Dealing with randomisation bias in a social experiment exploiting the randomisation itself: The case of ERA

Barbara Sianesi

University of Bath

2 October 2013





Motivation

- RCTs hailed as gold standard in programme evaluation
- Identifying "no randomisation bias" assumption (Heckman, 1992 and Heckman et al., 1999)
- Unusual position to empirically assess part of it: possibility that programme participation process has changed because of RA.
- Develop framework for analysis of non-participation under CIA assumption
- Show that RA, when combined with appropriate data, can offer ways to guide nonexperimental methods in addressing the shortcoming it gave rise to

CIA made up of two parts:

- 1. Remains testable under the experiment and offers a way to correct nonexperimental estimates that fail to pass the test
- 2. Very weak (at least in our application)

 \rightarrow usefulness of judicious combination of non-experimental methods and experimental set-up in overcoming the latter's shortcoming when admin outcome data available

• Issue which motivated paper arose in the ERA experiment:

The Employment Retention and Advancement (ERA) program

ERA treatment

Offer of a package of time-limited support once in work

Eligibles

- 1) LT unemployed mandated for ND25+
- 2) Unemployed volunteering for NDLP
- 3) [LPs on WTC working PT who volunteer for ERA]

Tested

Large-scale (*N*=16,000), multi-site (6 districts) RA social experiment Intake: Nov 2003 – Jan 2005 (pilots end Oct 2007)

Non-participation in the ERA study





Issues raised by non-participation

Policymaker interested in <u>impact of offering ERA for all those eligible</u> (in the 6 districts) to receive the offer.

(ERA as integral component of the New Deal)

• But... ERA tested only on a potentially selective subset of the eligibles

Because ERA was a study and involved RA some eligibles were denied or 'refused' participation in something which in normal circumstances one could not be denied or one could not 'refuse': becoming *eligible* for financial incentives and personal advice

 \rightarrow Assess scope for **randomisation bias** in the experimental estimate for the parameter of interest

[<u>Impact of offering ERA eligibility on the study participants</u> (in the 6 districts) \rightarrow Has non-participation affected the extent of **external validity** of the experimental results, and hence their representativeness and policy relevance?]

ERA study offers rare chance to look at this issue!

- offer (*ITT*)
- whole population (*ATE*)
- admin data

Research questions

- Impact on all eligibles
 - Impact on the non-participants
- Impact on all eligibles *versus* experimental impact on participants

Contribution

Moving beyond an experiment's limitations by climbing on its shoulders

- Impact estimates under selection-on-observables (CIA)
- Follow-up data for non-participants
 - \rightarrow Exploit experiment to test CIA (under weak assumption) (plus help with choice of how to summarise X)
 - → When test fails, use information from experiment to adjust non-experimental estimates
 - → Ancillary analysis: assess claim that long and detailed labour market histories can control for most selection bias
- No follow-up data for non-participants
 - → Reweighting estimator to deal with non-participation *and* survey/item non-response
 - \rightarrow Exploit experiment for subgroups for whom
 - follow-up (admin) earnings outcomes are available
 - equivalent survey and admin earnings outcomes are available

Sample and data

		ND25			NDLP	
Eligibles	7,796	100.0%		7,261	100.0%	
 Study non-participants 	1,790	23.0%		2,209	30.4%	
 Study participants 	6,006	77.0%	100.0%	5,052	69.6%	100.0%
 with survey outcome 	1,840		30.6%	1,745		34.5%
 without survey outcome 	4,166		69.4%	3,307		65.5%

Outcomes

- 12-month follow-up
- employment (ever employed and days) admin data
- annual earnings survey data
- fiscal year 2004/05 earnings admin data

Control variables

ERA district	
Inflow month	District-specific month from random assignment start when the individual started the ND25 Gateway or volunteered for NDLP
Local conditions	Total New Deal caseload at office, share of lone parents in New Deal caseload at office, quintiles of the index of multiple deprivation, local unemployment rate
Demographics	Gender, age, ethnic minority, disability, partner (ND25+), number of children (NDLP), age of youngest child (NDLP)
Current spell	Not on benefits at inflow (NDLP), employed at inflow (indicator of very recent/current employment), time to show up (defined as the time between becoming mandatory for ND25+ and starting the Gateway or between being told about NDLP and volunteering for it), early entrant into ND25+ programme (Spent <540 days on JSA before entering ND25+)
Labour market history	Past participation in basic skills, past participation in voluntary programmes (number of previous spells on: NDLP, New Deal for Musicians, New Deal Innovation Fund, New Deal Disabled People, WBLA or Outreach), past participation in ND25+
inflow)	 Active benefit history, inactive benefit history, employment history: (1) parsimonious summary (2) monthly employment dummies (3) dummies for sequences of employment/benefits/neither states; 90% of sample (4) dummies for ever employed in 12m window at any time in the past

