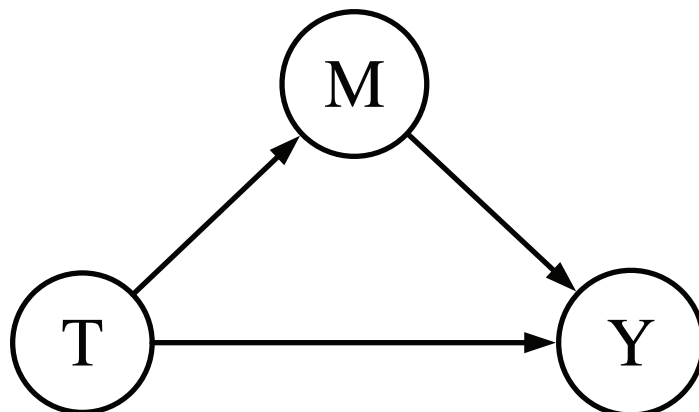


# Causal mediation analysis

Want to better understand a causal relationship? Causal mediation analysis explains these relationships by decomposing a treatment effect into a direct effect of the treatment on the outcome and an indirect effect through another variable, the mediator.

With Stata's new **mediate** command, you can fit causal mediation models for a variety of outcome and mediator outcome combinations.



- Linear, logit, probit, Poisson, and exponential-mean models
- Direct effects, indirect effects, total effects, proportion mediated, and potential-outcome means

- Continuous, binary, and count outcomes
- Continuous, binary, and count mediators
- Binary, multivalued, and continuous treatments

## Fit the model

Suppose we wish to evaluate the effect of physical exercise on self-perceived well-being and investigate potential causal mechanisms. Perhaps exercising causes an increase in certain hormones in the human body that, in turn, affects perceptions of well-being. To explore these relationships, we use a causal mediation model.

We specify the outcome, the mediator, and the treatment variable.

```
. mediate (wellbeing) (bonotonin) (exercise)
```

The natural indirect effect (NIE) of 9.8 is an estimate of the effect of exercise on well-being through production of the fictional hormone bonotonin. The natural direct effect (NDE) of 2.9 is an estimate of the effect of exercise on well-being through mechanisms other than bonotonin. Together, they sum to a total effect (TE) of exercise on well-being.

Perhaps we want to know the effect of exercise on well-being if we could control the bonotonin level. To estimate the controlled direct effect with bonotonin set to 10, we type

```
. estat cde, mvalue(10)
```

Viewer - view cma.smcl

File Edit History Help

view cma.smcl

view cma.smcl x

Dialog Also see Jump to

```
. mediate (wellbeing) (bonotonin) (exercise)
```

Causal mediation analysis Number of obs = 2,000

Outcome model: Linear  
Mediator model: Linear  
Mediator variable: bonotonin  
Treatment type: Binary

|     | wellbeing                         | Coefficient | Robust<br>std. err. | z     | P> z  | [95% conf. interval] |
|-----|-----------------------------------|-------------|---------------------|-------|-------|----------------------|
| NIE | exercise<br>(Exercise vs Control) | 9.799821    | .3943251            | 24.85 | 0.000 | 9.026958 10.57268    |
| NDE | exercise<br>(Exercise vs Control) | 2.891453    | .2304278            | 12.55 | 0.000 | 2.439823 3.343083    |
| TE  | exercise<br>(Exercise vs Control) | 12.69127    | .4005941            | 31.68 | 0.000 | 11.90612 13.47642    |

Note: Outcome equation includes treatment-mediator interaction.

CAP NUM INS

## Include covariates

We can control for confounders by including them in the model for the outcome, mediator, or both.

