
Hispanic	0.008	0.006	0.018
Age	32.46	32.58	31.87
Years of education	12.26	12.21	12.50
Employment			
Employment-employment	0.302	0.296	0.334
Employment-not employed	0.198	0.192	0.226
Not employed-employed	0.329	0.333	0.313
Not employed-not employed	0.170	0.179	0.127
Earnings second year prior	8,232	8,011	9,414
Earnings in prior year	8,088	7,985	8,614
Earnings following year	9,076	8,846	10,248
Earnings second year after	10,223	9,627	13,258
Program Experience			
WIA in prior two years	0.061	0.062	0.057
Comparison program participation in prior two years	0.181	0.166	0.258

DISCUSSION PAPER SERIES

IZA DP No. 4569

New Estimates of Public Employment and Training Program Net Impacts: A Nonexperimental Evaluation of the Workforce Investment Act Program

Carolyn J. Heinrich
Peter R. Mueser
Kenneth R. Troske
Kyung-Seong Jeon
Daver C. Kahvecioglu

November 2009

Example II:

Sensitivity of matching-based program evaluations to the availability of control variables

Michael Lechner and Conny Wunsch
SEW, University of St. Gallen

Roma, April 2011

Paper available as discussion paper

Swiss Institute for Empirical Economic Research
University of St.Gallen

Introduction (1)

- Many papers use matching methods that rely on *very informative data* (VID) to remove selection bias
- VID are costly (data collection, data manipulation)
- No consensus what VID means in practice

Introduction (3)

- ***If*** researchers can influence the data collection process (successfully) ***matching*** may be the method of choice
- If so, which control variables are required for a 'sufficiently' unbiased matching estimate of the effect of German (European?) labour market programs?
- **How sensitive are the results to the omission / inclusion of particular groups of variables?**

Literature on required control variables for m…

> Treated and controls from same labour market
  - Friedlander, Robins (1995), Heckman and Smith (1999)

> Recent unemployment history & transitions between employm., UE, OLF
  - Heckman, Ichimura, Smith, Todd (1998), Heckman, LaLonde, Smith (1999), Heckman, Smith (1999), Dolton, Smith (2010)

> Pre-treatment outcome (incl. some of the prev. variables)
  - Mueser, Troske, Gorislavsky (2007)

> All variables measured in the same way for treated and controls
  - Heckman, Ichimura, Todd (1997)

> Each papers focuses on particular variables, none has all of them (omitted variable bias)

Literature on required control variables … (2)

> Which are the methods used by these papers to judge whether these factors play a role?

1) *Upps-they-are-different* method

   Include specific variables and find that results change

   (problematic if some other groups of variables are left out)

2) Experimental estimate as benchmark

Alternative: Experiments as benchmark

> Exam.: LaLonde ('86), Dehejia, Wahba ('99), Smith & Todd ('05)

> Advantage: If experiment went well, unbiased estimate of truth

> Disadvantages

- Comparison seems only possible in the US
  - Lack of relevant experiments in Europe
  - But US programs, assignment rules, and UI insurance are hardly comparable to the European setting
- Available experiments may be very old and no longer relevant
- Experimental groups too small
  - Lots of sampling noise ➔ lack of power
- **No sufficiently informative observational data base available that could be matched to the experimental data**

Strategy of this paper to improve on this literature

> Large and very informative data which contains not only a particular dimension, but all relevant dimensions (improved German administrative data)

> Typical program types with 'standard' selection (*Training and job search assistance*)

> Common outcomes (employment, earnings)

> Vary informational content (add & remove blocs of variables)

> Compare results of 'standard estimation' & use ***EMCS*** approach

> But: No true experimental benchmark to define the 'truth'

Implementation

1. Evaluation as in Lechner/Miquel/Wunsch (LMW, JEEA 2011) with new comprehensive data

    - starting from full model remove blocs of variables
    - starting from parsimonious model add blocs of variables
    - check specifications from well-published studies

2. *Simulation based on real data and full model as DGP (EMCS): same exercise as before but*

    - shut down potential unobserved heterogeneity
    - use nonparticipants: true effect is zero
    - analyze how strength of selection affects potential biases

Programs considered

> Idea: Use standard programs

> Job search assistance (JSA)

