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

EALE 2012

Barbara Sianesi

The Employment Retention and Advancement (ERA) program

ERA treatment

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

Eligibles

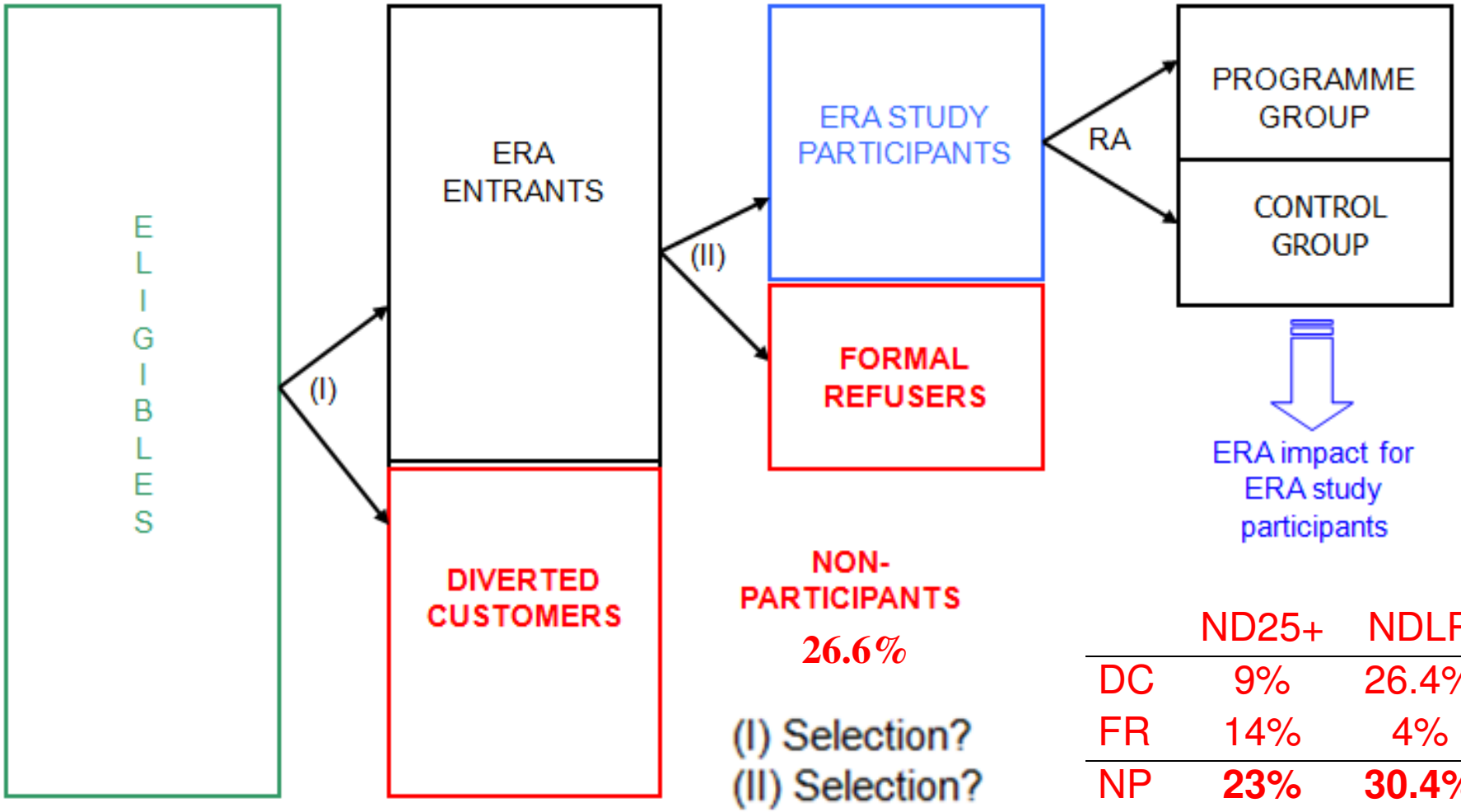
1) LT unemployed mandated for ND25+
2) Unemployed volunteering for NDLP
3) [LPs on WTC working PT who volunteer for ERA]

