Preacher and Hayes 2008: Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models

* Hypothesis using mediation is very common in behavioral science/psychology
* Mediation is when a predictor affects a dependent variable INDIRECTLY through at least one intervening variable (the mediator).
  + Assessing multiple simultaneous mediators is difficult, and has not been studied, but is clearly needed.
* Overview of simple and multiple mediation, as well as several approaches to investigate the process.
  + Additionally how to contrast 2+ mediators within a model.
* Correlations b/w variables is IMPORTANT, as correlation is a necessary but not sufficient condition to prove relationship.
  + HOW or WHY a causal is also interesting, but generally involves *mediation* analysis, how some variables affect others through intervening (mediating) variables.

Simple Mediation

Diagram

Description automatically generated

* How X affects Y through M.
  + NOTE: It is important to establish the causal order of X, M, and Y on theoretical/procedural ground
  + Total effect of X on Y is the sum of direct and indirect effects, *c = c’+ ab*, thus *c’* is the difference b/w the total effect of X on Y and the indirect effect of X on Y through M.
    - *c’ = c – ab* : These identities hold in regression and SEM where M and Y are continuous, but not where one or more of the dependent variables are binary, we then need to use logistic or probit regression, in which case the identity does not hold (MacKinnon & Dwyer, 1993).

General overview of how to test mediation hypothesis have been proposed (see MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002, for an overview).

* Commonly used path is *causal steps strategy* (Baron and Kenny 1986)
  + Investigator estimates paths of model in figure 1 using OLS regression or SEM, and assess the extent to which some criteria are met.
  + Variable M is a mediator if X significantly accounts for variability in M, X significantly accounts for variability in Y, and M significantly STILL accounts for variability in Y after controlling for X, and the effect of X on Y decreases substantially when M is entered simultaneously as a predictor of Y.
  + The last criterion is satisfied when the first and third criteria are satisfied, and when the signs of the effects are consistent w/ the proposed mediation process.
  + Using the diagram above, criteria requires paths *a, b,*  and *c*, to be significant, *c’* to be smaller than *c* by a nontrivial amount.
  + Note some authors (Collins, Graham, & Flaherty, 1998; Judd & Kenny, 1981; Kenny

et al., 1998; MacKinnon, 1994, 2000; MacKinnon, Krull, & Lockwood, 2000; Shrout & Bolger, 2002) argue that a significant total effect of X on Y (quantified as *c* in the diagram) is NOT necessary for mediation to occur

Other approaches are not based on individual paths in the mediation model, but instead on the product term *ab*, because this product is equal to the difference b/w total and direct effect!

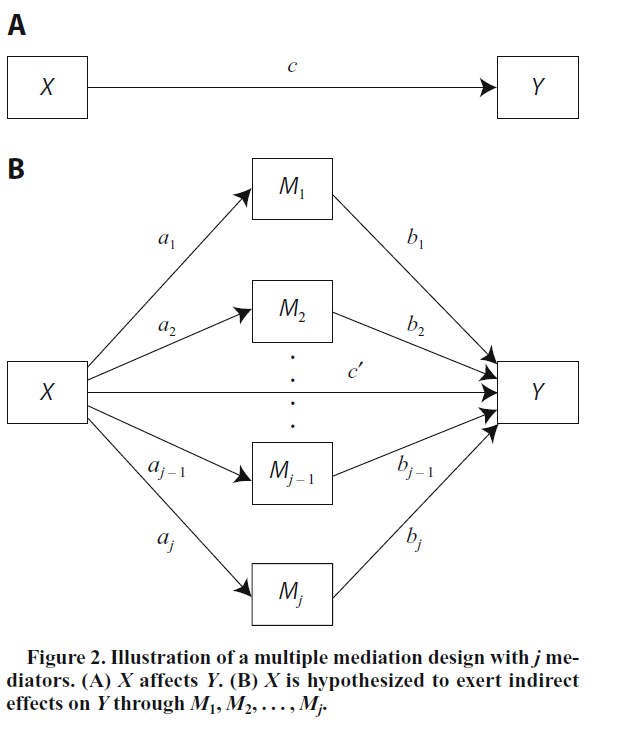
* The Sobel Test (Sobel, 1982, 1986) aka *product of coefficients* approach requires directly computing the ratio of ab to it’s estimated standard error (SE). Several formulas exist to estimate them, but the differences are negligible.
  + A *p* value is computed in reference to standard normal distribution, and significance itself supports the mediation hypothesis.
* However, using a standard normal for deriving *p*  for the indirect effect could be a problem b/ the sampling distribution of *ab* is only normal in LARGE samples.
  + *Distribution of the product approach* is a solution, based on the inference of the mathematical derivation of the distribution of the product of two normally distributed variables. Acknowledging how the distribution of products will be skewed and not requiring the assumption of normality. R code available (MacKinnon, Fritz, Williams, & Lockwood, 2007).

*Bootstrapping,* the nonparametric resampling procedure, is ANOTHER method for testing mediation that DOES NOT require normality on the sampling distribution.

* Computationally intensive, requires repeatedly sampling from the data set and estimating the indirect effect in each resampled data set.
  + By repeating thousands of times, an empirical APPROXIMATION of the sampling distribution for *ab* is built and used to construct confidence intervals for the indirect effect.
  + Details can be found with Bollen and Stine (1990), Lockwood and MacKinnon (1998), MacKinnon et al. (2004), Shrout and Bolger (2002), and Preacher and Hayes (2004, 2008).

Overall, these methods have been examine with simulations to asses Type I error rates and power, and the distribution of product approach or bootstrapping are seen as better than Sobel test or causal steps approach

* Because the first two have higher power while maintaining reasonable control over Type I error rate.
* Especially the causal steps strategy CANNOT be recommended except in large samples, see Preacher and Hayes (2004) and MacKinnon et al. (2002)

Multiple Mediation

* Simultaneous mediation by multiple variables w/ several
* Can have several X’s purportedly affecting a single Y
* The analytic methods for multiple mediation are MORE complex than for simple mediation
* Illustrated through figure 2, there are many paths, the indirect effects of X on Y via the j number of mediators.
  + Specific indirect effect of X on Y via mediator *i* however is the product of the two unstandardized paths linking x to y via that mediator
  + The effect of X on Y through M1 is a1b1, the TOTAL indirect effect of x on y is the sum of the specific indirect effects, the TOTAL effect of X on Y is the sum of the direct effect and all j of the specific indirect effects.
    - C = c’ + sum of indirect effects, the total indirect effect is c – c’
* Testing multiple mediation model instead of several separate simple mediation models is good!
  + Testing total effect of X on Y is like regression w/ several predictors
  + If an effect is found, can conclude that these j variables mediate the effect of X on Y.
  + And, what extent specific M variables mediate the X into Y effect, conditional on the presence of other mediators in the model.
  + Likelihood of parameter bias due to omitted variables is reduced!
    - Otherwise, can have biased parameter estimates, and has the issue of examining multiple comparisons instead of less comparisons (inflates type I error rate)
* It’s difficult to tease apart individual mediating effects that can often overlap in content.
  + The specific effect of M3 for example, isn’t just the effect of M3 alone, but conditional on all the other Mi­ in the model.
  + Thus, multicollinearity is an issue (same as in multiple regression!)

Investigating multiple mediation should involve 2 parts:

1. Investigating the total indirect effect (deciding which set of mediators translates the effect of X on Y
2. Testing hypothesis regarding individual mediators in the context of a multiple mediator model (the specific indirect effect associated w/ each mediator)
   1. A significant total indirect effect is NOT needed to investigate specific indirect effects.
   2. It is possible to find specific indirect effects to be significant in the presence of a nonsignificant total indirect effect!

Several approaches exist, primarily similar in ways to the original methods for testing simple mediation.

Causal Steps Approach

Generally used to find out whether or not c – c’ is a mediation effect. Here, the investigator asks whether the paths defining a specific indirect effect (ai and bi) are significant. If either of the paths through variable Mi is NOT different from 0, then Mi is not a mediator for effect of X on Y.

* Note: It’s possible for one variable to act as mediator M1, and another to act as a suppressor M2, in effect M1 and M2 cancel each other out (see MacKinnon et al., 2000)
* This is an EASY to understand extension, but it has the same problems as the single mediator version, thus it’s not used particularly commonly. It relies on a set of tests for individual *a* and *b* paths rather than testing the specific indirect effects, AND yields no point estimate or SE of the mediation effect.

Product-of-Coefficients Approach

* This works just fine for multiple mediators! Uses multivariate delta method to derive the SE of the total indirect effect (aka c – c’).
* The specific indirect effects can be investigated later through individual mediators.
* Total indirect effect for a model including the three mediators is the sub of the specific indirect effects!
  + ***F*** = a1b1 + a2b2 + a3b3
  + A more complex formula, using methods from Bollen (1987, 1989) generates the variance of the effect, **F**.
    - The square root of var(F) is the first order SE of the total indirect effect in a 3-mediator model. Assuming normality for the total indirect effect
    - A second-order ver. Of the multivariate delta method can be used, but the accuracy is only slightly improved.
  + Can either add or remove terms as necessary for larger or smaller than 3 mediator models from formula 1 in the paper.
* If using path analysis or SEM to fit, the residuals should be allowed to covary.
  + Obviously, because they all mutually depend on X! Thus covariance is somewhat expected.
  + If covariances are constrained to zero and actually correlated, the model will be misspecified, and the SE’s will be very biased, but the point estimates of a or b coefficients will still be correct.

Distribution of the Product Strategy

Can test specific indirect effects, however the distribution of sums or differences of products (needed to test hypothesis about total indirect effects/pairwise contrasts is NOT currently solved. Does not exist!

