CODERS' CORNER

MIND & BEHAVIOUR
RESEARCH GROUP

csae
CENTRE FOR THE STUDY OF
AFRICAN ECONOMIES

MEDIATION ANALYSIS

When we analyse results of an experiment, we are often interested in understanding *how* the effects of our treatment/intervention came to be. That is, are there any intermediate variables (*mediators*) in our data that have affected our final outcome of interest as a result of the treatment?

Our goal in mediation analysis is to understand how much of the total treatment effect is due to: (i) an indirect effect operating through one or several observed mediators, and (ii) the direct effect of the intervention not captured by these observed mediator(s).

How can mediation analysis be useful in an experiment that has a behavioural component?

With multiple follow-ups on behavioural characteristics (e.g. self-efficacy, time preferences) and socioeconomic variables (e.g. consumption, savings, investment), we could use mediation to test whether socioeconomic outcomes in later rounds can plausibly be explained by changes in the psychological variables at intermediate follow-up rounds after the behavioural intervention.
For example, we might be interested in knowing whether changes in self-efficacy measured after watching a video showing role models to raise participant's self-beliefs and aspirations may plausibly be a mediator for investment decisions measured later in time.

This post gives a quick overview of how to run mediation analysis and reports an adapted version of the Stata code provided by [Acharya et al (2016)](#), which provides more details on this method.

What assumptions do we need for mediation analysis to be credible?

Mediation rests on the following two assumptions:
1. No omitted variables for the effect of treatment on the outcome.
2. No omitted variables for the effect of the mediator on the outcome, conditional on the treatment, pretreatment confounders, and intermediate confounders.[1]

In an experimental setting, the first assumption usually holds because the treatment has been randomly assigned. The second assumption might be a bit stronger, as we cannot test it, requiring a theoretical argument to support it.

Simple two-step procedure:
*   **Step 1:** regress the outcome on the mediator, the treatment variable(s), a set of controls, and the interaction between the mediator and all other variables. Obtain the predicted value of the outcome fixing all mediators to zero. This is the 'demediated' outcome.
*   **Step 2:** In the second step, regress the demediated outcome on the treatment variable(s). The coefficients from this regression give the estimate of the *average conditional direct effect*, as defined in Acharya et al. (2016).[1] Note that the standard errors from this regression will be biased due to the fact that they ignore the first estimation. To avoid this, we can use nonparametric bootstrap, performing both stages of the estimation in each bootstrap replication.

In Stata, we can implement this procedure in several lines of code:

```
reg outcome mediator treatment x_1  x_2
gen ytilde = outcome  - _b[mediator] * mediator
reg ytilde treatment  x_1
```

[1] This second assumption is referred to in the literature as sequential unconfoundedness

The coefficient on treatment in the last regression is our estimate of the *average conditional direct effect.* x_1 is a list of pre-treatment confounders, whereas x_2 are intermediate (after treatment) confounders that may be relevant.

To get bootstrapped standard errors, we enclose these three steps in a Stata program and pass this program to the bootstrap command:

```
program define deboot, rclass
      reg outcome mediator treatment x_1  x_2
      gen ytilde = outcome  - _b[mediator] * mediator     3
      reg ytilde treatment  x_1
      return scalar deffect = _b[treatment]
end

bootstrap deffect=r(deffect), reps(1000) seed(12345): deboot
```
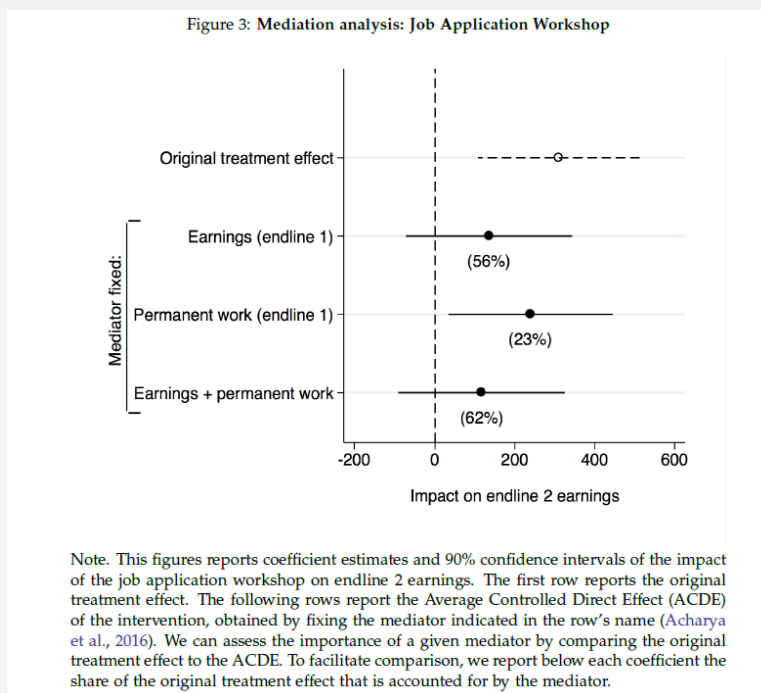
Some practical notes:

In an experiment involving more than one follow-up round of data (for example, a midline survey straight after intervention and an endline follow-up survey some time later), potential mediators for an outcome observed at endline may be observed variables that were significantly affected by treatment at midline.

For example, if we exposed participants to a video that aimed to change their aspirations, we may be interested in understanding whether any significant changes (if any), on the collected measures on self-beliefs after the screening of the video account for the changes (if any) on other socio-economic outcomes at endline.

Practically speaking, when the estimates of the average conditional effects are very similar relative to the average treatment effects, we can rule out that the treatment effect operates through our hypothesised mediator.

An example from a recent paper:

Abebe et. al (2019) use mediation analysis to understand the effects on wage earnings of a job application workshop to help unemployed young job-seekers in Addis Ababa better signal their skills to employers. As shown in the figure below, taken from their paper, they find that the 62% of the total treatment effects of the job application workshop four years after the workshop is mediated by the increase in earnings and likelihood of being in a permanent contract one-year after the workshop.

Figure 3: **Mediation analysis: Job Application Workshop**

Note. This figures reports coefficient estimates and 90% confidence intervals of the impact of the job application workshop on endline 2 earnings. The first row reports the original treatment effect. The following rows report the Average Controlled Direct Effect (ACDE) of the intervention, obtained by fixing the mediator indicated in the row's name (Acharya et al., 2016). We can assess the importance of a given mediator by comparing the original treatment effect to the ACDE. To facilitate comparison, we report below each coefficient the share of the original treatment effect that is accounted for by the mediator.

---

[2] **Average controlled direct effect (ACDE)** refers to the effect that the interventions would have on an outcome if the mediators are fixed at some particular value.
[3] Note that the code in Acharya et al. (2016) code omitted: "* mediator".

References:

Acharya, A., M. Blackwell, and M. Sen (2016). Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects. American Political Science Review 110(3), 512–529.

Abebe, G., Caria, S., Fafchamps, M., Falco, P., Franklin, S., & Quinn, S. (2019). Curse of Anonymity or Tyranny of Distance? The Impacts of Job-Search Support in Urban Ethiopia. *Mimeo*. Available at: http://www.simonrquinn.com/research/EthiopiaJobsExperiment.pdf

Giulio Schinaia, DPhil Candidate in Economics, Balliol College, and Research Assistant at the Mind and Behaviour Research Group, Blavatnik School of Government, Oxford
24 February 2020