- Counselling services, referral to vacancies, availability checks, one-day trial internship, job search training

> Training (TR)

- Occupational skills training, skill upgrading, programs that combine workplace and class-room training
- Planned duration of up to 6 months

Data

> German administrative data 1990-2006

- 2% random sample of all employees subject to social insurance
- employers social insurance records 1990-2006
- register of benefit recipients 1990-2006
- jobseeker register 2000-2006
- register of program participants 2000-2006
- additional firm specific information from (100%) population

> Rich set of regional characteristics for the 439 German districts

Details

- West Germany only (different and non-standard situation in E.G.)
- Entries in UE between April 2000 and December 2002
- Eligible prime age workers
- Treatment definition as in LMW (JEEA 2011)
  - Non-participant
    - not having joined any program 1 year after begin of UE
    - still unemployed at random start date (drawn from distribution of respective comparison program)
  - Participants: Joining JSA / Training in the 1st year of UE
- Estimator as in LMW (JEEA 2011)
  - Regression adjusted radius matching (best in Huber, Lechner, Wunsch, 2010)
- Inference based on bootstrap (and simulations) with 499 replications

Sample sizes

> Gender and program-specific analysis

- different selection process
- different effects

	Women	Men
JSA participants	1452	2267
JSA nonparticipants	22067	32660
Training participants	1570	1754
Training nonparticipants	20816	30189

> Different sample size for non-participants due to random start dates

Selection into programs

> Eligibility
  - Qualify for UI or unemployment assistance

> Caseworker (according to the legal rules)
  - Jobseekers employment prospects
  - Qualification needs
  - Chances of successful completion of the program
  - Local labour market conditions

> Job seeker (rational behaviour)
  - Employment and earnings prospects with / without the program
  - Cost in terms of effort, time and lost leisure
  - Risk of benefit sanction of non-compliance
  - May extend or renew UI eligibility

Split covariates in 14 blocs (1)

No.	Bloc	Variables
0	Baseline characteristics	Age, school degree, vocational degree, nationality, number of kids, age of youngest child <6, marital status
1	Timing of entry into unemployment & program	Half-month & quarter of entry into unemployment, time to treatment, interaction terms
2	Last employment: non-firm characteristics	Skill profile, full/part-time, occupation
3	Last employment: firm characteristics	Firm age, size, closed firm, fraction females, low-income, temporary & part-time jobs, age distribution, mean tenure, fraction of jobs destroyed, industry, most frequent occupation
4	Short-term employment history (up to 2 years before unemployment)	Half-month employed/out of labour force (olf)/ in program in the 6/24 months before, no employment/unemployment in last 2 years, time since last unemployment/olf in last 2 years, unemployed/olf in month 6/24 before, number of unemployment/olf spells employer changes
5	Long-term employment history (up to 10 years before unemployment)	Half-month employed/unemployed in the last 10 years before, in program/fortnights olf in the last 4/10 years before, no unemployment/olf in last 10 years, time since last unemployment/olf in last 10 years, mean employment/unemployment/olf duration in last 10 years, number of unemployment/out of labour force/program spells/employer changes in last 10 years, difference between potential & actual labour market experience, total time in last firm
6	Earnings history	Earnings in last job, average earnings in last 10 years, sum of earnings in last year/2 years
7	Industry & occupation-specific experience	Number of occupation/industry changes, tenure in last occupation/industry, total duration in last occupation/industry

Split covariates in 14 blocs (2)

8	Pre-treatment outcomes	Employed/earnings 4 years before, cumulated employment/earnings/ UI receipt/UI benefits over 4 years before
9	Benefits & UI claim	Amount of benefit, remaining potential UI benefit duration, no UI claim
10	Compliance with benefit conditions, employability & mobility	Fully mobile within Germany, average job referrals per day, no referrals, at least one type of non-compliance with benefit conditions in past
11	Health	Has health impairments, impairments affect employability, recognised disability status, total duration reported in sick during receipt of benefits in past, did not report in sick during receipt of benefits in past
12	Characteristics of job looked for	Skill profile, full/part-time, occupation
13	Region dummies	State (*Bundesland*)
14	Detailed regional information	GDP growth 1994-2002, travel time to next big city on public transport, fraction of foreigners, unemployment rate, agglomeration area, rural area, net migration