Bootstrapping

Super practical method! One assumption needed for use of SE’s derived via delta method (or a limitation of this multivariate extension for the product-of-coefficients strategy) is the need for multivariate normality!

* Not just the paths, but the sampling distributions of the total and specific indirect effects must be assumed to be normal for p-o-e strategy!
* We can use bootstrapping to solve multiple mediation, because we don’t HAVE to assume normality!

To bootstrap the sampling distribution of the specific and total indirect effects take a sample of size n cases with replacement from the original sample.

* A given case can be selected as part of a bootstrap sample 0, 1, 2 or even MORE times!
* Using this resample of size N, reestimate all j values of ai and bi­ and then calculate the product (ai \* bi) and the sum, from our resampled dataset.
  + Repeat this process k times, where k is at least 1000, yielding k estimates of the total and specific indirect effects of X on Y.
* Distributions of these K estimates serve as empirical, nonparametric approximations of the sampling distributions of the indirect effects of interest.
* Our bootstrap confidence interval for the population specific indirect effect through M1 is derived by sorting the k values of ai \* bi from low to high.
  + The lower and upper 100(alpha/2)% of the distribution are then found and taken as the lower and upper limits of the 100(1-alpha)% CI for the population indirect effect, where alpha is our desired nominal Type I error rate.
  + For example, with alpha of .05, we have 95% CI, and if K is 1000, we use the 25th and 976th values of ai \* bi in our sorted distribution. This is a PERCENTILE bootstrap CI.
    - Note, these can be asymmetrical b/c it’s based on empirical estimation rather than the assumption that the distribution is normal
* Bootstrapping is generally superior to product-of-coefficients in small/moderate samples

Contrasting Indirect Effects in Multiple Mediator Models

Sometimes need to test hypothesis that two indirect effects are equal in size.

* For example, which theory has greater impact in multiple mediation of the same model.

MacKinnon (2000) provides the only statistical treatment of contrast hypotheses such as these.

Example of Multiple Mediation

Testing hypothesis about early employee socialization (preentry knowledge, helpfulness of socialization agents) and how it affects socialization outcomes (job satisfaction, commitment, and cliarity).

* Proposed mediators are different types of socialization content expertise: Organizational goals and values, people, history, job performance proficiency, and politics (5 mediators)
* Theory was that early socialization experience is related to socialization outcomes, through how they affect the various elements of organizational socialization.
* Example is sub-set of hypothesis linking helpfulness of socialization agents to future job satisfaction.
  + Mediated by job proficiency, good work relationships, and understanding of workplace politics.
  + Only found indirect effect for the ‘people’ dimension.
* The total indirect effect of X on Y is f = a1b1 + a2b2 + a3b3 = .1074
  + Solved using equation 1, with var(f) = 0.0009719
  + Then Z = f/sqrt(var(f)) = 3.445
    - Reject H0 that the indirect effect is zero, p = 0.0006
* However, when directly comparing whether or not the effect of politics was significantly different than the effect of people, we can contrast the sampling variance, and determine that for both of them, there is 0 contained w/in the interval, thus even though people is significant, there is no significant difference b/w people and politics on how helpfulness affects job satisfaction.

Note, one big advantage of using SEM w/ latent variables is that unlike regression, you explicitly model measurement error. This allows you to test hypothesis using latent constructs rather than imperfect measured indicators!

MacKinnon 2002: A Comparison of Methods to Test Mediation and Other Intervening Variable Effects

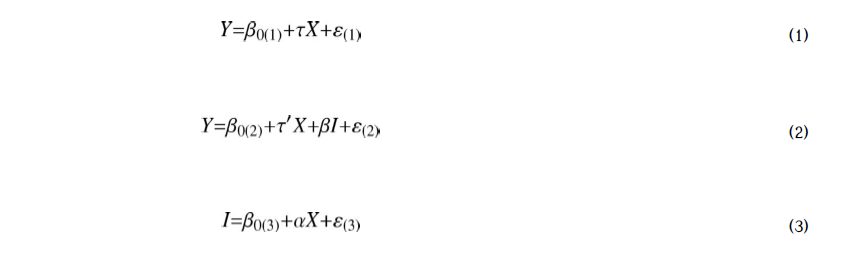
* Overview of a Monte Carlo study comparing various methods to test statistical significance of intervening variable effect (aka a mediator). The mediator transmits the effect of an independent variable to a dependent variable.
  + 2 methods based on distribution of product and 2 methods based on difference-in-coefficient methods have the most accurate type I error rates and greatest statistical power.
* X > I > Y , the effect of independent variable X on intervening variable I to affect dependent variable Y
  + Not used as commonly as should be, b/c either people don’t know the methods, there are too many methods to choose, or some ppl feel like the methods have too low statistical power.

There are three general schools of thought for methods when analyzing mediation.

1. Causal Steps Approach: Specifies a series of tests of links in a ‘causal chain’. Traditionally based on the work of Judd and Kenny (1981), and Baron and Kenny (1986) and is very commonly used.
2. Difference-in-Coefficients: Methods such like those that compare the difference b/w a regression coefficient before and after adjusting for the mediator (Freedman & Schatzkin, 1992;McGuigan & Langholtz, 1988;Olkin& Finn, 1995). Some of these methods test hypothesis about intervening variables that diverge from what psychologists are ‘used to’.
3. Product-of-Coefficients: Uses the product of coefficients involving paths in a path model (aka the indirect effect; Alwin & Hauser, 1975;Bollen, 1987;Fox, 1980;Sobel, 1982,1988).
   1. We use term intervening variable to refer to all non causal steps approaches for analyzing mediation.

Main goal is to simulate and thus compare/contrast Type I error rates and Statistical Power for all of these different mediation methods

* If power is too low, will not detect real effects in the population
* With Type I error too high, you can risk finding nonexistent effects.



Basic intervening Variable Model

* X is independent variable, Y is dependent, and I is intervening, our three B0 are the population regression intercepts for equation 1,2, and 3. t represents the relationship b/w independent and dependent variables after adjustment for intervening variables in equation 2, alpha represents the relationship b/w independent and intervening variables in equation 3, and beta represents tine relation b/w intervening and dependent variables adjusted for the effect of independent variable in equation 2.
  + We can do a series of tests ala Judd and Kenny, or Baron and Kenny
  + Can test each path involved in the effect (alpha and beta)
  + Test the product of the two paths (alpha x beta).
* Biggest similarity is that some me thods use the difference in the independent variable coefficients (t – t’), in equations 1 and 2 to estimate the value of the intervening variable effect.
  + If our coefficient (t’) does NOT differ significantly from 0 when our intervening variable is in our model, then the results are the same as a model where the effect is completely transmitted through the mediator.

Causal Steps Test of Mediation

Causal steps test is simple, and several series must be true for our mediator to be relevant. The sequence is pretty simple from Judd and Kenny. Testing X > M > Y.

Three main things are required from Judd and Kenny:

1. The treatment affects the outcome variable
2. Each variable in the causal chain affects the variable that follows it in the chain, when we control for all previous variables
3. The treatment exerts NO effect on the outcome when our mediating variables are controlled for.

Baron and Kenny have 3 different requirements:

1. Variations in levels of independent variable significantly account for variation in our mediator.
2. Variations in our mediator significantly account for variation in our dependent variable.
3. When our paths alpha and beta are controlled, a previously significant relationship between independent and dependent variables is NO LONGER significant.
   1. Generally assumed that there is a significant relationship b/w our independent and dependent variables.

Main difference b/w these two variations is that J and K need to show COMPLETE mediation, wherein there is no effect after accounting for the mediators. B and K find that only partial mediation is acceptable as well, and more realistic in social sciences.

Additional variation exists (Cohen & Cohen, 1983, p. 366): Researchers claim evidence for intervening variables effects when tests of each path are jointly (alpha AND beta) significant. This tests whether or not the independent variable relates to the mediator, and if the mediator is related to the dependent variable. However, this provides no test to see if alpha x beta product or overall X > Y relationship.

Generally, these variations are good and show SOME evidence towards the necessary conditions for strong inference of a causal effect through a mediator. But have some weaknesses:

* Does not provide a joint test of the three conditions.
* Does not estimate the size of indirect effect of X on Y (vs the direct effect)
* Does not provide standard errors for confidence intervals.
* Has a lot of trouble w/ multiple mediation models and evaluability.
* Cannot detect in cases of mediation where the indirect effect and direct effect ‘cancel’ each other out if the effects are in different directions.

Difference in Coefficients Test of the Intervening Variable Effect

We can also test mediators by comparing relationship b/w independent and dependent variable before and after controlling for our mediator. We have pairs of coefficients to compare in this case, such as the regression coefficients and the correlation coefficients.

In general, these procedures test a diverse set of null hypothesis about mediators.

Freedman and Schatzkin formula examines the difference b/w adjusted and unadjusted regression coefficients, and can examine standard error based on the variance and covariance of these adjusted and unadjusted regression coefficients.

Same w/ McGuigan and Langholtz, for standardized variables.

Clogg et al., examines ‘collapsibility’ which is whether or not we can ignore/collapse across a third variable when examining relationship b/w two variables. In mediation’s case, it’s testing to see if adding a mediator significantly changes the relationships between two variables. This can also be used to test whether or not the beta coefficient is significant.

Olkin and Finn use the multivariate delta method to find large sample standard error of the difference b/w a simple correlation, and the same correlation when adjusting for a third variable.