Viewer - view cma.sml

```
. mediate (wellbeing basewell age i.hstatus)
          (bonotonin i.gender i.hstatus)
          (exercise)
```

Causal mediation analysis

Number of obs = 2,000

Outcome model: Linear  
Mediator model: Linear  
Mediator variable: bonotonin  
Treatment type: Binary

|            | wellbeing                         | Coefficient | Robust<br>std. err. | z     | P> z  | [95% conf. interval] |
|------------|-----------------------------------|-------------|---------------------|-------|-------|----------------------|
| <b>NIE</b> | exercise<br>(Exercise vs Control) | 9.718811    | .369158             | 26.33 | 0.000 | 8.995274 10.44235    |
| <b>NDE</b> | exercise<br>(Exercise vs Control) | 3.092069    | .1676971            | 18.44 | 0.000 | 2.763389 3.42075     |
| <b>TE</b>  | exercise<br>(Exercise vs Control) | 12.81088    | .3723542            | 34.41 | 0.000 | 12.08108 13.54068    |

Note: Outcome equation includes treatment-mediator interaction.

CAP NUM INS

When controlling for covariates, the estimated total effect is 12.8. If everyone in the population exercised, well-being would be, on average, 12.8 points higher than if no one exercised. Of this, a 9.7-point increase is due to the mediating effect via bonotonin, and a 3.1-point increase is due to other mechanisms. We could type **estat proportion** to learn that the effect via bonotonin is 76% of the total effect.

## Alternative statistics

Add the **all** option to report alternative direct effects, indirect effects, and potential-outcome means.

Viewer - view cma.sml

```
. mediate (wellbeing) (bonotonin) (exercise), all
```

Causal mediation analysis

Number of obs = 2,000

|               | wellbeing                         | Coefficient | Robust<br>std. err. | z      | P> z  | [95% conf. interval] |
|---------------|-----------------------------------|-------------|---------------------|--------|-------|----------------------|
| <b>PMeans</b> | Y0M0                              | 57.11317    | .2753201            | 207.44 | 0.000 | 56.57355 57.65278    |
|               | Y1M0                              | 60.00462    | .3157888            | 190.02 | 0.000 | 59.38569 60.62356    |
|               | Y0M1                              | 66.68199    | .3258477            | 204.64 | 0.000 | 66.04334 67.32064    |
|               | Y1M1                              | 69.80444    | .2898927            | 240.79 | 0.000 | 69.23626 70.37262    |
| <b>NIE</b>    | exercise<br>(Exercise vs Control) | 9.799821    | .3943251            | 24.85  | 0.000 | 9.026958 10.57268    |
| <b>NDE</b>    | exercise<br>(Exercise vs Control) | 2.891453    | .2304278            | 12.55  | 0.000 | 2.439823 3.343083    |
| <b>PNIE</b>   | exercise<br>(Exercise vs Control) | 9.568827    | .3884522            | 24.63  | 0.000 | 8.807475 10.33018    |
| <b>TNDE</b>   | exercise<br>(Exercise vs Control) | 3.122447    | .2418591            | 12.91  | 0.000 | 2.648412 3.596482    |
| <b>TE</b>     | exercise<br>(Exercise vs Control) | 12.69127    | .4005941            | 31.68  | 0.000 | 11.90612 13.47642    |

CAP NUM INS

The pure natural indirect effect (PNIE) and total natural direct effect (TNDE) are alternative decompositions of the total effect. Potential-outcome means are the expected values of the outcome under specific conditions. For example, the Y0M0 of 57.1 is the expected well-being if no one exercises.

## Alternative models

By default, **mediate** assumes a linear model for the outcome and mediator and a categorical treatment variable.

For a binary outcome or mediator, add the **logit** or **probit** option.

```
. mediate (y1 x1 x2, logit)
          (m1 x1 x3, logit)
          (treat)
```

Then type **estat or** or **estat rr** to calculate effects on the odds-ratio or risk-ratio scale.

For count variables, use the **poisson** option.

```
. mediate (y2 x1 x2, poisson)
          (m2 x1 x3, logit)
          (treat)
```

Then type **estat irr** to calculate effects on the incidence-rate-ratio scale.

For exponential-mean models, use the **expmean** option.

```
. mediate (y3 x1 x2, expmean)
          (m1 x1 x3, probit)
          (treat)
```

When the treatment is continuous, we specify the **continuous()** suboption to indicate control and treatment levels of interest

```
. mediate (y x1 x2)
          (m x1 x3)
          (treat2, continuous(0 -2 -1 1 2))
```

We can visualize the treatment effects at each value of the treatment variable we specified with **estat effectsplot**.