Selectior

$$SD(x^k) = \left| \frac{Mean^{Part}(x^k) - Mean^{NonPart}(x^k)}{\sqrt{Var^{Part}(x^k) + Var^{NonPart}(x^k)}} \right| \times 100$$

Table 5.1: Descri *lations*

		Job search assistance						Training					
		Men			Women			Men			Women		
		P	NP	SD	P	NP	SD	P	NP	SD	P	NP	SD
Age in years		33	37	24	34	38	24	35	37	10	37	38	7
Schooling:	No degree	.12	.12	0	.07	.08	3	.09	.11	5	.03	.08	14
	Upper secondary degree	.20	.15	9	.29	.28	1	.20	.16	8	.34	.28	8
	University entry degree	.12	.11	1	.17	.16	2	.18	.11	15	.21	.15	11
Vocational degree:	No degree	.37	.34	4	.32	.32	0	.27	.33	10	.22	.33	17
	University/college degree	.03	.04	2	.05	.05	1	.08	.03	14	.06	.05	5
Foreign citizen		.15	.17	3	.09	.11	6	.13	.15	5	.07	.12	12
At least one child		.24	.23	1	.38	.33	9	.23	.23	0	.41	.32	13
Married		.34	.43	12	.40	.49	13	.40	.43	4	.46	.48	3
Beginning of unemployment*		37	32	18	35	32	12	30	32	7	28	31	13
Time to treatment in half-months		6.8	5.3	19	6.9	5.5	18	7.9	6.5	18	7.8	6.3	21
Remaining potential UI benefit duration in days		276	315	14	302	332	11	308	315	2	335	333	0
No vacancy referral		.16	.34	30	.17	.36	32	.18	.33	24	.22	.36	22
Any form of non-compliance with benefit conditions		.24	.19	9	.11	.10	3	.19	.19	0	.07	.10	7
Health problems (yes / no)		.17	.22	9	.15	.21	11	.16	.22	11	.14	.22	15
Looking for low- to medium-skilled job		.45	.43	3	.41	.40	2	.35	.42	10	.29	.41	18

Selection: Descriptive statistics (2)

		Job search assistance						Training					
		Men			Women			Men			Women		
		P	NP	SD	P	NP	SD	P	NP	SD	P	NP	SD
Last job:	Earnings	833	867	5	599	603	1	938	863	11	669	599	11
	Unskilled worker	.41	.37	5	.23	.21	3	.33	.37	6	.13	.22	16
	Clerk	.18	.16	4	.35	.35	0	.31	.16	27	.50	.35	21
	Firm size	269	321	2	233	270	3	232	320	4	271	269	0
	Fraction laid off by firm	.27	.25	5	.24	.23	2	.26	.24	4	.26	.23	6
2 years before:	# of unemployment spells in	.65	.78	10	.43	.58	14	.61	.80	15	.39	.59	18
	# of out-of-labour-force spells	.80	.78	1	.72	.75	2	.68	.79	8	.63	.76	11
4 years before:	Employed	.56	.56	0	.51	.54	5	.58	.56	4	.57	.54	3
	Earnings	786	910	9	564	627	6	920	900	1	669	625	4
Cumulated over 4 years before: Employment		59	60	2	59	60	2	62	60	5	61	60	4
	Earnings / 10000	52	57	9	38	40	4	61	57	7	44	40	9
	UI receipt	7.5	9.9	13	5.9	7.5	12	7.4	1.0	17	5.6	7.7	15
	UI benefits	1469	2038	16	809	1100	12	1430	2050	18	815	1122	13
Cumulated over 10 years before: # of UE spells		1.7	2.1	14	1.0	1.3	14	1.5	2.1	19	.9	1.4	15
	# of out-of-labour-force spells	2.8	2.6	6	2.4	2.3	1	2.4	2.6	6	2.0	2.3	12
	# of occupation changes	3.7	3.3	10	2.9	2.7	6	3.6	3.3	8	2.7	2.7	3
	# of industry changes	2.2	1.9	13	2.0	1.8	8	2.1	1.9	10	1.8	1.8	1

Selection: Descriptive statistics (3)