In summary, each difference in coefficient methods provides an estimate of the mediator AND it’s standard error. However, the null hypothesis might be strangely formatted and not resemble traditional psychological sciences ones. E.g. the Clogg test assumes fixed X and I, which isn’t realistic for mediating variables. The difference b/w simple and partial correlation represents a UNIQUE test of our mediating effect, because it seems like there is no relationship between the mediator and our dependent variable, but there is a mediation effect that exists! Main weakness is that these methods do NOT provide a framework for generalizing the tests to estimate appropriate coefficients and test significance for multiple mediators.

Product of Coefficients Tests for the Intervening Variable Effect

The third general approach is to test significance of mediator effect by dividing the estimate of the mediator effect (alpha x beta) by the standard error, and then comparing this value against a normal distribution.

The basic form is derived by Sobel (1982) using the multivariate delta based on 1st order taylor series approximation.

A more exact standard error can be calculated by using the first and second order Taylor series approximation (Aroian 1944) of the product of alpha and beta.

In both cases, the mediating variable effect is divided by the standard error, then compared against a normal distribution to test for significance (H0: alpha x beta = 0)

Goodman (1960) derived an unbiased variance of the product of two normal variables, subtracting the product of variances, giving a slightly modified equation (same as Arorian, but w/ subtracting the product instead).

MacKinnon, Lockwood, and Hoffman (1998) showed that using (alpha x beta)/(std error) as our method generally results in low power, as the distribution of the product of alpha and beta is NOT normally distributed, but is asymmetric with high kurtosis.

* Given multivariate normality of X, I, and Y, the paths for alpha and beta are independent
* Given the theory of products of random variables, (MacKinnon et al., 1998; MacKinnon & Lockwood, 2001) proposed some variants that should be more accurate.
  + An empirical distribution of the (alpha x beta)/(std error)
    - Significant simulations have resulted in a reference table of estimated critical values. For example, the empirical critical value is .97 for the .05 significance level rather than 1.96
  + Distribution of the product of two standard normal variables
    - The distribution for the product of two z statistics, one for alpha, and one for beta, if we assume both are normal, then za x zb can be directly tested for significance based on our theoretical distribution of the product for two normal random variables.
    - Thus, we convert alpha and beta into z scores, multiply them against each other, and use a critical value based on our table distribution for product of random variables. For example the critical level for alpha x beta = 0 for .05 in P = za x zb is 2.18 instead of 1.96!
  + Asymmetric confidence limits for the distribution of the product, alpha x beta.
    - The same as the previous, we calculate two z statistics, which are then used to find critical values for our product of two random variables to find lower and upper significance levels! If our confidence interval does not include zero, the mediator is considered significant!

Product of Coefficients method allow for estimates of our intervening variable effect, and the standard error of the mediator itself. Our model follows from path analysis where our mediator is the PRODUCT of coefficients that we hypothesize to measure causal relations. This works just fine for multiple mediator models! Two main problems, sampling distribution is NOT the normal distribution, and our H0 is very complex.

Simulation Study of Methods

Focused primarily on Type I error rate and statistical power. We also looked at mediator effect estimates and our standard errors for these.

Study varied as a 2 x 4 x 4 x 4 x 5 design. Factors of independent variable type (continuous/binary), effect size of path alpha (0, small, medium, large), path beta, and path t’, as well as sample size (50, 100, 200, 500, 1000) for 640 conditions, with 500 replications total.

Type I error rate was simulated by looking at the NO effect for alpha and beta, as since there is no effect, and 500 reps, we should expect 25 reps to show a significant effect (5%)

Otherwise, the # of times that each method found a significant effect was a fair measure of statistical power (as there was indeed, a ‘real’ effect in all other conditions in the simulation). The higher # of times the method let us reject a false H0, the greater the power!

Results

There was no difference in methods b/w the binary case and the continuous case (keep this in mind!)

Most estimates of the mediator had minimal bias, except for za x zb , as the point estimates for this were larger than point estimates for the mediator. Bias decreased as sample size and effect size increased for all methods.

All methods had standard error being generally accurate, except for Freedman/Schatzkin and Clogg estimates, being much smaller than true values for all conditions. Goodman’s method also yielded undefined standard errors. For example, Goodman’s unbiased error was undefined 40% of the time when the true effect was 0, and 10% of the time when effect size was small w/ small sample size.

Generally, standard errors for the product of regression coefficient in standardized variables were all very close to TRUE values for all conditions. The standard errors derived using the multivariate delta method were generally accurate!

Power and Type I Error

Generally, any condition where alpha ≠ beta but both > 0 had similar results.

Causal steps method has Type I error rates at very low values for all sizes. Baron and Kenny/Judd and Kenny had lower power for small and medium effect sizes. B and K had greater power as t’ increased, J and K had less power as t’ increased. The joint significant test was similar in that it had lower Type I error rates like the other causal steps methods, but it also had the most power in general. Generally power capped out at 0.80.

Difference in Coefficient methods had low Type I error rates, and had .80 or greater power and could detect small effects when sample size got to 1000, medium at 100, and large at 50. Even though standard error seems to underestimate TRUE standard error, the Type I error rates were the best, and had great statistical power.

Product of Coefficients methods had low Type I error rates and adequate power, similar to difference in coefficients methods. The za x zb  test with the z scores had good Type I error and the MOST power of ALL tests.

Overall, the two distribution methods, and the Clogg/Freedman and Schatzkin methods were the best w/ Type I error and power. But, Clogg method assumes fixed effects for X and I, so it might not work in all cases . Same for Freedman and Schatzkin.

Statistical Performance

J and K/ B and K are too underpowered. This is b/c requiring a total significant effect of X on Y leads to a lot of Type II errors. These methods are likely to miss real effects, but UNLIKELY to commit a Type I error. Good for specific use cases, but the alternative causal steps method, testing if alpha and beta are JOINTLY significant, has more power and more accurate Type I error rates.

Power rates for difference in coefficients methods are higher than B and K and J and K, but Type I error rates are TOO conservative except for clogg/freedman and schatzkin tests. Has the most accurate Type I error rates and greatest power for most situations. These methods underestimate standard error, but that compensates for too low critical values in the standard reference distribution! Product of coefficients method is higher power, but the Type I error rates are too low.

However, Clogg/Freedman and Schatzkin has an exception, when true pop values for alpha are 0, and beta is nonzero, the methods conclude that there is a mediator FAR too often (if alpha is = 0, should be no mediator). Because the test of significance is equivalent to whether or not beta is statistically significant.

In summary – tests of mediation trade off two competing issues! Non-normal sampling distribution of our alpha x beta effect leads to tests that are associated w/ lower empirical levels of significance than stated levels, when H0 is true, and low power when H0 is false. Second, the test for the null hypothesis for alpha x beta = 0 is complex, because it is a COMPOUND form where a = 0, b = 0, a = 0 and b ≠ 0, and b = 0 and a ≠ 0. In contrast, using otherwise overly conservative critical values turns out to empirically compensate for the inflation in Type I error rate due to this compound null hypothesis!

Statistical Recommendations

Generally, use either z’ = (alpha x beta) / (std error), for maximum power and increased Type I error rate if alpha or beta parameter is 0. Otherwise, use asymmetric confidence interval test w/ accurate Type I error rates, good power, and provides estimates of the magnitude of the mediation effect.

Causal Inference

Requirements for causal inference are complex and controversial! Generally, the traditional J and K / B and K methods demonstrate that the causal processes (X > I > Y) are consistent w/ the data. However, this only holds up if the residuals in equation 2 and 3 are independent, which is a very easy assumption to violate! Holland (1988) analyses these assumptions for further reading.

Establishing conditions for causal inference REQUIRES a more complex design. Where ideally, both the treatment and mediator are manipulated in a randomized experiment.

“For example, imagine a hypothesized model in which commitment leads to intentions, which, in turn, leads to behavior. Subjects could be randomly assigned to a high or low commitment to exercise program condition, following which their intentions to exercise would be measured. Following this, subjects could be randomly assigned to a condition in which the same exercise program was easy versus difficult to access and the extent of their behavioral compliance with the program could be measured.”

Adding design features like randomization and temporal precedence can be useful to rule out alternative causal explanations!

Imai 2010a: Identification, Inference, and Sensitivity Analysis for Causal Mediation Effects

Causal mediation analysis is a standard applied by researchers in many disciplines. Determining the alternative causal mechanisms by examining the roles of intermediate variables. Under certain assumptions, we can prove that the average causal mediation effect can be NONPARAMETRICALLY identified! The Linear Structural Equation Model (LSEM) can be interpreted as an ACME (average causal mediation effect) estimator once we add some parametric assumptions. You can also use a specific sensitivity analysis in the LSEM framework (determining if there is an unmeasured confounder).

What is a causal mechanism?

* For example, using fumigants increases farm yields, but what is the intermediate phenomena? Suspect that it is reduction of eelworms that causes it.
* Helps explain which of various competing theories lead to an outcome, generally different theories have different causal paths underlying the same cause-effect relationship.

Note: Some of the assumptions simply cannot hold, the treatment could be randomized, so we can ignore the treatment assignment, but we cannot ignore the mediator! Since there could be unmeasured pre-treatment variables that could confound our relationship, sensitivity analysis is a crucial step.

Example – Political Issue Framing and Political Opinions

Hypothesized that different frames for the same news story alter subject’s political tolerance by affecting more general political attitudes. News clips shown about KKK rally, and 2 versions, one where KKK rally was shown as free speech issue, other as a violent public disruption. Hypothesis is specifically that tolerance is mediated by subject’s attitudes towards how important free speech is, and how important public order is.

We would like to identify causal mediation effects, rather than total causal effect or controlled direct effects. However, we are only able to randomize the news stories, NOT the subjects pre-existing attitudes (what we propose as the mediator). Thus, there could be unobserved covariates that confound our proposed relationship. For example, the subject’s political ideology affects both their public order attitude and tolerance for Klan rally under BOTH treatment conditions!