Tested

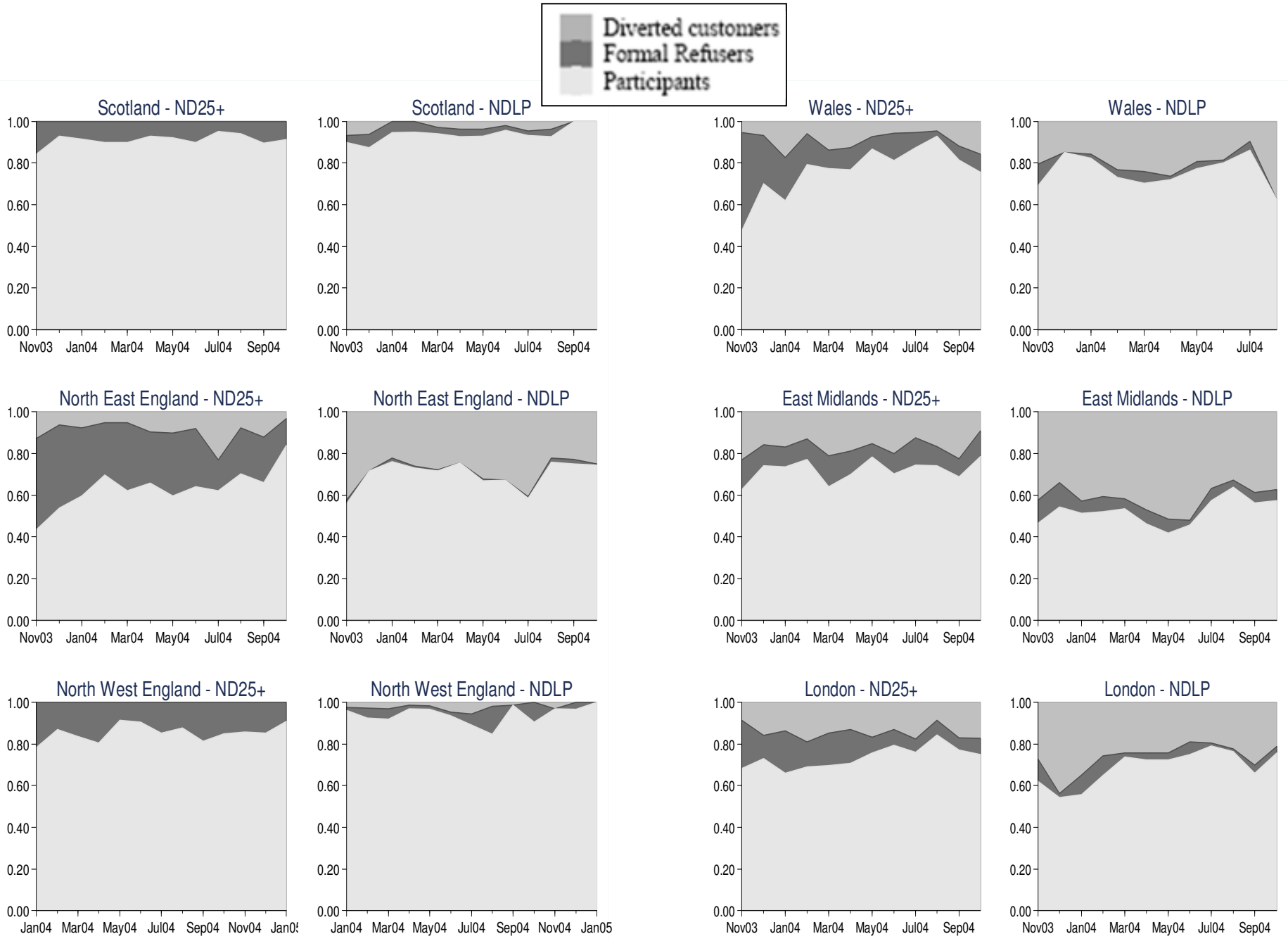
Large-scale (*N*=16,000), multi-site (6 districts) RA social experiment