	Job search assistance						Training					
	Men			Women			Men			Women		
	P	NP	SD	P	NP	SD	P	NP	SD	P	NP	SD
Baden-Wurttemberg	.12	.12	0	.13	.14	3	.12	.11	1	.15	.14	2
Bavaria	.09	.23	28	.12	.21	17	.15	.23	14	.17	.21	7
Lower Saxony, Bremen	.17	.16	2	.15	.15	0	.19	.16	5	.15	.15	1
Schleswig-Holstein, Hamburg	.19	.07	25	.20	.08	25	.11	.07	10	.11	.07	9
Hessen	.07	.08	2	.07	.08	3	.08	.08	0	.07	.08	2
Rhineland- Palatinate, Saarland	.08	.08	1	.07	.07	1	.08	.08	0	.11	.07	9
Local unemployment rate in %	8.8	8.3	12	8.5	8.2	7	8.5	8.3	5	8.2	8.2	0
# of observations	2267	32660		1452	22067		1754	30189		1570	20816	

> Overall

- Selection is substantial but not extreme
- Positive selection for training

Selection model: The propensity score

- 165 variables in the probits
- All blocs of variables are jointly significant given the other variables of the model (Wald-tests)

- Same is true for regressions / probits of the outcome equation

Implementation (1)

> Use full specifications
  - remove single blocs
  - remove groups of blocs of related variables
> Use baseline specification
  - add blocs of variables
  - add groups of blocs of related variables
> Compare full specification with (approximations to) specifications used in the literature
  - Sianesi (2004), Mueser, Troske, Gorislavsky (2007), LaLonde (1986), Dehejia, Wahba (1999), Heckman, Smith (1999), Dolton, Smith (2010)
> **57 specifications**

Implementation (2)

> Need benchmark to compare estimates

> Option 1: Results of full model

- very realistic
- estimates are noisy ➔ conclusions become noisy
- sampling uncertainty not of much interest in our comparison, because other researchers may have larger (smaller) samples
- inference with 499 bootstrap samples

Implementation (3)

> Option 2: Placebo data

- simulate treatment among controls using estimated propensity score
- true value is zero
- procedure is repeated 500 times → measure of uncertainty

> Sometimes we have to use regressions to better understand implications of the 57 results

$$E(Bias) = \alpha + \sum_{k=1}^{14} (Bloc_k \ omitted)\beta_k + ... + error$$

Table 6.1: Regression results for the simulations

	4 years after program start		Average in year 4 after program start		Cumulated effects over the first 48 months after program start			
Bloc of variables removed from full specification	employ-ment rate in %	half-monthly earnings in EUR	months employed in%	half-monthly earnings in EUR	half-months employed	earnings in EUR	half-months on UI	benefit re-ceipt from UI in EUR
Timing of entry into unemploy-ment & program	***0.41***	***8.4***	***0.22***	***4.9***	***0.14***	***97***	***-0.24***	***-103***
Last job: Non-firm characterist.	***0.06***	***5.8***	**0.02**	***4.4***	0.01	***175***	***-0.06***	***-8***
Firm characteristics	***-0.11***	***-5.2***	***-0.31***	***-8.5***	***-0.44***	***-530***	***-0.06***	***-36***
Labour market history: 2 years	***0.15***	***-1.6***	0.02	***-4.1***	***-0.16***	***-348***	***-0.06***	***-28***
10 years	*-0.05*	***-3.7***	***-0.19***	***-6.6***	***-0.27***	***-425***	***-0.14***	***-72***
Earnings history	***0.26***	***10.5***	***0.28***	***11.3***	***0.31***	***581***	***0.11***	***69***
Industry- & occupation-specific experience	**-0.06**	***-4.3***	***-0.13***	***-5.6***	***-0.16***	***-297***	***-0.07***	***-37***
Pre-treatment outcomes	0.02	***2.0***	**-0.05**	0.6	***-0.13***	***-57***	***-0.03***	***-16***
Benefits & UI claim	***0.09***	***2.6***	***0.07***	***2.7***	-0.02	***89***	***0.07***	***51***
Compliance with benefit condit., employability & mobility	*0.04*	***-6.8***	-0.03	***-7.3***	***-0.17***	***-388***	***0.02***	***-6***
Health	***0.54***	***12.5***	***0.66***	***13.7***	***0.71***	***741***	***0.13***	***59***
Characteristics of job looked for	***0.08***	***3.0***	*0.04*	***1.6***	0.01	23	***-0.05***	***-23***
Region dummies	***-0.06***	***-6.1***	***-0.12***	***-7.7***	***-0.21***	***-439***	***-0.02***	***-35***
Detailed regional information	***-0.29***	***-5.7***	***-0.28***	***-5.6***	***-0.33***	***-285***	***-0.01***	***-4***
History: Employment	***0.08***	***2.0***	***0.09***	***2.2***	***0.07***	***101***	-0.02	-4
Unemployment	0.02	***-1.6***	***-0.13***	***-4.6***	***-0.35***	***-436***	***-0.08***	***-58***
Out-of-labour-force	***0.10***	***4.0***	***0.27***	***7.6***	***0.41***	***538***	***0.14***	***77***