How to Identify Average Causal Mediation Effect (ACME)

Preconditions: Random sample of size n from a population, for each unit, *i*  that we observe four traits, (Ti, Mi, Xi, and Yi). Ti is the binary treatment variable = 1 if receives treatment, 0 otherwise. Our mediator is Mi, and Y­I is our outcome variable. Finally, Xi is our vector of observed pre-treatment covariates (similar to what we would use to calculate a propensity score!). *M, X,* and *Y* denote the support of the distributions.

Given this, what is a mediator? It MUST be a post-treatment variable that occurs before the outcome is realized! Other than this very obvious requirement, what a mediator is, is based on previous theory. For example, treatment is receiving vaccine, outcome is whether or not subject gets the flu. Scientist may say that antibodies are the mediator (Vaccine > antibodies > flu). However… the parents signing form for risks of vaccine could also be a mediator (in theory) – Hypothesis could be, getting informed of the risks will make parents LESS likely to have the child get the 2nd dose of vaccine, thereby increasing likelihood of getting flu.

Define this using potential outcomes framework. Mi(t) is potential value of mediator for unit I under treatment status t. Yi(t,m) is the OUTCOME for unit i under a specific combination of mediator and treatment value. Thus Mi =Mi(Ti) and Yi = Yi(Ti,Mi(Ti)). If there is j different values of mediator, there are 2j values for the outcome, only which ONE can be observed.

Thus, the causal mediation effect for unit i given treatment t is:

Sigma\_i(t) ≡ *Yi(t,Mi(*1*))* −*Yi(t,Mi(*0*))*  for t = 0 or 1.

This sigma\_i(t) is the natural indirect effect, sigma\_i(0) is the pure indirect effect, and sigma\_i(1) is the total indirect effect. Essentially, sigma\_i(t) is the difference between the potential outcome given treatment status t, and our potential outcome if the treatment is the SAME, but with a mediator value that would result under the other treatment status. We can observe the first one… but the second one is by definition unobservable.

Our outcome depends on the value of treatments and the mediators, and not on HOW they are realized. E.g. this assumption is violated if our outcome responds to the value of our mediator differently depending on if it was directly assigned or occurred at random.

Equation 1 formalizes how mediation effects represent the indirect effects of treatment through the mediator. Specifically, we wish to find the average causal mediation effect, ACME. T is described as the total causal effect, or the OUTCOME (Y) given that the treatment leads to the mediator existing MINUS the outcome given no treatment and no mediator! Eta here is defined as the natural direct effect or pure/total direct effect. This means this is the causal effect of the treatment on the outcome when the mediator is FORCIBLY SET to the value that WOULD occur under treatment status t (0 or 1), a.k.a. it’s the direct effect of our treatment WHEN THE MEDIATOR IS HELD CONSTANT.

Equation 3 is critical, because it shows how the TOTAL CAUSAL EFFECT is equawl to the SUM of our mediator effect under one treatment condition, and the natural direct effect under our OTHER treatment condition! Note that the causal mediation effect and natural direct effect is NOT the same as the ‘controlled direct effect’ of the mediator. Unlike mediation effects, controlled direct effects of mediator are for specific values of our mediator, not potential values. This is useful if we want to get how the causal effect of mediator on outcome changes as a function of treatment! AKA the causal mediation effect examines whether our mediator mediates the causal relationship between our treatment and outcome, and the controlled direct effect looks at whether or not the treatment MODERATES (interacts with) the causal effect of mediator on outcome! (very interesting!)

Main Identification Result

Under this framework, with specific assumptions, we can nonparametrically identify the ACME!

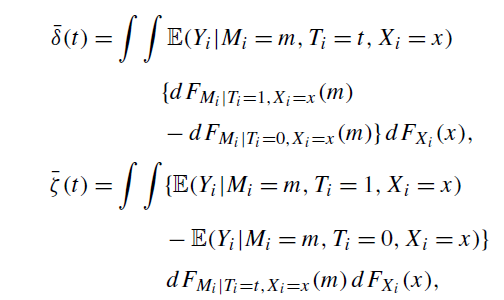
1. Assumption 1 (Sequential Ignorability): The outcome given a particular treatment prime and mediator, given a particular treatment, are independent of our treatment, GIVEN that our pretreatment covariates are x
   1. Second element, Our outcomes given treatment prime and a mediator are independent of our mediator (based on a treatment) GIVEN that Ti = treatment and Xi = x.
   2. For all treatment t,t’= 0,1 and all x contained within sample space of pre-treatment covariates. Note that probability of treatment given x is greater than 0, and ???? (can’t understand last element)

Our treatment is assumed to be ignorable (we don’t need it to determine our outcome, because it’s independent!) given the pre-treatment covariates, and our mediator is assumed to be ignorable GIVEN that we have the observed value of our treatment AND the pretreatment covariates.

* Note, unlike standard sequential ignorability assumption the condition independence in equation 5 MUST hold without conditioning on observed values of post-treatment confounders.

This results in the following Theorem!

Theorem 1 (Nonparametric Identification): Under Assumption 1, the ACME and the average natural direct effects are nonparametrically identified as follows for t =0,1:



where FZ(·) and FZ|W(·) represent the distribution function of a random variable Z and the conditional distribution function of Z given W.

This identification result does NOT hold under standard sequential ignorability assumption! Also, we must condition on the post-treatment counfounders Zi as well as the pretreatment covariates Xi. This limitation matters because you can’t assume the absence of post-treatment confounders based on the experimental setup. You can address this by conditioning on the pre-treatment variables alone in some cases!

For formula (9), we see that there is a no-interaction assumption. That given the same mediator, and a difference in treatment, there is a singular value Bi, that results, and is NOT the product of interaction.

Implications for Linear Structural Equation Model

Based on the original framework by Baron and Kenny, the following sequence of linear equations (eq 11 through 13).

Text

Description automatically generated

See MacKinnon (2008) for a review and Imai, Keele and Tingley (2009) for a critique of this literature.

Note that the original model can be additionally conditioned on any observed pre-treatment co-variates by including them as ADDITIONAL REGRESSORS in each equation.

Under this model, Baron and Kenny supposed that mediation can be tested by fitting the three linear regressions and testing the H0, b1=0, b2=0, and y =0. If all of these null hypothesis are rejected, then B2y could fairly be interpreted as a mediation effect!

* Equation 11 is apparently redundant, by subbing equation 12 into 13, we find that we don’t have to test b1=0. This is since the ACME CAN be nonzero even when the average total causal effect is zero
  + Aka what happens when the mediation effect offsets the direct effect of the treatment!
* Additionally, we see that the assumed natural direct effect is beta3, whereas the total causal effect is beta3 + beta2 \* gamma.
  + We also note no-interaction assumption between our treatment and our average causal mediational effect!
    - Which is the same as the no-interaction assumption for our average natural direct effects.

Connecting our parametric and non-parametric identification w/ each other, we find that the ACME equals the product of two terms representing the average effect of Ti on Mi and that of Mi on Yi (holding Ti at t), respectively!

Estimation and Inference

Parametric Estimation

Using the LSEM given by equations 12,13 and Assumption 1, the estimation of ACME is very easy since errors are independent. Following Baron and Kenny, can estimate equations 12 and 13 by fitting 2 separate linear regressions. Then, standard error for our estimated ACME can be calculated either approximately using the delta method, or exactly using the variance formula for Goodman.

Natural direct and total effects, standard errors, can be obtained via regressions of Yi on Ti and Mi, and Yi on Ti (equation 13 and 11 respectively).

Nonparametric Estimation

Simple example, nonparametric estimator for a discrete mediator with J distinct values. Estimate the ACME separately for each stratum of mediator, as defined by the pre-treatment covariates Xi. Then aggregate the total stratum of estimates to obtain an estimated composite ACME. In these cases, a nonparametric estimator can be gotten by plugging in the sample analogues for the population quantities in Theorem 1 (equation 18 [explain further?]).

By the law of large numbers, our estimator asymptotically converges to the true ACME under assumption 1.

Additionally, we can also derive the asymptotic VARIANCE of our estimator as well, not just a point estimate! We need the realized values of our treatment variable for this. This follows theorem 3. Note we can consistently estimate our variance by replacing unknown population quantities with the corresponding counterparts in the sample data itself!

You can also use nonparametric regression to model the means given the outcome based on treatment, mediator, and covariables and then use formula 19 to estimate the estimator. However, there is no simple expression for asymptotic variance for this estimator itself, so instead of using a clean formul9a, we require a bootstrap procedure instead.

Finally, can nonparametrically model our distribution for our mediator when it is not discrete as well. This uses a Monte Carlo draw of our mediator Mi from it’s predicted distribution based on the fitted model (see Imai, Keele and Tingley 2009)

ALSO! Note that the R Package, Mediation, is managed directly from the results in this paper, written by Imai (2010)!

Simulation Study

* Small scale monte-carlo experiment to investigate the performance of our estimators in equations 18 and 19, and our variance estimator in theorem 3.
* Varies in both sample sizes (50,100,500) and whether or not the estimator is parametric or nonparametric, repeated for 50,000 iterations. Half the sample receives treatment and half is in the control group.

Study finds that both estimators work very well with these parameters. Essentially 0 bias for nonparametric estimators, with a small increase in bias for parametric estimators. Converges towards no bias at larger sample sizes (500). Variance is larger for nonparametric estimator than the parametric one. 95% Confidence Intervals converge to nominal coverage once sample size increases.