Methodology



Average effect on participants Average effect on non-participants Average effect on *all* eligibles

$$ATE_{1} \equiv E(Y_{1} - Y_{0} | Q=1)$$

$$ATE_{0} \equiv E(Y_{1} - Y_{0} | Q=0)$$

$$ATE \equiv E(Y_{1} - Y_{0}) = (1-p) \cdot ATE_{1} + p \cdot ATE_{0} \quad p \equiv \Pr\{Q=0\}$$

Follow-up data

 $ATE = (1-p) \cdot \{ E(Y \mid R=1) - E(Y \mid R=0) \} + p \cdot \{ E(Y_1 \mid Q=0) - E(Y \mid Q=0) \}$

Akin to getting the **ATNT** using matching methods

Assume (CIA-1)

(CIA-1) $E(Y_1 | Q=0, X) = E(Y_1 | Q=1, X)$ and (CS)

Identification

 $E(Y_1 | Q=0) = E_X[E(Y_1 | Q=0, X) | Q=0] = (CIA-1) = E_X[E(Y_1 | Q=1, X) | Q=0]$ $= (RA) = E_X[E(Y_1 | R=1, X) | Q=0] = (CS) = E_X[E(Y | R=1, X) | Q=0]$

Implementation

Match to each non-participant one or more similar programme group member(s).

Testing (CIA-1)

Under (CIA-\beta) of no selection into ERA study based on realised unobserved individual gains, once allowing for arbitrarily heterogeneous impacts based on rich set of *X*:

• validity of (CIA-1) can be tested by testing (CIA-0):

E(Y | Q=0, X) = E(Y | R=0, X) (OLS or matching)

Test can also help in choosing how to summarise X.

• if (CIA-0) – and hence (CIA-1) – fail: $E(Y_0 | Q=0, X) = E(Y_0 | Q=1, X) + \alpha(X)$ $\alpha(X) \neq 0$ $E(Y_1 | Q=0) = E[E(Y | R=1, X) | Q=0] + \underbrace{E[\alpha(X) | Q=0]}_{\checkmark}$

> Bias in the matching estimate Identified from $\alpha(X) = E(Y | Q=0, X) - E(Y | R=0, X)$

No follow-up data

 $ATE = (1-p) \cdot ATE_1 + p \cdot E(Y_1 - Y_0 | Q=0)$

Akin to **attrition**

 \rightarrow reweigh Y of the responding participants on the basis of the X of the full eligible group to make them representative – in terms of X – of the full eligible population

Motivation: Survey earnings

- all we had originally
- clean definition of components, incl. all part-time and self-employed
- same horizon (and all post-treatment)

Assume (CIA-β)

(CIA- β) $E(Y_1 - Y_0 | Q=1, X) = E(Y_1 - Y_0 | Q=0, X)$ hence $= E(Y_1 - Y_0 | X)$

Allowing for selective survey/item non-response based on X

(NR) $E(Y_1 | R=1, S=1, X) = E(Y_1 | R=1, S=0, X)$ and $E(Y_0 | R=0, S=1, X) = E(Y_0 | R=0, S=0, X)$

Identification

$$ATE = E[E(Y_1 - Y_0 | X)] = (CIA - \beta) = E[E(Y_1 - Y_0 | Q = 1, X)]$$

=(RA)= E[E(Y_1 | R=1, X)] - E[E(Y_0 | R=0, X)]
=(NR)= E[E(Y_1 | R=1, S=1, X)] - E[E(Y_0 | R=0, S=1, X)]
= E[E(Y | R=1, S=1, X)] - E[E(Y | R=0, S=1, X)]

Implementation

1) Reweighting

Weigh Y of responding participants so as to reflect distribution of X in eligible population.

$$ATE = E[\omega_1(X) \cdot S \cdot R \cdot Y - \omega_0(X) \cdot S \cdot (1 - R) \cdot Y]$$

with $\omega_k(X) \equiv \frac{P(Q=1)}{P(Q=1|x)} \frac{P_{RS|Q}(k, 1|1)}{P_{RS|Q,X}(k, 1|1, x)}$ for $k=0, 1$