Intake: Oct 2003 – Apr 2005 (pilots end Oct 2007)

Non-participation in the ERA study



	ND25+	NDLP
DC	9%	26.4%
FR	14%	4%
NP	**23%**	**30.4%**


Diverted customers
Formal Refusers
Participants
Scotland - ND25+
Scotland - NDLP
Wales - ND25+
Wales - NDLP
North East England - ND25+
North East England - NDLP
East Midlands - ND25+
East Midlands - NDLP
North West England - ND25+
North West England - NDLP
London - ND25+
London - NDLP
1.00
0.80
0.60
0.40
0.20
0.00
Nov03 Jan04 Mar04 May04 Jul04 Sep04
Jan04 Mar04 May04 Jul04 Sep04 Nov04 Jan05

Issues raised by non-participation

- Policymaker interested in impact of offering ERA for all those eligible to receive the offer. (ERA as an integral component of the New Deal)

- But… ERA tested only on a potentially selective subset of the eligibles

2 ways to view non-participation

- **Impact of offering ERA eligibility on the *eligibles*** (in the 6 districts)
  → Assess the scope for **randomisation bias** (Heckman, 1992 and Heckman *et al.*, 1999) in the experimental estimate for the parameter of interest

- **Impact of offering ERA eligibility on the *study participants*** (in the 6 districts)
  → 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 the participants
- Take up of ERA services by the non-participants

Methodology

Impact estimates under **selection-on-observables**: matching and re-weighting techniques

When follow-up data for non-participants not available: issue of survey and/or item non-response