Results

Regression analysis of placebo data

> Every single bloc does not lead to major bias

- although it could be 'non-small' (health, info about current UI spell)

> Omitting several blocs may lead to substantial biases

> Biases more pronounced for cumulative outcomes

- they include lock-in effect as well

Table 6.2: Bias of effects results for selected specifications obtained from simulations

	Outcome variables	4 years after program start		Average in year 4 after program start		Cumulated effects over the first 48 months after program start			
Specification of propensity score	Correlation of *p(x)* with *p(x)* of full model	employ-ment rate in %	half-monthly earnings in EUR	months employ-ed in%	half-monthly earnings in EUR	half-months em-ployed	earnings in EUR	half-months on UI	benefit receipt from UI in EUR
		Training - men							
True model	1.00	0.0	-3	-0.1	-2	-0.06	-30	0.01	1
Sianesi (2004)	0.85	***1.4***	***41***	***1.1***	***38***	***1.16***	***1845***	***-0.13***	-8
Mueser, Troske, Gorislavsky (2007)	0.62	***1.6***	***45***	***1.3***	***38***	***1.07***	***1504***	***-0.28***	***-172***
LaLonde (1986), Dehejia, Wahba (1999)	0.44	***1.2***	***38***	***0.7***	***24***	0.03	***322***	***-0.90***	***-461***
Heckman, Smith (1999)	0.55	***1.3***	***44***	***0.9***	***31***	***0.60***	***952***	***-0.63***	***-404***
Dolton, Smith (2010)	0.38	***1.8***	***90***	***1.1***	***73***	***0.68***	***2751***	***-1.14***	***-375***
Baseline with very inflexible employment, unemployment & out-of-labour-force history	0.42	***1.2***	***53***	***0.6***	***39***	0.03	***1049***	***-0.97***	***-444***
Baseline with inflexible employment, unem-ployment & out-of-labour-force history	0.48	***1.3***	***57***	***0.9***	***49***	***0.69***	***2054***	***-0.48***	***-200***
		Training - women							
True model	1.00	-0.1	-2	-0.1	-2	-0.06	-58	0.00	3
Sianesi (2004)	0.83	***0.8***	***27***	***0.7***	***27***	***0.79***	***1289***	***-0.10***	***26***
Mueser, Troske, Gorislavsky (2007)	0.68	***1.7***	***30***	***1.7***	***30***	***1.36***	***1310***	***-0.26***	***-83***
LaLonde (1986), Dehejia, Wahba (1999)	0.50	***1.9***	***23***	***1.6***	***17***	***0.76***	***237***	***-0.27***	***-126***
Heckman, Smith (1999)	0.62	***1.6***	***24***	***1.6***	***24***	***1.41***	***1109***	***-0.21***	***-156***
Dolton, Smith (2010)	0.44	***1.7***	***68***	***1.6***	***70***	***1.52***	***3297***	***-0.54***	19
Baseline with very inflexible employment, unemployment & out-of-labour-force history	0.50	***1.0***	***35***	***0.9***	***35***	***0.47***	***1474***	***-0.51***	***-105***

Table 6.2: Bias of effects results for selected specifications obtained from simulations