2) Matching

Construct weights to realign X via matching

- exact specifications of pscore and response probabilities not needed
- can assess actual comparability



- (I) Subgroup for whom FY 2004/05 admin earnings *outcomes* are available
 - → post April group (35% of ND25+ and 41% of NDLP eligibles)
 - Is (CIA-0) test passed for admin earnings? (Plus guidance on how to construct *X*)
- (II) Subgroup for whom 'equivalent' 1st-year survey and FY 2004/05 admin earnings outcomes are available (i.e. cover same horizon)
 - \rightarrow March/May group (25% of ND25+ and NDLP eligibles)
 - Experimental impacts
 - admin earnings for March-May respondents vs admin earnings for full March-May group
 - admin earnings for March-May respondents *vs* survey earnings for March-May respondents
 - admin earnings for full March-May group vs survey earnings for March-May respondents
 - Is (CIA-0) test passed for admin earnings?
 - Compare *ATE* for full March-May group in terms of admin earnings to *ATE* for full March-May in terms of survey earnings (derived from respondents)

Findings: Employment outcomes (admin)

ND25+

	ATE.	С	CIA-0 test			
	AIL_1	OLS	Matching	θ_0		
DAYS EMPLOYED						
Raw	4.0	-9.4	***	0.834		
All other X's plus						
summary	4.6*	-7.9***	-9.7***	0.829		
monthly employment	4.8**	-7.6***	-9.4***	0.835		
ever employment	5.0**	-7.6***	-9.4***	0.835		
sequence	4.8**	-7.9***	-8.8***	0.843		
summary + monthly employment	4.8**	-7.7***	-9.2***	0.837		
summary + ever employed	5.0**	-7.7***	-9.3***	0.837		
summary + sequence	4.8**	-8.0***	-8.8***	0.843		
EVEREMPLOYED						
Raw	0.014	-0.06	2***	0.808		
All other X's plus						
summary	0.017	-0.044***	-0.056***	0.825		
monthly employment	0.017	-0.043***	-0.053***	0.831		
ever employment	0.019*	-0.042***	-0.053***	0.831		
sequence	0.017	-0.043***	-0.052***	0.835		
summary + monthly employment	0.017	-0.044***	-0.052***	0.835		
summary + ever employed	0.019*	-0.043***	-0.053***	0.831		
summary + sequence	0.017	-0.044***	-0.053***	0.833		

	n Formal Diverted 4TF					CIA-0 test						
	P	rormai	Diverieu	AIL_1	Raw	$\theta_{\rm raw}$	OLS	Matching	θ_0			
DAYS EMPL.												
All	23.0	13.6	<i>9.4</i>	4.6*	4.0	0.834	-7.9***	-9.7***	0.829			
Scotland	8.7	8.7	0.0	8.6	-17.3	0.690	-8.2	-8.3	0.828			
NE England	34.9	26.1	8.8	-10.3	-33.9***	0.565	-23.6***	-27.8***	0.616			
NW England	14.6	14.6	0.0	7.5	-7.0	0.864	-1.8	-3.0	0.937			
Wales	20.7	11.1	9.6	-13.6	-12.0	0.816	-16.3	-7.9	0.864			
East Midls	27.5	10.7	16.8	8.0	-4.3	0.934	-5.7	-7.7	0.885			
London	25.8	11.1	14.8	8.9**	-3.6	0.915	-3.7	-2.8	0.932			
no NE Eng	21.0	11.5	9.5	6.5***	-5.7*	0.894	-4.7	-5.3	0.901			
Ever empl.												
All	23.0	13.6	9.4	0.017	-0.062***	0.808	-0.044***	-0.056***	0.825			
Scotland	8.7	8.7	0.0	0.047	-0.096*	0.726	-0.039	-0.041	0.861			
NE England	34.9	26.1	8.8	-0.036	-0.191***	0.541	-0.149***	-0.172***	0.571			
NW England	14.6	14.6	0.0	0.033	-0.024	0.915	0.010	0.010	1.038			
Wales	20.7	11.1	9.6	-0.035	-0.027	0.923	-0.017	-0.004	0.987			
East Midls	27.5	10.7	16.8	0.031	-0.073**	0.817	-0.060**	-0.071**	0.819			
London	25.8	11.1	14.8	0.022	-0.017	0.929	-0.009	-0.010	0.958			
no NE,EM	18.8	11.8	7.0	0.023*	-0.041**	0.858	-0.007	-0.011	0.956			

