The Stata Journal (*yyyy*) *vv*, Number *ii*, pp. 1–14

Causal Mediation Analysis

Raymond Hicks
Niehaus Center for Globalization and Governance, Princeton University
Princeton, NJ/USA
rhicks@princeton.edu
Dustin Tingley
Department of Government, Harvard University
Cambridge, MA/USA
dtingley@gov.harvard.edu

Abstract. Estimating the mechanisms that connect explanatory variables with the explained variable, also known as "mediation analysis," is central to a variety of social science fields, especially psychology, and increasingly fields like epidemiology. Recent work on the statistical methodology behind mediation analysis points to limitations in earlier methods. We implement in Stata computational approaches based on recent developments in the statistical methodology of mediation analysis. In particular, we provide functions for the correct calculation of causal mediation effects using several different types of parametric models, as well calculating sensitivity analyses for violations to the key identifying assumption required for interpreting mediation results causally.

Keywords: mediation, causal mechanism, direct effects, sensitivity analysis

1 Introduction

The **mediation** package is designed to estimate the role of causal mechanisms that transmit the effect of a treatment variable on an outcome. Causal mechanisms are central to many studies in the social and life sciences, and the statistical analysis of mechanisms is widespread.[1] By positing, and empirically testing causal mechanisms, scholars start to be able to explain *why* a relationship exists between two variables. The `mediate` and `medsens` commands contained in the **mediation** package, implement the procedures described in Imai, Keele, and Tingley (2010a) and Imai, Keele, and Yamamoto (2010c) for a common set of statistical models.

Earlier approaches to mediation analysis largely relied on some form of structural equation modeling. Unfortunately, these earlier methods were not derived from a formal framework for causal inference and did not did not permit sensitivity analyses with respect to key identification assumptions. Furthermore, earlier methods were difficult to correctly extend to non-linear models such as those with binary outcome variables. The tools in the **mediation** package enable users to conduct sensitivity analyses and cover several common statistical models that handle binary dependent variables. Mediation and sensitivity analysis are each implemented with a single line of syntax, making the

1. For example, a canonical paper on the topic by Baron and Kenny (1986) has over 24,849 citations according to Google Scholar (date accessed, 9/13/2011).

procedure simple for users. In this paper we discuss the foundations of these methods and how to use the **mediation** package. A longer, non-technical, introduction is available elsewhere (Imai et al. 2011)

2 Background

2.1 Notation

The underlying theoretical results that the **mediation** package is based upon formulates the identification of causal mechanisms in the common framework of potential outcomes. First consider a sample of units in an experiment that are either in the treatment $T_i = 1$ or control $T_i = 0$ condition. Thus the outcome for observation i in the treatment condition can be denoted as $Y_i(1)$, or more generally $Y_i(T_i)$. In practice we usually only observe each unit in one condition; thus for the above example $Y_i(0)$ is not observed. This implies that the *unit-level* treatment effect is unobservable, and researchers typically focus on estimation of the average treatment effect (ATE) over a population, $\mathbb{E}(Y_i(1) - Y_i(0))$.

Mediation analysis moves beyond calculation of average treatment effects and instead seeks to quantify the effect of a treatment that operates through a particular mechanism. Let $M_i(t)$ denote the potential value of a mediator of interest for unit i under the treatment status $T_i = t$. Similarly, let $Y_i(t, m)$ denote the potential outcome if the treatment and mediating variables equal t and m. Here we only observe one of the potential outcomes, and the observed outcome, Y_i, is $Y_i(T_i, M_i(T_i))$, which depends upon both the treatment status and the level of the mediator under the observed treatment status.

2.2 Quantities of Interest

They key quantity of interest is the calculation of how much of the treatment variable is transmitted by the mediating variable. Following Robins and Greenland (1992) and Pearl (2001) we define indirect effects or *causal mediation effects* for each unit i as,

$$\delta_i(t) \quad \equiv \quad Y_i(t, M_i(1)) - Y_i(t, M_i(0)), \tag{1}$$

for each treatment status $t = 0, 1$. This causal quantity is the change in the outcome corresponding to a change in the mediator from the value that would be realized under the control condition, i.e., $M_i(0)$, to the value that would be observed under the treatment condition, i.e., $M_i(1)$, while holding the treatment status constant at t. For example, if $M_i(1) = M_i(0)$ then the treatment has no effect on the mediator and the causal mediation effect would be zero. Importantly, because the treatment is fixed and only the mediator changes, we isolate the hypothesized mechanism. We can also define the *direct effects* of the treatment as

$$\zeta_i(t) \quad \equiv \quad Y_i(1, M_i(t)) - Y_i(0, M_i(t)), \tag{2}$$

for each unit i and each treatment status $t = 0, 1$. This represents all other causal mechanisms linking the treatment to the outcome.

