A CACE in point: Estimating causal effects via a latent class approach in RCTs with noncompliance using Stata

Stata Conference 2021

Dr Patricio Troncoso
Heriot-Watt University
University of Manchester (Honorary)
@p_tronc
p.troncoso@hw.ac.uk

2021/08/05
Outline

- RCTs under noncompliance and CACE.
- Using Stata `gsem` to estimate a CACE model.
- Comparison with other software packages.
Randomised controlled trials (RCT) under noncompliance

A standard RCT compares the outcome of interest in two groups:

- The treatment group, who is randomly selected to receive the treatment
- The control group, who is randomly selected to continue with standard treatment or no treatment at all

In the presence of perfect compliance, a dummy for trial arm assignment is enough to estimate the effect of the intervention.

- That is known as "Intention to Treat" (ITT), where we use standard regression models
- But some participants may not receive the treatment in full or comply with all requirements
- As expected, this is a common occurrence that can bias the ITT estimate
What is CACE?

Complier Average Causal Effect -> Proposed by Imbens and Rubin (1997)

The difference between the outcome in those participants who complied with the intervention and those who would have complied if assigned to treatment
What is CACE?

Complier Average Causal Effect -> Proposed by Imbens and Rubin (1997)

The difference between the outcome in those participants who complied with the intervention and those who would have complied if assigned to treatment

- ITT = (a + b) - (c + d)
What is CACE?

Complier Average Causal Effect -> Proposed by Imbens and Rubin (1997)

The difference between the outcome in those participants who complied with the intervention and those who would have complied if assigned to treatment

- ITT = (a + b) - (c + d)
- Per protocol = a - (c + d)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Compliance = 1</th>
<th>Compliance = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Complier</td>
<td>b) Non-complier</td>
<td></td>
</tr>
<tr>
<td>c) &quot;Would-be&quot; complier</td>
<td>d) Non-complier</td>
<td></td>
</tr>
</tbody>
</table>
What is CACE?

Complier Average Causal Effect -> Proposed by Imbens and Rubin (1997)

The difference between the outcome in those participants who complied with the intervention and those who would have complied if assigned to treatment

- ITT = \((a + b) - (c + d)\)
- Per protocol = \(a - (c + d)\)
- As treated = \(a - (b + c + d)\)
What is CACE?

Complier Average Causal Effect -> Proposed by Imbens and Rubin (1997)

The difference between the outcome in those participants who complied with the intervention and those who would have complied if assigned to treatment

- ITT = \( (a + b) - (c + d) \)
- Per protocol = \( a - (c + d) \)
- As treated = \( a - (b + c + d) \)
- CACE = \( a - c \)
What is CACE?

Complier Average Causal Effect -> Proposed by Imbens and Rubin (1997)

The difference between the outcome in those participants who complied with the intervention and those who would have complied if assigned to treatment

<table>
<thead>
<tr>
<th>Compliance = 1</th>
<th>Compliance = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Complier</td>
<td>b) Non-complier</td>
</tr>
<tr>
<td>c) &quot;Would-be&quot; complier</td>
<td>d) Non-complier</td>
</tr>
</tbody>
</table>

- \( \text{ITT} = (a + b) - (c + d) \)
- \( \text{Per protocol} = a - (c + d) \)
- \( \text{As treated} = a - (b + c + d) \)
- \( \text{CACE} = a - c \)

But compliance is not observable in the control group, which is why:

- a) we need to make assumptions exploiting randomisation, and
- b) adopt a probabilistic approach
Assumptions of CACE

Random assignment

- This allows us to assume "equivalent" groups with respect to the outcome before the intervention

Monotonicity

- There are no defiers (those who do the opposite of assignment)
- Similarly, there are no always-takers (those who participate regardless of assignment)

Stable unit treatment value (SUTVA)

- The outcome for any participant is independent of the group assignment of other participants

Exclusion restriction

- The treatment has zero effect on the outcome in the non-complier group
What are the characteristics of a CACE statistical model?