Sensitivity Analysis

ACME works under assumption 1, but these assumptions may be TOO strong. For example, in situations where treatment is randomized, but the mediator is NOT randomized. In this case, equation 4 of assumption 1 is satisfied, but equation 5 may not for 2 reasons. 1: Unmeasured pre-treatment covariates that confound the relation between mediator and outcome, or 2: Observed or unobserved post-treatment confounders! These are the sources of some concern for mediation analysis (see Green, Ha and Bullock, 2010)

We can address this by assessing sensitivity of an estimated Average Causal Mediational Effect to unmeasured pre-treatment confounding. Note this does not address post-treatment confounding, which is best dealt with by experimental design.

Parametric Sensitivity Analysis

One of our elements does not hold unless equation 5 holds. We can assess sensitivity of our conclusions to violations of equation 5 by using ei2 =|= ei3 as our element for correlation. Correlation must range b/w -1 and 1. Also note! If correlation is NOT zero, this also violates assumption 1, which is another good check! We can check sensitivity by altering the correlation value directly, and then see if our ACME estimate is different.

Essentially, we can show that if the treatment is randomized, the ACME is identified given a particular value for p (Theorem 4). Additionally, we have a corollary, that under LSEM, our data generating process is not informative at all about either our sensitivity parameter (p) or the ACME without equation 5 holding true!

Additionally, our parameter p, can be defined/interpreted as the magnitude of an unobserved confounder (which is kinda the point!). Thus, the amount or influence of unobserved pre-treatment confounders can be calculated as coefficients of determination. These represent the proportion of previously unexplained variance that becomes explained by any unobserved confounders accounted for.

Empirical Application – Political Science Example

* Finds that the estimated ACME under free speech condition (delta 0) is larger than the effect under public order condition (delta 1) for both parametric and nonparametric estimators. The 95% CI for nonparmetric estimation of public order effect includes zero! Fail to reject the H0 for delta 0 = delta 1 under our parametric analysis, with p value of 0.238. The estimated ACME is between our previous two estimates and does NOT contain zero.
* Sensitivity Analysis: Researchers randomized the news stories, but the attitudes were NOT randomized, instead observed (the people came in w/ the attitudes they already had!). Subjects pre-existing ideology affects their attitudes towards public order AND tolerance for the Klan, in BOTH treatment conditions!
  + Tested to see how sensitive our estimates were to violations of independence assumption (which is fair, given the experimental procedure and methods)
  + Assumptions about the DIRECTION of the ACME under assumption one holds unless p is less than -0.68. Thus, even after holding the sample variability constant, our CI covers the value of zero only when -0.79 < p < -0.49. Thus our original finding of a negative ACME is robust to violations of Assumption 1 under the LSEM.

Conclusion

* Found a new identification condition for the ACME which is EASY to interpret, but weaker than existing results in some cases.
* Prove that estimates based on standard LSEM can have valid causal interpretations given the proposed framework.
* Nonparametric estimation strategies can be derived to find the ACME
  + Allowing researchers to not require such strong form assumptions as in the standard LSEM
* Parametric sensitivity analysis techniques have been developed to assess how sensitive estimates are to violation of various LSEM assumptions.

Imai 2010b: A General Approach to Causal Mediation Analysis

* Causal mediation has been formulated and implemented using linear structural equation models.
* This is bad for 3 reasons!
  + Lack of general definition of causal mediation effects independent of a particular model
  + Inability to specify key identification assumptions
  + Difficulty of extending the framework to nonlinear models
* Alternative approach exists, that is GENERAL!
  + Offers definition, identification, estimation, and sensitivity analysis, without referencing a specific statistical model
  + We link these 4 elements into a single framework, accommodating linear and nonlinear relationships, parametric and nonparametric models, continuous/discrete mediators, and various types of outcome variables.

One main goal of social science is causal inference, with randomized experiments as the gold standard. However, this is only a ‘black box’ view of causality!

* Estimation of causal effects allow researchers to examine whether a treatment causally affects an outcome!
* It CANNOT tell us how and why such an effect happens
  + This is important b/c identification of these mechanisms is needed to test competing theoretical explanations of the same causal effects!
* Our solution is mediational analysis
  + Identifies intermediate variables or mediators that lie in the causal pathway

Traditionally uses Linear Structural Equation Modelling (LSEM; e.g., Baron & Kenny, 1986; Hyman, 1955; James, Mulaik, & Brett, 1982; Judd & Kenny, 1981; MacKinnon, 2008; MacKinnon & Dwyer, 1993). However, this doesn’t work for TWO reasons!

1. By definition, LSEM cannot offer a general definition of causal mediation effects that are applicable BEYONE specific statistical models
   1. This is because the key identification assumption is stated in the context of a particular model
   2. Making it hard to separate limitations of research design from those of the specific statistical model!
2. Second the methods in LSEM don’t generalize to nonlinear models
   1. Including logit and probit models, for discrete mediators and outcomes as well as non- or semiparametric models.

Now we can use a single framework, placing causal mediation analysis within the counterfactual framework of causal inference and create a formal definition for causal mediational effects! Slightly extends our results of Imai (proved under sequential ignorability assumption the average causal mediation effects are nonparametrically identified).

Sequential Ignorability

* Conditional on the observed pretreatment covariates, the treatment is independent of all potential values of the outcome and mediating variables.
* The observed mediator is independent of all potential outcomes given the observed treatment and pretreatment covariates.
  + This is vital b/c it’s nonparametric identification, setting a minimum base of assumptions required for mediation effects to be causal without respect to statistical models used by researchers.

Additionally, develops general estimation procedures for causal mediation effects for linear and nonlinear relationships, para- and nonparametric, and continuous/discrete mediators. Additionally, have developed sensitivity analysis to be used for applied researchers

Example: JOBS Search Study II

* JOBS II: Randomized field experiment that investigates the efficacy of job training interventions for unemployed workers.
  + Designed to increase re-employment and enhance mental health of job seekers
* N = 1,801 unemployed workers, assigned randomly to treatment or control
  + Treatment: Job skills workshops with job search skills and coping strategies for dealing with setbacks.
  + Control: Booklet describing job search tips
* Measures:
  + Continuous measure of depression
  + Employed or not (binary)?
* Original hypothesis was that workshop attendance leads to better mental health and employment outcomes
  + This is mediated through increased subject confidence in their ability to search for a job.
    - This is measured with continuous measure of job-search self-efficacy.
* Additional pretreatment covariates matter
  + Depression, for example here.
  + Demographic variables
  + Level of economic hardship

Statistical Framework for Causal Mediation Analysis

* Describes the ‘counterfactual’ framework for causal inference.
  + Define causal mediation using potential outcomes notation
  + Show minimum set of conditions under which product of coefficients method and it’s variants let us estimate causal mediation effects
  + Differences with previous approach used by Rubin and Rosenbaum.

The Counterfactual Framework

* The causal effect of the job training program for each worker can be defined as
  + The difference between two outcomes
    - A: The worker participates in the job training program
    - B: The worker does NOT participate
* T­I can be our binary treatment variable
  + = 1 if worker i participated in program, = 0 otherwise
* Then Yi(t) denotes the outcome, or employment status, that results under treatment status t
  + For example, Yi(1) is worker i’s employment state if the worker participates in the program.
    - Although there are 2 values for each worker, only 1 is observed, if the worker did not participate, then only Yi(0) is observed!
  + Thus, if we use Thus, if we use Yi to denote the observed value of employment status, then we have Yi = Yi(Ti­) for all i
* Given this setup, the causal effect of the job training program can be defined as
  + Yi(1) – Yi­(0), but since only one of the states is observable, even randomized experiments cannot identify this unit-level causal effect!
  + Thus, researchers focus on the identification and estimation of the AVERAGE causal effect, which is defined as
    - E[Yi(1) – Yi(0)], this expectation is taken with respect to the random sampling of units from a target population
  + If our treatment is randomized (like in JOBS II), then Ti is statistically independent of potential outcomes!
    - See formal notation!
  + When this independence is true, the average causal effect can be identified by the observed mean difference between treatment and control groups. This is the ‘familiar result’ that the difference-in-means estimator is unbiased for the average causal effect in randomized experiments
* The above notation assumes no interference between units!
  + In this context, this means that worker i’s employment status is not influenced by whether another worker, j, participates in the training program.
  + This assumption is apparent as the potential values of Yi are written as a function of Ti, which does not depend on Tj for i is not equal to j.
  + Best to address this assumption in research design
    - We should ensure that participants in the experiment were not in the same household, for example (b/c then they would be related)
* All results here hold this assumption true

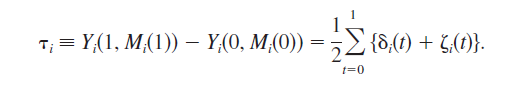
Defining Causal Mediation Effects

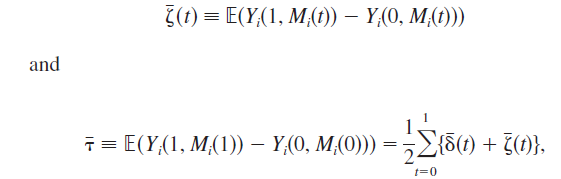
* Lets say we are interested in the mediating effect of job training program on depression
  + Mediating variable is level of confidence that worker has to do job search activities
* One hypothesis is that doing the job training program leads to less depression by improving the amount of self-confidence to search for a job
  + Mi denotes the observed level of job search self-efficacy, measured after the training program but before the outcome variable.
  + Because job search self-efficacy can be affected by the program participation, there are TWO values
    - Mi(1) and Mi(0), only one of which will be observed!
    - Mi = Mi(Ti), if worker i participates in the program, we will observe Mi(1) but NOT Mi(0).
* What about our outcomes? Previously our Y was only a function of the treatment, but in a causal mediational analysis, the outcomes depend on the mediator as WELL AS the treatment variable.
  + Thus, Yi(t,m) is the potential outcome that happens if the treatment and mediating variables equal t and m respectively!
  + For example, in the JOBS II study, Yi(1 , 1.5) is the degree of depression symptoms that are observed if worker i participates in the training program and has a job search self-efficacy score of 1.5!
* Any given outcome is only one of multiple potential outcomes, and the observed outcome equals Any given outcome is only one of multiple potential outcomes, and the observed outcome Yi equals Yi (Ti, Mi(Ti))
  + Lastly, no interference between units is assumed throughout. The potential mediator values for each unit DO NOT depend on the treatment status of other units, and the outcomes of any unit also do not depend on the treatment status and mediator values of other units.