	Outcome variables	4 years after program start		Average in year 4 after program start		Cumulated effects over the first 48 months after program start			
Specification of propensity score	Correlation of *p(x)* with *p(x)* of full model	employment rate in %	half-monthly earnings in EUR	months employed in%	half-monthly earnings in EUR	half-months employed	earnings in EUR	half-months on UI	benefit receipt from UI in EUR
Job search assistance - men									
True model	1.00	0.1	1	0.1	1	0.09	65	0.01	-1
Sianesi (2004)	0.91	***0.4***	***5***	***0.3***	***4***	***0.33***	***225***	***-0.04***	***-30***
Mueser, Troske, Gorislavsky (2007)	0.68	***1.0***	***18***	***0.6***	***14***	***0.28***	***414***	-0.02	3
LaLonde (1986), Dehejia, Wahba (1999)	0.42	***0.4***	***-8***	***-0.3***	***-21***	***-1.06***	***-1574***	***-0.45***	***-222***
Heckman, Smith (1999)	0.62	***1.8***	***11***	***0.7***	***-8***	-0.01	***-989***	***-0.27***	***-143***
Dolton, Smith (2010)	0.54	***0.7***	***-17***	***-0.4***	***-37***	***-1.24***	***-2443***	***-0.71***	***-319***
Baseline with very inflexible employment, unemployment & out-of-labour-force history	0.35	***1.4***	-2	-0.1	***-32***	***-1.38***	***-2682***	***-0.67***	***-342***
Baseline with inflexible employment, unemployment & out-of-labour-force history	0.42	***1.7***	***10***	***0.6***	***-12***	***-0.37***	***-1280***	***-0.13***	***-76***
Job search assistance - women									
True model	1.00	-0.1	1	-0.2	0	-0.06	50	0.00	-5
Sianesi (2004)	0.89	0.1	***6***	0.0	***6***	-0.01	***264***	***0.17***	***109***
Mueser, Troske, Gorislavsky (2007)	0.64	***1.0***	***16***	***0.7***	***15***	***0.29***	***491***	-0.03	***67***
LaLonde (1986), Dehejia, Wahba (1999)	0.42	***0.3***	***-4***	0.0	***-9***	***-1.16***	***-1049***	***-0.13***	-9
Heckman, Smith (1999)	0.58	***1.2***	*3*	***0.7***	0	***0.20***	***-230***	***-0.10***	**18**
Dolton, Smith (2010)	0.52	***0.4***	***-5***	0.0	***-7***	***-0.43***	***-438***	***-0.70***	***-112***
Baseline with very inflexible employment, unemployment & out-of-labour-force history	0.40	***0.7***	***8***	***0.5***	***4***	***-0.72***	***-489***	***-0.34***	**-18**
Baseline with inflexible employment, unemployment & out-of-labour-force history	0.45	***1.1***	***15***	***0.8***	***12***	0.01	***268***	0.00	***102***

Note: *Italics*: significant on the 10% level, **bold**: significant on the 5% level, ***bold italics***: significant on the 1% level

Results: Other studies

> Training

- All studies appear to be too optimistic
- Problem most severe for cumulative outcomes
- Dolton & Smith (2010) most problematic

> Job search assistance

- Difference less severe, but they are still there

> All studies emphasizing the problem of omitted variables suffer from omitted variable bias

Conclusions

> Confirmation of what has been identified by literature as important variables (*on top of standard demographics*)

- pre-treatment outcomes (Mueser, Troske, Gorislawsky, 2007)
- transitions between labour market states (many studies)
- flexible modelling of labour market history (Dolton, Smith, 2010)

Conclusions

> Other factors

- Health
- Short ***and*** long run labour market history in many dimensions (earnings, type of job, olf, etc.)
- Information on last employer in many dimensions
- Timing of UE & program
- Information on current UE spell in general
- Information on previous employment spells in general

Conclusions

- However, given typical sample sizes, whether these factors lead to a *significantly* different effect cannot be answered in general
- **MSE:** Although samples are not small, sample uncertainty may also be an important concern
  - For our sample sizes, investing in getting larger samples is probably at least as important as investing in getting more informative data
  - This conclusion might change if selection would be stronger

Thank you

for your attention!

Michael Lechner
SEW – University of St. Gallen
April 2011

Swiss Institute for
Empirical Economic Research
University of St.Gallen