To estimate a CACE model, we need the following:

- An appropriate regression model for the outcome of interest
- A binary variable indicating compliance in the treatment arm that is treated as known in the model.
- A probabilistic model for compliance in the control group with reasonable predictors
CACE and latent class models: An example

"JOBS II": RCT to prevent depression as a result of job loss: Little & Yau (1998). Data available here

Our outcome of interest is "depress" and "r" is the treatment dummy.

"c" is a dummy to indicate compliance. From this, we derive "comp" as compliance with the intervention in the treatment arm and leave it as missing (unobserved) in the control group.

There are also a few covariates, such as age and education level, etc.

We could fit a standard latent class regression model for depress with latent compliance as such:

```
gsem (depress <- r depbase risk) ///
   (C <- age educ motivate econ assert single nonwhite) ///
   (comp <- , logit), ///
lclass(C 2)
```

But this is not a CACE model because of the following:

- The treatment effect ("r") is estimated freely
- The observed compliance in the treatment arm is not treated as known
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

gsem (depress <- r depbase risk) ///
   (C <- age educ motivate econ assert single nonwhite) ///
   (comp <- , logit), ///
   lclass(C 2)
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

```
gsem (depress <- r depbase risk) ///
    (C <- age educ motivate econ assert single nonwhite) ///
    (comp <- , logit), ///
lclass(C 2)
```

**Step 1: extend the regression model for depression**
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

```plaintext
gsem (1.C: depress <- r depbase risk) ///
     (2.C: depress <- r depbase risk) ///
     (comp <- , logit) ///
     (C <- age educ motivate econ assert single nonwhite), ///
lclass(C 2)
```

**Step 1:** extend the regression model for depression **into 2 paths**
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

```
gsem (1.C: depress <- r depbase risk) ///
   (2.C: depress <- r depbase risk) ///
   (comp <- , logit) ///
   (C <- age educ motivate econ assert single nonwhite), ///
lclass(C 2)
```

**Step 1:** extend the regression model for depression into 2 paths
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

```
gsem (1.C: depress <- r depbase risk) ///
(2.C: depress <- r depbase risk) ///
(comp <- , logit) ///
(C <- age educ motivate econ assert single nonwhite), ///
lclass(C 2)
```

**Step 1:** extend the regression model for depression into 2 paths

**Step 2:** fix the effect of the intervention in the non-compliers class
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

```
gsem (1.C: depress <- r@0 depbase risk) ///
(2.C: depress <- r depbase risk) ///
(comp <- , logit) ///
(C <- age educ motivate econ assert single nonwhite), ///
lclass(C 2)
```

**Step 1:** extend the regression model for depression into 2 paths

**Step 2:** fix the effect of the intervention in the non-compliers class to zero:

This is the "exclusion restriction" assumption

- **1.C:** regression path for non-compliers. Treatment effect fixed at zero (@0)
- **2.C:** regression path for compliers. Treatment effect freely estimated.
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

```bash
gsem (1.C: depress <- r@0 depbase risk) ///
    (2.C: depress <- r depbase risk) ///
    (comp <- , logit) ///
    (C <- age educ motivate econ assert single nonwhite), ///
    lclass(C 2)
```

**Step 1:** extend the regression model for depression into 2 paths

**Step 2:** fix the effect of the intervention in the non-compliers class to zero:
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

```
gsem (1.C: depress <- r@0 depbase risk) ///
(2.C: depress <- r depbase risk) ///
(comp <- , logit) ///
(C <- age educ motivate econ assert single nonwhite), ///
lclass(C 2)
```

**Step 1:** extend the regression model for depression into 2 paths

**Step 2:** fix the effect of the intervention in the non-compliers class to zero:

**Step 3:** extend the latent class model for compliance
Making the most of gsem path notation

We can use gsem path notation to make a standard LC model into a CACE model:

```stata
gsem (1.C: depress <- r@0 depbase risk) ///
       (2.C: depress <- r depbase risk) ///
       (1.C: comp <- _cons@-15, logit) ///
       (2.C: comp <- _cons@15, logit) ///
       (C <- age educ motivate econ assert single nonwhite), ///
lclass(C 2)
```