We can define causal mediation effects for each unit i, as follows

* + for t = 0,1
* We see our causal mediation effect represents the indirect effect of the treatment on the outcome through the mediating variable.
  + The key to understanding equation 1 is the following counterfactual question!
  + “What change would occur to the outcome if one changes the mediator from the value that would be realized under the control condition, Mi(0), to the value that would be observed under the treatment condition, Mi(1), while holding the treatment status constant, at value *t*?
    - If our treatment had no effect on the mediator, that is Mi­(1) = Mi(0), then the causal mediation effect would be zero
    - Although Yi(t, Mi(t)) can be observed for any units with Ti = t, we cannot observe Yi(t, Mi(1 – t )), which is the outcome, given our treatment, with the mediation value that would have occurred under the other treatment, by definition this is UNOBSERVABLE!
* In the JOBS II study, d(1) represents the difference between two potential depression levels for worker i who participates in the training program
  + For this worker, Yi(1, Mi(1)) equals an observed depression level if the worker actually participated in in the program, where Yi(1, Mi(0)) represents the depression level that would result if worker i participates, but has the mediator value that would result under no participation (which is again, IMPOSSIBLE!).
* Similarly, d(0) represents an impact worker i’s depression level due to the change in mediator induced by the participation in the program, while suppressing the direct effect of program participation (which is SUPER IMPOSSIBLE!)
  + This definition formalizes, independent of other statistical models, the intuitive notion about mediation held by applied researchers
    - That the treatment indirectly influences the outcome through the mediator!

We can define the direct effect of treatment for each unit as follows:

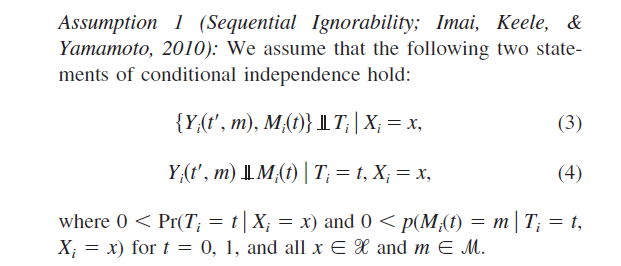
* for t = 0, 1
  + In the jobs study, S(1) represents the direct effect of the job training program on worker i’s depression level, while holding the level of job search self-efficacy constant at the level that would be realized under program participation.
* Addressing both of these elements, we can then decompose the total effect of our treatment into the causal mediation and direct effects.
* for t = 0,1
  + For the JOBS II study, this formula represents the average causal mediation effect amongst all workers of the population, of which the analysis sample CAN be considered as representative!
* Likewise, averaging over the relevant population of workers, we can define the average direct and total effects as

respectively

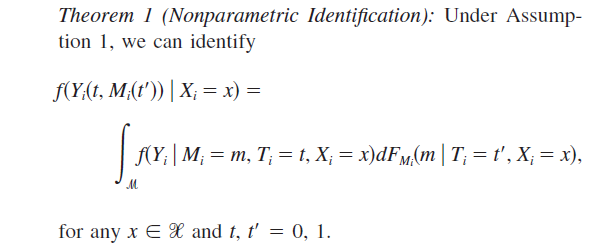
* As under our no interaction assumption, (i.e. avg delta = avg delta (1) = avg delta (0) and avg sigma = avg sigma (1) = avg sigma (0)), the average causal mediation effect and average direct effects sum to the average total effect
  + Total effect = average causal mediation effect (delta) + average direct effect (sigma)
* Note that the average total effect MAY be close to 0 in some cases, but that does not necessarily imply that the average causal mediational effect are also small
  + It’s possible that the effects have opposite signs, and thus OFFSET each other, yielding a small average TOTAL effect. This is important because it implies that a policy can be improved by modifying it such that an effective mediator plays a larger role to increase it’s overall effectiveness.

Sequential Ignorability Assumption

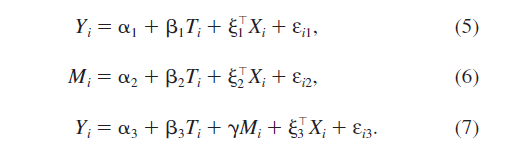
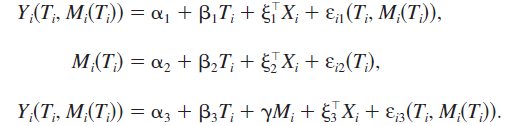
This is our main assumption, allowing us to make valid inferences about the causal mediation effects defined above.

* What assumptions are needed to give the average mediation effect a CAUSAL interpretation?
  + Randomized experiments just need to assume no interference between units to estimate the average treatment without bias
* Causal mediation analysis requires an ADDITIONAL assumption!
  + Let Xi be a vector of the observed pretreatment confounders for unit i, where chi denotes the support of the distribution of Xi (the range of values Xi can take on).
  + In the JOBS II data, Xi  includes pretreatment depression, and demographic characteristics such as education, race, marital status, etc.
* Given these confounders, our assumption can be written as
* 
* The main advantage of this assumption over other alternatives is the ease of interpretation!
* Assumption 1 is called sequential ignorability, because two ignorability assumptions are made sequentially
  + First, given the observed pretreatment confounders, the treatment assignment is assumed to be ignorable, that is, statistically independent of potential outcomes and potential mediators.
    - In the JOBS II study, this assumption is satisfied because workers were randomly assigned to the treatment and control groups
    - It would not be satisfied in any studies where the subjects may self-select into the treatment group!
    - Obtaining as many pretreatment confounders as possible helps improve the credibility of the ignorability of treatment assignment
  + The second part states that the mediator is ignorable given the observed treatment and pretreatment confounders. That is, the second part of the sequential ignorability assumption is made conditional on the observed value of the ignorable treatment and the observed pretreatment confounders.
    - The ignorability of the mediator may NOT hold even in some randomized experiments!
    - In the JOBS II study, the randomization of the treatment assignment does NOT justify this second ignorability assumption, because the post-treatment level of workers job search self-efficacy is NOT randomly assigned by the researchers.
    - In other words, the ignorability of the mediator implies that for workers with the same treatment status and pre-treatment characteristics, the mediator can be regarded AS IF it were randomized
  + This second stage is a STRONG assumption and should not be made lightly
    - It is always possible that there could be unobserved variables that confound the relationship between the outcome and the mediator variables even after conditioning on the observed treatment status and the observed covariates.
    - Additionally, the conditioning set of covariates must be pretreatment variables!
      * Without an additional assumption, cannot condition on posttreatment confounders even if we observe them.
  + Similar to the ignorability of treatment assignment in observational studies, it is difficult to know for certain whether the ignorability of the mediator holds EVEN AFTER the researchers collect as many pretreatment confounders as possible
    - This assumption is referred to as ‘nonrefutable’ because it CANNOT be directly tested from the observed data
    - Instead, we can test for it using sensitivity analyses that allow us to quantify the degree to which the empirical findings are ROBUST to a potential violation of the sequential ignorability assumption
    - Sensitivity analyses are appropriate because they allow us to probe whether a substantive conclusion is robust to violations of our sequential ignorability assumption!

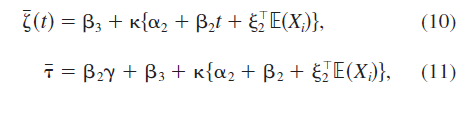
Nonparametric Identification Under Sequential Ignorability

* Nonparametric identification = Without any additional distributional or functional form assumptions to estimate the average causal mediation effect
* This is important for 3 reasons!
  + First: Possibility of constructing a ‘general’ method of estimating the average treatment effect for outcome and mediating variables of any type and using any parametric or nonparametric models
  + Second: Implies that we can estimate causal mediation effect while using weaker assumptions about the correct form or distribution of our observed data.
  + Third: Nonparametric identification analysis reveals the key role of the sequential ignorability assumption irrespective of the statistical models used by researchers.
* We can generalize this using a Theorem that is an extension of Assumption 1
* 
  + This proof is a generalization of Imai 2010a
  + This proof shows that under sequential ignorability, the distribution of the required potential outcome (the quantity in the left hand side of the equation) can be expressed as a function of the distributions of the observed data.
    - That is the conditional distributions if Mi given (Ti, Xi), and that of Yi­ given (Mi, Ti, Xi)
  + Thus, this assumption allows us to make inferences about counterfactual qualities that we do not observe (the potential outcomes and mediators of workers in the opposite treatment status) using the quantities we do observe (observed outcomes and mediators for workers in a particular treatment status)
* These conditional distributions are given by a set of the linear regression models, but as it is not based in any specific model, this enables us to develop a GENERAL estimation procedure for causal mediation effects under various nonlinear conditions.