What should be clear is that while we observe $Y_i(t, M_i(t))$ for units with $T_i = t$, we do not observe the counterfactual outcome $Y_i(t, M_i(1-t))$ in the typical research design with one observation per unit. This makes identifying causal mechanisms more difficult than identifying treatment effects, and requires an additional assumption known as sequential ignorability, discussed below. In practice, just as with treatment effects, we are interested in an average of the mediation effect. This is denoted the *average causal mediation effects* (ACME) $\bar{\delta}(t)$ and is formally defined as $\bar{\delta}(t) \equiv \mathbb{E}(Y_i(t, M_i(1)) - Y_i(t, M_i(0)))$. Similarly, the *average direct effects* (ADE) are defined as $\bar{\zeta}(t)$ is $\mathbb{E}(Y_i(1, M_i(t)) - Y_i(0, M_i(t)))$.

2.3 Identification Assumption

The ACME or ADE is not identified in the standard design, where the treatment is randomized/ignorable conditional pre-treatment covariates, and the mediator/outcome variables are measured. This is because a potential outcome required for the calculation of indirect and direct effects is never observed. An additional assumption is therefore required, sequential ignorability (SI). The assumption can be written as,

Assumption 1 (Sequential Ignorability (Imai et al. 2010c))

$$\{Y_i(t', m), M_i(t)\} \perp\!\!\!\perp T_i \mid X_i = x, \tag{3}$$

$$Y_i(t', m) \perp\!\!\!\perp M_i(t) \mid T_i = t, X_i = x, \tag{4}$$

where X_i is a vector of the observed pre-treatment confounders, $0 < \Pr(T_i = t \mid X_i = x)$ and $0 < p(M_i = m \mid T_i = t, X_i = x)$ for $t = 0, 1$, and all x and m in the support of X_i and M_i, respectively.

Assumption 1 applies two ignorability assumptions sequentially. In the first step, given the observed pre-treatment confounders, the treatment assignment is assumed to be ignorable – statistically independent of potential outcomes and potential mediators. This assumption is common, and also referred to as no omitted variable bias, exogeneity, or unconfoundedness. In experiments, the assumption is expected to hold since treatment is randomized. The second step assumes-given the actual treatment status and pre-treatment confounders-the observed mediator is ignorable. While the second step is similar to standard exogeneity assumptions, it is interesting to note that randomizing both the treatment and mediator does not identify the ACME (Imai et al. 2009, 2011).

2.4 Existing Methods and Practices

The standard approach to mediation analysis can be broken out into either 1) a set of steps whereby the statistical significance of slope estimates in a regression is evaluated or 2) the multiplication of slope coefficients along the causal path and a test of the

significance of the product. An extended discussion of these approaches is contained elsewhere (MacKinnon et al. 1995; Imai et al. 2010a). The key practical limitations of existing methods are 1) the difficulty in correctly extending to non-linear models (e.g., probit, etc.) and 2) the inability to conduct sensitivity analyses to the sequential ignorability assumption.

Regarding the first limitation, existing suggestions to estimate mediation effects with binary models using non-linear regression models (probit/logit) do not correspond to causal mediation effects (Imai et al. 2010a). When both the mediator and outcome variable are continuous and estimated with a linear regression, the mediation effect under the sequential ignorability assumption is equivalent to estimating two regressions,

$$M_i = \alpha_2 + \beta_2 T_i + \xi_2^\top X_i + \epsilon_{i2}, \quad (5)$$

$$Y_i = \alpha_3 + \beta_3 T_i + \gamma M_i + \xi_3^\top X_i + \epsilon_{i3}, \quad (6)$$

and taking the product of the coefficient on the treatment variable in the first model with the coefficient on the mediator model in the second, $\beta_2\gamma$. Sans explicit recognition of the role the sequential ignorability assumption plays, this is well known (MacKinnon et al. 2007; Baron and Kenny 1986). When outcome variables are binary and mediators continuous, the product of coefficients (or some transformation of them) does not correspond to average causal mediation effects (Imai et al. 2010a; Pearl 2011) despite advice to the contrary (MacKinnon et al. 2007; Kenny 2008). Similarly, the product of slope coefficients cannot be used when the mediator is binary and a non-linear model is used (probit/logit) (Li et al. 2007; Imai et al. 2010a). The fact that methods developed for linear models do not extend to the use of non-linear models in the context of mediation analysis is broadly accepted (Kohler et al. 2011).[2]