DAYS EMPLOYED

	p	ATE_1	ATE_0	ATE	$ATE_1 \neq ATE$
All but NE England	0.210	6.5**	9.7***	7.2***	no
All districts			10.1***	5.9***	*
All districts, adjusted	0.230	4.6*	6.3*	5.0*	no
NE England			8.1	-3.9	**
NE England, adjusted	0.349	-10.3	-15.4**	-12.1*	no

EVER EMPLOYED

	р	ATE_1	ATE ₀	ATE	$ATE_1 \neq ATE$
All but NE Eng, E Midls	0.188	0.023*	0.026*	0.024**	no
All districts		-	0.056***	0.026**	***
All districts, θ -adjusted	0.230	0.017	0.007	0.015	no
NE England			0.092**	0.009	***
NE England, θ -adjusted	0.349	-0.036	-0.062**	-0.045	no
E Midlands			0.083***	0.045**	*
E Midlands, θ -adjusted	0.275	0.031	0.013	0.026	no

NDLP

	ATE ₁	OLS	CIA-0 test Matching	θ_0
DAYS EMPLOYED			—	
Raw	-0.1	2	3.8	1.033
All other X's plus				
summary	-2.2	-10.4***	-11.2**	0.914
monthly employment	-2.4	-10.2***	-10.2**	0.921
ever employment	-2.5	-11.0***	-12.1**	0.907
sequence	-2.4	-10.8***	-11.7**	0.910
summary + monthly employment	-2.7	-10.6***	-11.1**	0.915
summary + ever employed	-2.2	-10.8***	-12.4**	0.906
summary + sequence	-2.1	-10.4***	-11.2**	0.914
EVER EMPLOYED				
Raw	0.003	0.	.004	1.009
All other X's plus				
summary	-0.006	-0.041***	-0.041**	0.928
monthly employment	-0.007	-0.040***	-0.042**	0.925
ever employment	-0.007	-0.041***	-0.042**	0.925
sequence	-0.007	-0.043***	-0.044**	0.922
summary + monthly employment	-0.008	-0.042***	-0.044**	0.922
summary + ever employed	-0.007	-0.042***	-0.045***	0.921
summary + sequence	-0.007	-0.042***	-0.043**	0.924

-		F 1	D: 1		CIA-0 test					
	p	Formal	Diverted	AIE_1	Raw	θ raw	OLS	Matching	θ_0	
DAYS EMPL			•							
All	30.4	4.0	26.4	-2.2	3.8	1.003	-10.4***	-11.2**	0.914	
Scotland	5.3	2.8	2.5	9.6	-75.0***	0.478	-71.1***	-64.2**	0.490	
NE England	29.2	1.0	28.2	0.0	2.7	1.023	-14.7	-18.8*	0.864	
NW England	6.2	3.7	2.5	21.1**	38.4*	1.336	31.6*	27.6	1.224	
Wales	23.6	3.6	20.1	-16.6	20.3	1.141	-4.9	-7.6	0.955	
East Midls	47.1	5.9	41.2	-15.5**	4.9	1.044	-11.1*	-10.7	0.916	
London	31.0	4.9	26.1	-3.5	12.9	1.127	-3.4	-6.4	0.947	
no Sctl, NE, EMidls	23.4	4.3	<i>19.1</i>	-1.9	13.3*	1.117	-2.9	-9.2	0.931	
EVER EMPL										
All	30.4	4.0	26.4	-0.006	0.004	1.009	-0.041***	-0.041**	0.928	
Scotland	5.3	2.8	2.5	0.041	-0.130	0.786	-0.063	-0.056	0.895	
NE England	29.2	1.0	28.2	-0.020	-0.003	0.994	-0.063**	-0.071*	0.880	
NW England	6.2	3.7	2.5	0.063*	0.165**	1.319	0.130*	0.130	1.242	
Wales	23.6	3.6	20.1	-0.044	0.031	1.049	-0.052	-0.038	0.946	
East Midls	47.1	5.9	41.2	-0.036	-0.001	0.998	-0.049**	-0.043	0.923	
London	31.0	4.9	26.1	0.000	0.046	1.105	-0.026	-0.030	0.942	
no NE, EMidls	21.3	4.1	17.2	0.009	0.033	1.066	-0.018	-0.030	0.947	