Causal Interpretation of the Product of Coefficients and Related Methods

* The potential outcomes framework encompasses standard mediation based on the classic single mediator LSEM as a ‘special case’, for example:
* 
* After fitting each linear equation via least squares, the product of coefficients method uses beta\_hat2 x y\_hat as an estimated mediation effect
  + Similarly, the difference of coefficient methods yields numerically identical estimate by computing beta\_hat1 – beta\_hat3 in this linear case. Because beta\_hat1 = beta\_hat2\*y\_hat + beta\_hat3, and beta1 = beta2\*y + beta3 always holds, equation 5 is redundant given equations 6 and 7
* This product of coefficients method yields a valid estimate for causal mediation effect under the potential outcomes framework, but we need the sequential ignorability, and additional no-interaction assumption (mean delta(1) = mean delta (0)).
  + Note that this can be interpreted as a valid estimate that is ‘asymptotically consistent’ of the causal mediation effect as long as our assumption of linearity holds.
* To understand this connection between product of coefficients method and causal mediation as defined earlier, we can see the outcomes within the LSEM framework below:
* 
  + We show that under assumption 1, the average causal mediation effects are identified as Beta2 \*Y and Beta3 respectively, for t = 0, 1.
  + Thus, as long as we adopt the assumptions of linearity and no-interaction ALONG with sequential ignorability, the product of coefficients method provides a valid estimate of the causal mediation effect.
* However, theorem 1 implies that only the sequential ignorability assumption is needed, thus, how can we relax these assumptions?

Relaxing the No-Interaction Assumption

* Judd and Kenny (1981) and Kramer et al (2008) have an alternative to the standard product of coefficients method by relaxing the no-interaction assumption.
  + “No interaction between the treatment and mediator is often unrealistic”, thus, they replace equation 7 with equation 8, below:
  + 
  + This interaction could occur in JOBS II for example, if, the average mediation effect via the improvement workers mental health depends on whether or not they get the job training program.
* Kraemer argues that in addition to beta hat2, either y hat or k hat must be statistically indistinguishable from zero in order to conclude that average mediation effects exist
* Although our interaction term, TiMi is a reasonable suggestion, the procedure can be improved such that the hypothesis test is conducted DIRECTLY on the average causal mediation effects, as seen below:
* For t = 0, 1
* Estimation of this quantity in the JOBS II example would give different mediation effects, taking into account the interaction between program participation and job search self-efficacy. We can also estimate the average direct and total effects with:
* For T = 0,1
* The consistent estimates of sigma and theta can be obtained by replacing the coefficients of equations 10 and 11 with their least squares estimates, and expected value by the sample average of Xi, which we denote by Xbar.
  + Although we can easily relax no interaction, extending the LSEM mediation framework to models for discrete outcomes is more difficult

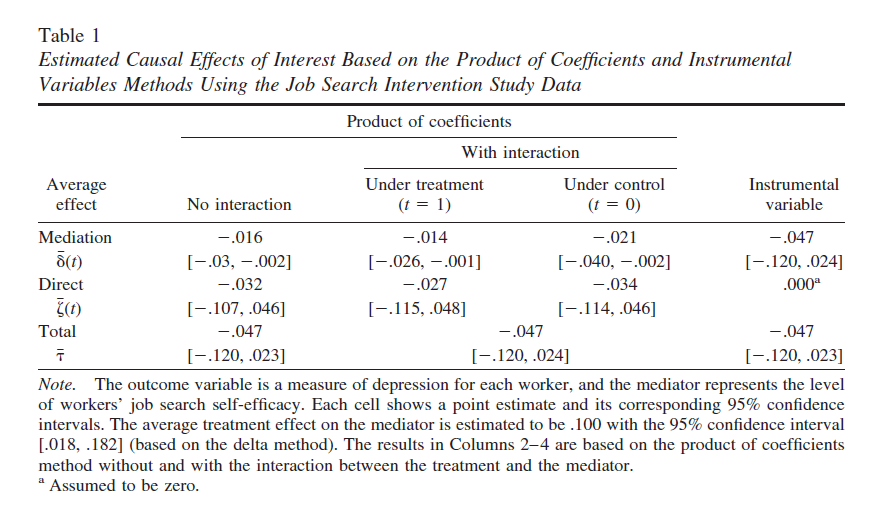
Relationship with Instrumental Variables

* Instrumental variables could be used for causal mediation analysis (e.g., Albert, 2008; Jo, 2008; Sobel, 2008).
* Using instrumental variables to estimate causal mediation effects requires an alternative set of identification assumptions
  + These differ from our original assumption 1 in important ways!
  + In particular, although the existence of unobserved confounders is allowed, the direct effect is assumed to be zero.
* This means, that the instrumental variables approach eliminates, a priori, alternative causal mechanisms
  + This is a serious problem, as knowing the causal mechanism is vital to social science research.

An Application to JOBS II: Product of Coefficients Method

* Outcome variable = Measure of depression
* Mediator = level of worker job search self-efficacy
  + Both measures range from 1-5
* We increase credibility for sequential ignorability by including the entire set of covariates

Table 1 indicates the estimated quantities based on both product of coefficients and instrumental variables methods.

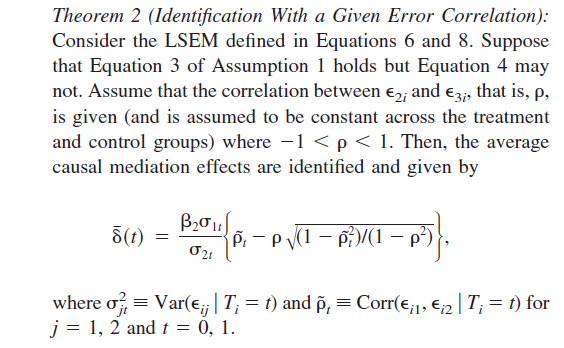


* First, standard product of coefficients method is used, assuming no interaction between the mediator and treatment.
  + Finds a small and statistically significant negative mediation effect (1st column)
  + Thus, since the average treatment effect on the mediator is negative, the results imply that the program participation on average DECREASES SLIGHTLY the depression symptoms, by increasing job search self-efficacy.
  + Average direct and total effects are estimated to be negative, with larger effect sizes (but are statistically indistinguishable from 0)
* Next, we relax the no-interaction assumption, allowing the average causal mediation effect to depend on the treatment status!
  + Second and third column have results
  + We see here that the findings are similar, and there is little evidence for the presence of an interaction
* Lastly, we see the instrumental variables method
  + It is unreasonable to hold the assumption of no direct effect in this application however. These results are somewhat different, and the ACME is estimated to be negative, but not statistically different from 0

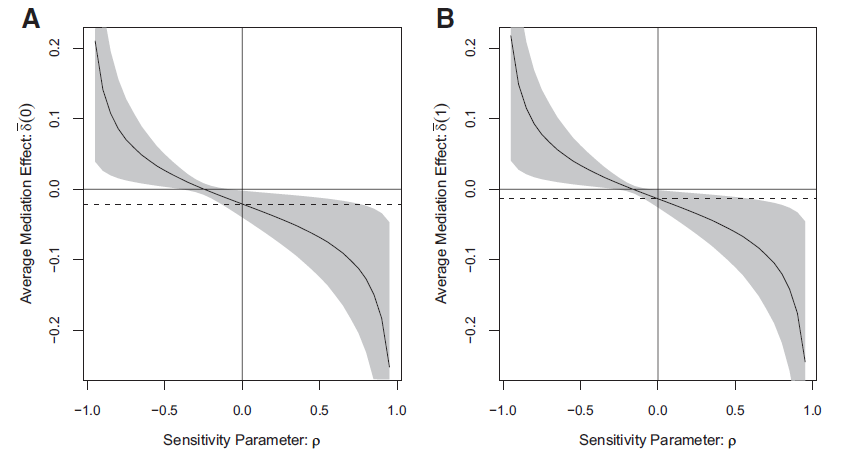
Sensitivity Analysis!

* This is considered a very important aspect!
* The potential outcomes framework clarifies the role of key identification assumptions
* As shown above, randomization of treatment does NOT identify causal mediation effects
  + Even in randomized experiments, an additional assumption (e.g. sequentially ignorability using our techniques) is required for identification
* Thus, the sequential ignorability assumption is very important! If it is not satisfied, the estimated quantity CANNOT be given a CAUSAL interpretation.
  + In particular, the 2nd part of Assumption 1 is ‘nonrefutable’. By definition, we cannot test this assumption with our observed data.
* Was sequential ignorability violated in our JOBS II study?
  + The second part might be violated by stating that ‘individuals who improved their sense of skill by one point in our intervention, may have different observed/unobserved characteristics than those individuals who had a similar improvement in our control condition’
* Sensitivity analysis addresses these nonrefutable assumptions by showing what conditions would have to exist in order for our assumptions to be proven wrong.
  + The goal is to ‘quantify the degree to which the key identification assumptions MUST be violated for the original conclusion to be reversed’
  + If an inference is sensitive, a slight violation of the assumption may lead to different conclusions!
* We can calibrate this degree of sensitivity as well.
* Given the importance of sequential ignorability, mediation analysis is incomplete without sensitivity analysis. (but it’s pretty easy to do in R!)