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%

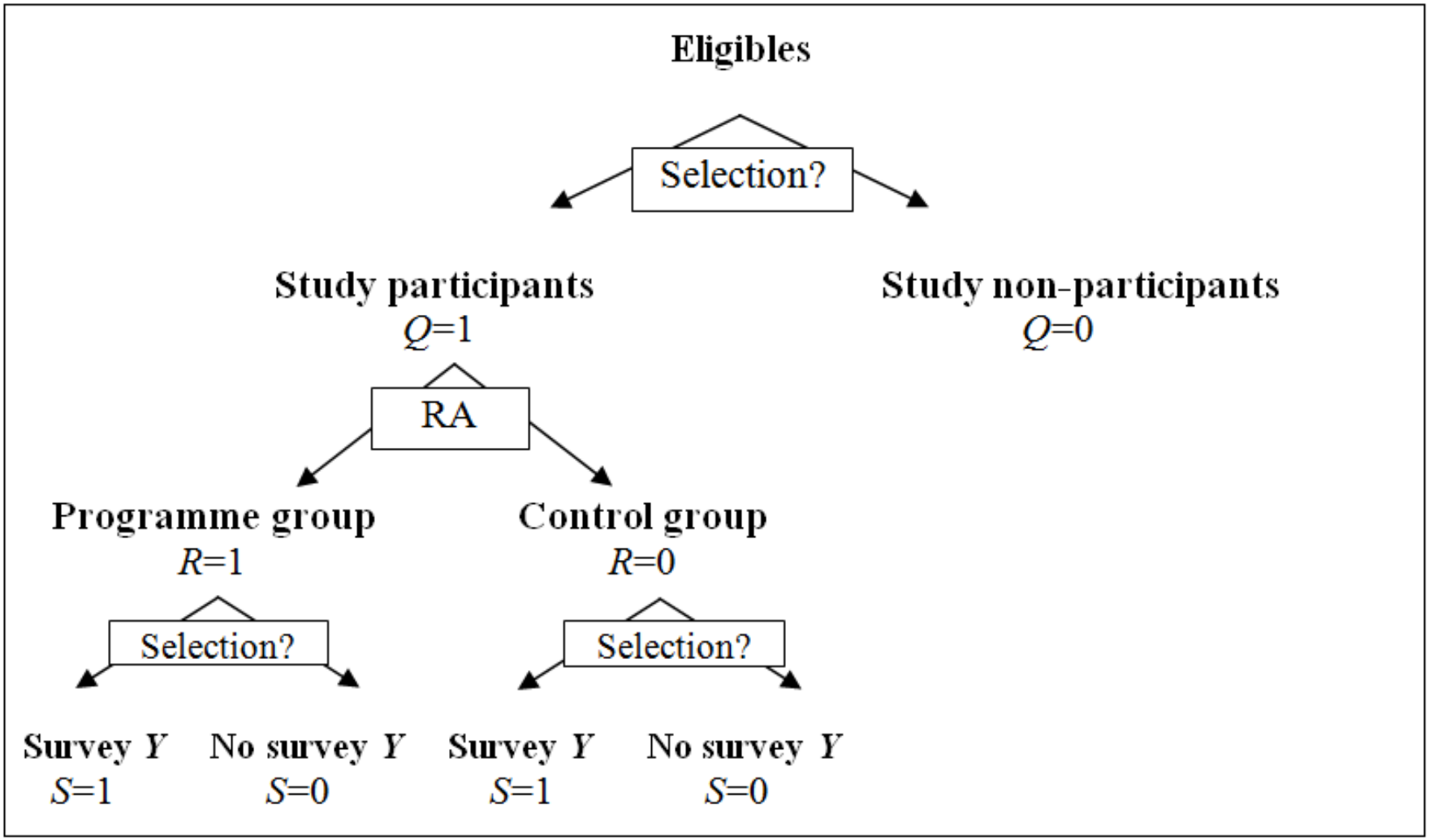
Outcomes

- 12-month follow-up
- employment (ever employed and days) – admin data
- benefits (days) – admin data
- annual earnings – survey 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
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 (3 years pre-inflow)	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+, active benefit history dummies (JSA and compensation from NDYP, ND25+, Employment Zones and WBLA and Basic Skills), inactive benefit history dummies (Income Support and Incapacity Benefits); employment history dummies
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

Methodology



$ATE \equiv E(Y_1 - Y_0)$ average effect on *all* eligibles

$ATE_1 \equiv E(Y_1 - Y_0 \mid Q=1)$ average effect on study participants

$ATE_0 \equiv E(Y_1 - Y_0 \mid Q=0)$ average effect on non-participants

$ATE = (1-p)\cdot ATE_1 + p\cdot ATE_0$ $\quad p \equiv \Pr\{Q=0\}$

Admin data:

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

Survey data:

$$ATE = (1-p)\cdot\mathbf{ATE_1} + p\cdot\mathbf{E(Y_1 - Y_0 \mid Q=0)}$$

Follow-up data

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

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

Assume

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

Implementation

Match to each non-participant one or more similar programme group member(s) based on $p(x) \equiv P(Q=0 \mid X) = P(Q=0 \mid Q=0 \vee Q=1, X)$ or $P(Q=0 \mid Q=0 \vee R=1, X)$.