DAYS EMPLOYED

	р	ATE_1	ATE_0	ATE	$ATE_1 \neq ATE$
NW Eng, Wales, London	0.234	1.9	-3.8	0.5	no
All districts			-2.1	-2.2	no
All districts, θ -adjusted	0.304	-2.2	-8.8**	-4.2	*
Scotland	0 0 7 0		72.1	12.9	no
Scotland, θ -adjusted	0.053	9.6	19.4	10.1	no
NE England		0.0	5.7	1.7	no
NE England, θ -adjusted	0.292	0.0	-14.5	-4.2	no
East Midlands			-4.4	-10.3	no
East Midlands, θ -adjusted	0.471	-15.5**	-15.3*	-15.4**	no
EVER EMPLOYED		ATE	ATE	ATE	ATE - ATE
All but NE Eng. EMidle	$\frac{p}{0.213}$	$\frac{AIE_1}{0.011}$	$\frac{AIL_0}{0.007}$	$\frac{AIE}{0.010}$	$\frac{AIL_1 \neq AIL}{n_0}$
All districts	0.213	0.011	0.007	0.010	no
All districts, θ -adjusted	0.304	-0.006	-0.007	-0.007	no
NE England			0.033	-0.005	no
NE England, θ -adjusted	0.292	-0.020	-0.040	-0.026	no
East Midlands		0.001	0.020	-0.009	**
East Midlands, θ -adjusted	0.471	-0.036	-0.022	-0.029	no

Findings: Annual earnings (survey)

		Ν	D25+	ND	LP
$\Delta_{\mathrm{S=1},X}$		445.4**	$\Delta_{\mathbf{S}=1,X} \neq ATE$	788.1*** <i>Δ</i>	$S=1,X \neq ATE$
ATF	Weighting	579.6**	no	762.1***	no
111 L	Matching	551.2***	no	708.5***	no

Sensitivity analyses: ND25+

(CIA-0) test in terms of 2004/05 earnings (admin)

	X + history	Raw	θ raw	OLS	Matching	θ_0
Post-April group	monthly empl	-147	0.937	-240	-208	0.910
March-May group	summary + monthly	-465*	0.776	-275	-109	0.938

Full March-May group:

	р	ATE_1	ATE_0	ATE _a	$ATE_1 \neq ATE_a$
(A) 2004/05 earnings (admin)	0.248	183.9	531.7**	270.2	no

(B) annual earnings (survey)		$\Delta_{S=1,X} \neq ATE_{b}$	$ATE_a \neq ATE_b$	
$\Delta_{\mathrm{S}=1,X}$		273.1		
ATE _b	Weighting	819.6	no	no
	Matching	700.4**	no	no

Sensitivity analyses: NDLP

(CIA-0) test in terms of 2004/05 earnings (admin)

	X + history	Raw	θ raw	OLS	Matching	θ_0
Post-April group	summary	210	1.087	-82	-69	0.976
March-May group	summary	323	1.132	-10	52	1.019

Full March-May group:

	p	ATE_1	ATE_0	ATE_{a}	$ATE_1 \neq ATE_a$
(A) 2004/05 earnings (admin)	0.320	375.9	621.8	454.7*	no

(B) annual earnings (survey)		$\Delta_{S=1,X} \neq ATE_{b}$	$ATE_{a} \neq ATE_{b}$	
$\Delta_{\mathrm{S}=1,X}$		736.1		
ATE _b	Weighting	759.9	no	no
	Matching	566.0	no	no

Conclusions

Experiments

- can suffer from randomisation bias
- with admin outcome data offer ways to validate and if necessary correct non-experimental methods in addressing it

ND25+

• For employment outcomes, experimental set-up qualified conclusions from non-experimental methods throughtout:

Non-experimental methods $\rightarrow ATE_1$ under-estimates ATE

Once corrected $\rightarrow ATE_1$ representative of ATE

• Earnings results (gain) appear reliable; again ATE₁ representative of ATE

NDLP

- For employment outcomes, irrespective of correction: a zero (negative for E Midls durations) ATE₁ representative of ATE
- Earnings results appear reliable; sizeable ATE₁ representative of ATE

No evidence that ERA experiment has suffered from randomisation bias in terms of year-1 results.

No evidence that controlling for detailed histories from administrative data adequately deals with selection.