The Linear Structural Equation Models

* Looking at the standard LSEM framework, we can have sensitivity analysis for causal mediation based on the CORRELATION between error for mediation model, and error for outcome model.
  + This denotes the correlation between the two terms as Rho, our sensitivity parameter.
  + This correlation can arise if omitted variables affect both mediator and outcome variables
    - This is because the omitted variables will be part of BOTH of the two error terms!
* Thus, under sequential ignorability, Rho equals 0, and nonzero values of rho imply departures from our ignorability assumption.
  + Imai 2010 shows that we can express the ACME as a function of rho, and model parameters that be consistently estimated even though rho is nonzero.
* Thus the main question is such: How large does Rho have to be for the causal mediation effect to go away?
  + If small departures from zero in Rho produce an ACME that is substantively different from our estimate obtained under sequential ignorability… this suggests that our study is sensitive to the potential violation of the sequential ignorability assumption
  + Another criteria, does the CI for our mediation effect contain 0?
* Extending the result from Imai 2010 to the LSEM with no interaction assumptions relaxed, we derive the mediation effect as a function of Rho and other quantities that we can consistently estimate.
* 
* Theorem 2 establishes the linkage between average causal mediation effect and degree of correlation of our two different errors, given our identifiable model parameters.
  + Iterating further, we can obtain confidence intervals under various values of Rho.
* How small is ‘small enough’ to conclude that our result is valid?
  + There is no absolute threshold, as the magnitude of rho can only be interpreted relative to rho in other studies.
  + However… the magnitude of Rho from one study can be interpreted via coefficients of determination (a.k.a. R2!)
  + The alternative parameterization allows interpretation because one can understand the influence of potential omitted variables in terms of it’s explanatory power!

JOBS II Example: Empirical Illustration

* Looking at our continuous outcome and mediator, we ask whether our finding is sensitive to violations of sequential ignorability.
  + We relax the no-interaction assumption
* Delta(1) = 0 when Rho = -.165, and Delta(0) = 0 when Rho = -.245
* Figure 1 illustrates this point by plotting estimated average mediation effects and the 95% CI against various values of rho
* 
* Finding that for delta(0), the CI includes zero at a rho value of -.09 and for delta (1) at -.06.
  + This underscores the sensitivity of our estimate
  + In another study, the mediation effects are zero for a rho value of .48
* Thus, the estimated effects here are CONSIDERABLY more sensitive than in the other study.
  + It would take a smaller unobserved confounder to overturn the conclusion obtained under our ASSUMPTION of sequential ignorability for JOBS II

Generalization to Nonlinear Models

* We can show that the above methods can be applied and generalized to nonlinear models.
  + Difficulty for LSEM is that it does NOT extend to nonlinear models!
* For example, suppose in JOBS II the outcome variable is a BINARY measure about whether the person is employed or not?
  + LSEM using product of coefficients doesn’t generalize if we want to use logistic regression (perfect for binary) to model our outcome variable!
* This generalization can accommodate linear, nonlinear, parametric, and nonparametric models, continuous and discrete mediators, and various types of outcome variables.
  + This is possible because the method is NOT tied to specific statistical models.
  + Using nonparametric identification from Theorem 1, we develop 2 algorithms based on Monte Carlo simulation to estimate causal mediation effects that are applicable to any statistical models.
    - We can also adapt these to frequently used nonlinear models!

The Estimation Algorithms

* For each subject we observe Yi(T­­I, Mi­ (Ti))
* However… we need to infer the counterfactual quantity from this: Yi(T­­I, Mi­ (1 - Ti))!
  + Or what our outcome would be, given a treatment, and given the mediator value that would otherwise occur under the OTHER treatment! (which by definition… doesn’t exist!)
* Theorem 1 suggests that we can obtain one Monte Carlo draw of our potential outcome Yi(t­ , Mi­ (t’)) for any t and t’ using MODEL PREDICTIONS given the subjects pretreatment covariates Xi = x.
  + To do this, we must first sample Mi(t’) from our selected mediator model: f(Mi|Ti = t’, Xi = x)
  + Then given this draw of the mediator, sample Yi(t , Mi­ (t’)) from our outcome model, f(Yi|Ti = t, Mi (t’), Xi = x)
  + Our nonparametric identification result implies that this procedure doesn’t have to change, regardless of statistical models used for the mediator and outcome
* Once we obtain these Monte Carlo draws of potential outcomes, we can compute the relevant quantities of interest that are functions of these potential outcomes
  + Thus, we generate two general algorithms that can accommodate many situations that we encounter in practice
* First: We have an algorithm for parametric inference, in which parametric models (e.g. probit or logit) are specified for the mediator and the outcome variable.
  + We can describe this algorithm to estimate the average causal mediation effects!
* Note that this is based on the Quasi Bayesian Monte Carlo approximation of King, Tomz, and Wittenberg (2000) where our posterior distribution of quantities of interest is approximated by their sampling distribution.
  + Similar ideas have been used in previous models (see Bauer, Preacher, & Gil, 2006; MacKinnon, Lockwood, & Williams, 2004).
  + In contrast, this algorithm applies to ANY parametric statistical model of the researchers choice.

Algorithm 1 (Parametric Inference): Can be used for any parametric model!

Step 1: Fit models for the observed outcome and mediator variables

Step 2: Simulate model parameters from their sampling distribution

Step 3: Repeat the following 3 substeps:

1. Simulate the potential values of the mediator
2. Simulate the potential outcomes given the simulated values of the mediator
3. Compute the causal mediation effects

Step 4: Compute summary statistics such as point estimates and confidence intervals.

* The strength here is that this algorithm can be applied to ANY parametric statistical model!
* 1000 simulations per step (the default) works well here
  + Greater or lesser numbers could be needed depending on the complexity of the models themselves
  + Non/Semi-parametric models can be used as well, but an adaptation of the algorithm needs to be made to address the differences

Algorithm 2(Nonparametric Inference): Can apply to para and nonparametric models

Step 1: For each of the bootstrapped samples, repeat the following steps

1. Fit models for observed and mediator variables
2. Simulate the potential values of the mediator
3. Simulate the potential outcomes, given the simulated values of the mediator
4. Compute the causal mediation effects

Step 2: Compute summary statistics such as point estimates and confidence intervals.

* These algorithms are evaluated using sets of simulations in the paper itself
  + Two sets of simulations, the first a population model that has a mediator with a nonlinear effect on the outcome
    - Compared the performance of a semi-parametric model with that of a linear model with and without a quadratic term in the mediator
  + Other set of simulations has continuous mediator but binary outcome.
    - Compares the performance of probit model against product of coefficients method.
* In both simulations, Imai’s methods recover the population parameters with little bias!

Quantile Causal Mediation Effects

* Methods based on LSEM can show ACME under certain criteria (sequential ignorability)
  + However… what about distributional features of the outcome variable OTHER than the average
* For example: in JOBS II – What happens for individuals with a HIGH level of depression, rather than those of an AVERAGE level of depression?
  + Some individuals might respond to the intervention strongly, making the average a poor description of how specific types of people respond to treatment
* ‘Quantile Causal Mediation Effects’ address these issues!
  + ‘The difference between a certain quantile (e.g. median) of two relevant potential outcomes’
* In this context, Quantile regression allows for convenience in modeling the quantiles of the outcome distribution, while adjusting for various covariates.
  + Can replace equation 7 above with this quantile regression model
  + Can’t use the product of coefficients method… but algorithm 2 works just fine!

Quantile Regression Example: JOBS II Data

* In JOBS II we can see whether or not job training program DIRECTLY affects subjects’ levels of depression and whether job search self-efficacy mediated the relationship between outcome and treatment.
  + Both direct and indirect effects in the figure are averaged over all observed pretreatment covariates included in the original models.
* The magnitude of the estimated mediation effects increases slightly as one moves from lower to higher quantiles. But the change is very small, and consistent over time.
  + Contrasting this, the estimated direct effect varies a lot over the quantiles, and the CI’s are wide and always include zero, so it’s hard to be certain of it’s effect.

Nonparametric and Semiparametric Regressions

* LSEM framework allows estimation of ACME based on a set of linear regressions, and we need the assumption of linearity for this to work
  + How can we relax the assumption of linearity?
* For example: In JOBS II, change in depression mediated by increased job search self efficacy might be very small amongst those with high levels of job search self-efficacy. (a.k.a. ‘Ceiling effect’?)
  + For these subjects, program participation is unlikely to further increase mediating effects b/c of diminishing effect of treatment on mediator.
    - Conversely… mediation effects might be smallest amongst those with low job search skills, as they are unable to overcome societal and institutional thresholds that reinforce levels of depression.
* Instead of assuming linear relationships between variables, we can use non and semiparametric regressions to avoid linear functional form assumptions! (e.g., Keele, 2008)
  + Attempts to recover the ‘true relationship’ from the data while imposing much weaker functional form assumptions.
  + Could use quadratic terms in LSEM… but this transformation is often a poor approximation in actual practice.
    - Also, once you do, the product of coefficients method no longer applies to these and more complex situations.
  + We can use algorithm 2 to simulate parametric/nonparametric models however, and we can estimate the ACME using that!
* When we allow our mediator to have a nonlinear effect on our outcome, we can apply the generalized additive model (GAM) to estimate the average causal mediation effects
  + 
  + Instead of equation 7, where the s(x) function is a smooth and possibly nonlinear function that is ESTIMATED nonparametrically from the data (e.g. a Spline!)
* In the LSEM framework, we assume s(x) to be a linear function… but we relax the no-interaction assumption by fitting an additional following model instead of equation 8
  + In our analysis, we see that there is a mild threshold effect in the no-interaction model, self-efficacy must exceed average before we see depression go down.
  + When accounting for interaction, we see that there is a negative relationship at higher levels of our mediator.
    - Treatment group, pattern mirrors our no interaction model but an even stronger effect (but wider CI?!?)
    - We have a small but statistically significant negative mediation effect.
* We can indeed model nonlinearity (but it didn’t change our conclusion here)
  + Algo 2 allows us to relax basic model assumptions and STILL produce well-defined direct, mediation, and total effects, as long as we assume sequential ignorability!