Sensitivity analysis

(CIA-1') $E(Y_1 \mid Q=0, X) = \theta\, E(Y_1 \mid Q=1, X)$

allow participants and non-participants with the same X to differ in terms of some unobservable translating into a proportional difference of θ

No follow-up data

$ATE = (1-p)\cdot \mathbf{ATE_1} + p\cdot \mathbf{E(Y_1 - Y_0 \mid Q=0)}$

Akin to **attrition**
→ reweigh *Y* of the participants on the basis of the *X* of the eligibles to make them representative – in terms of *X* – of the full eligible population

Assume

(CIA-2) $E(Y_1 - Y_0 \mid Q=1, X) = E(Y_1 - Y_0 \mid Q=0, X)$ hence $= E(Y_1 - Y_0 \mid X)$

Implementation

The empirical counterpart can be derived in several ways:

1) **Reweighting**

Directly weigh the outcomes of the (responding) participants so as to reflect the distribution of X in the eligible population.

Ignoring survey/item non-response

$$\hat{ATE} = \left[\frac{(1-p)\,p_R}{\#(R=1)} \sum_{i \in \{R=1\}} \frac{y_i}{\left(1-p(x_i)\right) p_R(x_i)} \right] - \left[\frac{(1-p)(1-p_R)}{\#(R=0)} \sum_{i \in \{R=0\}} \frac{y_i}{\left(1-p(x_i)\right)\left(1-p_R(x_i)\right)} \right]$$

where $p_R \equiv P(R{=}1 \mid Q{=}1)$ and $p_R(x) \equiv P(R{=}1 \mid Q{=}1, x)$

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

$$\hat{ATE} = \left[\frac{1}{\#(R=1, S=1)} \sum_{i \in \{R=1, S=1\}} \frac{(1-p)\,p_{RS1}}{\left(1-p(x_i)\right) p_{RS1}(x_i)} y_i \right] - \left[\frac{1}{\#(R=0, S=1)} \sum_{i \in \{R=0, S=1\}} \frac{(1-p)\,p_{RS0}}{\left(1-p(x_i)\right) p_{RS0}(x_i)} y_i \right]$$

where $p_{RS1} \equiv P(R{=}1, S{=}1 \mid Q{=}1)$, $p_{RS0} \equiv P(R{=}0, S{=}1 \mid Q{=}1)$ and $p_{RS1}(x)$ and $p_{RS0}(x)$ are the corresponding probabilities conditional on x

2) **Matching**

Construct weights to realign X via matching

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

Can do it in 2 ways:

A) separately recover ATE_0 and then combine it with experimental ATE_1 to get the ATE or

B) recover ATE directly

Again, can do A) and B) both ignoring and allowing for survey and item non-response

(Detailed matching protocols in paper's Appendix)

Analysis of take-up

1. Are the non-participants individuals who even if offered ERA services would not take them up?
2. What kind of involvement would non-participants have had with JCP had they participated in the study and been assigned to the control group?

Assume CIA

Conditioning on *X*, participants and non-participants would have taken up the same amount of ERA services on average

Implementation

View take-up/involvement measures as outcomes and assess them as done for the admin outcomes (these are survey measures but non-response to these questions was <1%).

Findings

Experimental findings

	Raw	*Adjusted*
ND25+		
Ever employed	0.014	0.017
Days employed	4.0	4.6*
Days on benefits	-3.0	-3.0
Earnings	379*	393*
NDLP		
Ever employed	0.003	-0.006
Days employed	-0.1	-2.2
Days on benefits	-8.2**	-5.1
Earnings	885***	730***

Matching estimates for administrative outcomes

	ATE_1	ATE_0	ATE	$ATE_1 \neq ATE$
ND25+				
p =0.230				
Ever employed	0.017	0.056***	0.026**	***
Days employed	4.560**	9.984***	5.805***	*
Days on benefits	-2.966	8.862**	-0.250	***
NDLP				
p =0.304				
Ever employed	-0.006	0.015	0.000	
Days employed	-2.208	-1.957	-2.132	
Days on benefits	-5.078	8.881**	-0.831	***

Sensitivity analysis: ATE_θ, θ from 0.5 to 1.5

$E(Y_1 \mid Q{=}0, X) = \theta\, E(Y_1 \mid Q{=}1, X)$

ND25+

Ever employed		Days employed		Days on benefits	
θ	ATE_θ	θ	ATE_θ	θ	ATE_θ
0.50	-0.011	0.50	-0.783	0.50	-30.424
0.75	0.007	0.75	2.511	0.75	-15.337
0.88	**0.017**	**0.91**	**4.560**	**0.96**	**-2.966**
1.00	0.026	1.00	5.805	1.00	-0.250
1.25	0.044	1.25	9.099	1.25	14.836
1.50	0.062	1.50	12.393	1.50	.