**Step 1:** extend the regression model for depression into 2 paths

**Step 2:** fix the effect of the intervention in the non-compliers class to zero:

**Step 3:** extend the latent class model for compliance into 2 paths to treat observed compliance in the treatment arm as known

- **1.C:** path for non-compliers. Intercept fixed at \( \text{logit}=-15 \) (\_cons@-15\) to ensure non-compliers in the treatment arm (\text{comp}=0\) will be assigned to that class

- **2.C:** path for compliers. Intercept fixed at \( \text{logit}=15 \) (\_cons@15\) to ensure compliers in the treatment arm (\text{comp}=1\) will be assigned to that class.
Estimation and results: JOBS II dataset

This code is to replicate the results of Little & Yau (1998) using gsem (Stata 15 onwards):

/* Read the data */
infile depress risk r depbase age motivate educ assert single econ nonwhite x10 c c0 ///
using "http://www.gllamm.org/books/wjobs.dat", clear
/* Generate compliance indicator variable */
gen comp=c if r==1 /* compliance missing in control group */
/* Fit the model */
gsem (1.C: depress <- r@0 depbase@c1 risk@c2) ///
 (2.C: depress <- r depbase@c1 risk@c2) ///
 (1.C: comp <- _cons@-15, logit) ///
 (2.C: comp <- _cons@15, logit) ///
 (C <- age educ motivate econ assert single nonwhite), ///
lclass(C 2)

Note: @c1 and @c2 constrain the effect of depbase and risk on depress to equality across classes
Estimation and results: JOBS II dataset

These are the results for the Non-compliers class (class 1)

<table>
<thead>
<tr>
<th>Response</th>
<th>depress</th>
<th>Number of obs</th>
<th>502</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>Gaussian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Link</td>
<td>identity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response</td>
<td>comp</td>
<td>Number of obs</td>
<td>335</td>
</tr>
<tr>
<td>Family</td>
<td>Bernoulli</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Link</td>
<td>logit</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|               | Coef.       | Std. Err.    | z    | P>|z|   | [95% Conf. Interval] |
|---------------|-------------|--------------|------|-------|-----------------------|
| depress       |             |              |      |       |                       |
| depbase       | -1.463379   | .1826867     | -8.01| 0.000 | -1.821438 -1.10532    |
|               | .2624529     | .2791255     | 3.47 | 0.001 | .3973586 1.426155     |
| _cons         | 1.632537     | .2791255     | 5.85 | 0.000 | 1.085461 2.179613     |
| comp          |             |              |      |       |                       |
| _cons         | -15 (constrained) |   |      |       |                       |
| var(e.depress)| .506397     | .0322776     |      |       | .4469262 .5737814     |

The exclusion restriction for non-compliers is applied and the effect of the intervention ("r") on the outcome in this group is fixed to zero.
Estimation and results: JOBS II dataset

These are the results for the Compliers class (class 2)

<table>
<thead>
<tr>
<th>Class</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>depress</td>
</tr>
<tr>
<td>Family</td>
<td>Gaussian</td>
</tr>
<tr>
<td>Link</td>
<td>identity</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response</th>
<th>comp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>Bernoulli</td>
</tr>
<tr>
<td>Link</td>
<td>logit</td>
</tr>
</tbody>
</table>

|          | Coef. | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|----------|-------|-----------|-------|------|----------------------|
| depress  |       |           |       |      |                      |
| depbase  | -1.463379 | 0.1826867 | -8.01 | 0.000 | -1.821438, -1.10532  |
| x        | -0.3098673 | 0.1173219 | -2.64 | 0.008 | -0.5398141, -0.0799205 |
| risk     | 0.9117568 | 0.2624529 | 3.47  | 0.001 | 0.3973566, 1.426155  |
| _cons    | 1.81249 | 0.2971227 | 6.10  | 0.000 | 1.23014, 2.394839    |

|          |       |           |       |      |                      |
|----------|-------|-----------|-------|------|                      |
| comp     |       |           |       |      |                      |
| _cons    | 15 (constrained) |          |       |      |                      |

| var(e.depress) | 0.506397 | 0.0322776 | 0.4469262 | 0.5737814 |

The effect of treatment in the compliers class is **-0.3098673**

- This means that *compliers* are expected to score **0.31** less in the depression scale than "would-be" compliers
Other software packages

gllamm (Rabe-Hesketh, S. et al., 2004). JOBS II example available here


gsem

| Coef.     | Std. Err. | z   | P>|z| |
|-----------|-----------|-----|-----|
| -0.3098673| 0.1173219 | -2.64| 0.008|