NDLP

Ever employed		Days employed		Days on benefits	
θ	ATE_θ	θ	ATE_θ	θ	ATE_θ
0.50	-0.081	0.50	-20.027	0.50	-32.977
0.75	-0.040	0.75	-11.079	0.75	-16.904
0.96	**-0.006**	**0.99**	**-2.208**	**0.93**	**-5.078**
1.00	0.000	1.00	-2.132	1.00	-0.831
1.25	0.041	1.25	6.816	1.25	15.242
1.50	0.082	1.50	15.763	1.50	31.315

Matching and reweighting estimates for survey outcomes

Average ERA impact on earnings for all eligibles

		ND25+	NDLP
$\Delta_{S=1}$		393.2*	730.2***
ATE	allowing for non-response, weighting	559.9**	644.7**
	allowing for non-response, separate CS	580.2***	718.2***
	ignoring non-response, separate CS	442.8*	662.8***
	ignoring non-response, joint CS	443.5*	660.4**
% lost to joint CS		0.8	1.0

Take-up and involvement with JCP predicted for the non-participants

Under ERA

	ND25+		NDLP	
	Program group	Non-participants	Program group	Non-participants
Has had contact with JCP staff	84.8	83.7	85.3	86.4
Has ever initiated face to face visits	55.4	54.5	62.0	61.3
Had face to face contact with JCP staff ≥10 times	43.0	43.5	14.2	15.5
Received help/advice from JCP when not working	85.0	82.5***	77.2	74.8*
JCP staff did better-off calculation when not working	41.6	41.0	63.8	63.2
JCP staff suggested attend a Jobclub Centre	32.7	34.3	5.3	6.6*
JCP staff arranged education/training	30.4	31.3	14.6	17.8***
JCP staff offered help/advice w/out being requested	18.4	16.2**	26.3	27.6
Found advice from JCP staff overall very helpful	33.1	31.2	42.6	43.2
Found advice from JCP staff overall not at all helpful	4.7	5.0	3.4	2.5
Has heard of employment bonus	75.4	72.9**	72.8	71.0
Has heard of training bonus	43.0	40.1**	50.8	52.9

Without ERA

	ND25+		NDLP	
	Control group	Non-participants	Control group	Non-participants
Has had contact with JCP staff	78.2	78.2	71.9	74.6
Has ever initiated face to face visits	50.4	49.7	55.5	56.5
Had face to face contact with JCP staff ≥10 times	41.0	42.1	9.8	9.1
Received help/advice from JCP staff when not working	84.9	85.8	73.7	71.2
JCP staff did better-off calculation when not working	38.6	39.4	64.2	64.7
JCP staff suggested attend a Jobclub/Programme Centre	32.9	35.2	6.2	7.1
JCP staff arranged education/training	31.5	31.4	12.3	14.0
JCP staff offered help/advice without being requested	7.8	7.9	9.4	9.9
Found advice from JCP staff overall very helpful	23.6	22.8	31.1	35.1**
Found advice from JCP staff overall not at all helpful	5.8	5.2	4.1	3.7

Conclusions

- NDLP
  - story unchanged: 1st-year impact results *generalize* to eligible population

- ND25+
  - evidence of some randomization bias / some loss in external validity
  - experimental employment (and possibly earnings) estimates *underestimate* impact on all eligibles

- "Under normal operation, non-participants would not be interested in taking up ERA support and incentives."
  - no support for either intake group