Mplus

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>S.E.</th>
<th>Est./S.E.</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent Class 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEPRESS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TX</td>
<td>-0.310</td>
<td>0.130</td>
<td>-2.378</td>
<td>0.017</td>
</tr>
<tr>
<td>RISK</td>
<td>0.912</td>
<td>0.247</td>
<td>3.685</td>
<td>0.000</td>
</tr>
<tr>
<td>DEPBASE</td>
<td>-1.463</td>
<td>0.181</td>
<td>-8.077</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Latent Gold

gllamm

| Coef.     | Std. Err. | z   | P>|z| |
|-----------|-----------|-----|-----|
| -0.3098673| 0.1173219 | -2.64| 0.008|

Latent Gold
A recent application: The Good Behaviour Game
The Good Behaviour Game

GBG is a school-based, universal behaviour management intervention implemented by class teachers.

- Core components: classroom rules, team membership, monitoring behaviour and positive reinforcement.
- Children work in teams to win the GBG for agreed rewards.
- Played during normal classroom activities for a specified time period.
- The class teacher monitors rules: working quietly, being polite, etc.

GBG trial in England

- 3084 children in 77 schools
- 1524 children in 39 control schools
- 1560 children in 38 intervention schools
- Intervention: 2015-2017
- Follow-up: 2017-2019
- Primary data collection
- Administrative data: National Pupil Database
Pre-registered hypotheses

Children in English primary schools implementing the GBG would demonstrate significantly better:

- mental health (conduct problems, psychological wellbeing, etc.);
- rates of school absence from school (the focus of this presentation)

For more details:

The GBG trial registration is available here

The full study protocol is available here

The final report is currently under review and will be published in the Journal of Public Health Research
Compliance with the GBG protocol

The duration of GBG play sessions varied across classrooms widely, so we chose dosage as our main indicator of compliance:

- "Compliers" are those children in classrooms where the GBG was played for at least the median time across intervention schools.

- "Non-compliers" are those in classrooms where the GBG was played for less than the median time.

- We also used the 75th percentile as sensitivity analysis.
CACE model for absence from school in the GBG trial (1)

We used Stata gsem to run a CACE model for absence

- This is the non-compliers class. The effect of trial is fixed to zero.

- Absence. Poisson distribution:
  family(poisson)
  exposure(exposure_var)

- Clustering around schools. Robust standard errors:
  vce(cluster clust_var)
The results for the compliers class are:

1. The effect of **trial** for compliers is -0.656 (IRR=0.519).
2. This means that compliers in the GBG have an incidence rate of 51.9% that of would-be compliers in the control group.
3. The GBG seems to be successful in **reducing absence** for those who played the game for sufficiently long.
Final remarks

gsem is a flexible suite of Stata commands that allows us to estimate complex statistical models to assess trial efficacy under noncompliance (and much more).

We can make use of various gsem specifications to address different issues:

- Various distributional assumptions for outcome measures
- Standard error options for clustering
- Use of constraints to test different hypotheses (common coefficients and/or variances)
- Model comparison and selection using information criteria (AIC and BIC available)
- Fast, easy to use and widely available
Thank you!

Acknowledgements

This work was supported by the funding of:

- Education Endowment Foundation (no grant number)
- National Institute for Health Research (grant number 14/52/38)

Data access:

The Department for Education (DfE) granted access to the National Pupil Database to analyse the educational outcomes of participant children in the Secure Research Service, hosted by the UK Office for National Statistics.

Disclaimer:

This work was produced using statistical data from the UK Office for National Statistics (ONS). The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.
References and further reading