Regarding the second limitation, the blind application of earlier methods without respect to the non-randomization of the mediator has led some to advise abandoning the search for mechanisms (Bullock et al. 2010). Alternatively, we suggest conducting sensitivity analyses, which we show how to do using our package in section 4 or adopting different experimental designs (Imai et al. 2009). Sensitivity analysis allows the analyst to state how an estimated quantity would change for different degrees of violation of the key identification assumption (Rosenbaum 2002). Because the sequential ignorability assumption can never be tested directly, sensitivity analysis is a key component of conducting *causal* mediation analysis.

3 Causal Mediation Analysis

3.1 The algorithm

The **mediation** package calculates the average mediation and direct effects by simulating predicted values of the mediator/outcome variable, that we do not observe, and

2. A related procedure is discussed in Buis (2011). The main differences are that their procedure requires dichotomous outcome variables, represents quantities of interests in terms of log-odds, and does not provide a method for sensitivity analysis to violations of the key identification assumption (see below).

then calculating the appropriate quantities of interest (average causal mediation, direct and total effects, as well as the proportion of the total effect that is mediated). We implement the parametric algorithm described in (Imai et al. 2010a).

Algorithm 1 (Parametric Inference)

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 three steps for each draw of model parameters;*

1. *Simulate the potential values of the mediator*
2. *Simulate the potential outcomes given the simulated values of the mediator*
3. *Compute quantities of interest (average causal mediation effect, average direct effect, average total effect)*

Step 4: *Compute summary statistics such as point estimates (average) and confidence intervals*

The structure of the algorithm is a function of the theoretical results linking the sequential ignorability assumption and the mediation effect, and the calculation of uncertainty estimates is based on the quasi-Bayesian Monte Carlo approximation of King et al. (2000). While this algorithm can be applied to a range of parametric or semi-parametric statistical model, mediation currently implements the procedure for the common cases where OLS, probit, or logit models are used. Users wishing more flexibility, such as the use of quantile regressions, at this point would need to use the R package with the same name, which implements a non-parametric bootstrap version of algorithm 1.[3] Sampling weights may also be used.

4 Sensitivity Analysis

The preceding section discussed how to analyze data under the sequential ignorability assumption. We cannot test this assumption with the data and hence a sensitivity analysis should be conducted. The sensitivity analysis investigates how robust the results are to the violation of the SI assumption. The exact form of the sensitivity analysis will depend on the types of parametric models used for the mediator and outcome models. We briefly describe each of the three cases covered by the **mediation** package (continuous mediator and outcome, continuous mediator and binary outcome, and binary mediator and continuous outcome).[4]

3. Imai et al. (2010b) illustrate the use of this software.
4. The following assumes a binary (0/1) treatment variable.

4.1 Continuous mediator and outcome variables

When linear models are used for the mediator and outcome variables sensitivity analysis is based the linear structural equation model (LSEM) in equations 5 and 6. Here a violation of the SI assumption leads to a correlation between ϵ_{i2} and ϵ_{i3}, which we denote by ρ. ρ is equal to zero under SI . As shown in Imai et al. (2010c) the ACME can be expressed as a function of ρ using identifiable parameters. The procedure is essentially an application of the iterative feasible generalized least square algorithm of the seemingly unrelated regression (Zellner 1962). Asymptotic variance of the estimated average causal mediation effects can be expressed with the Delta method and the confidence intervals can be constructed. Additional mathematical details are provided in Imai et al. (2010c).

4.2 Binary mediator and continuous outcome variables

If the mediator is modeled as a probit regression with an independently and identically distributed error term with standard normal distribution, and a linear normal regression with error variance equal to σ_3^2 for a continuous outcome variable, then a sensitivity analysis is available Imai et al. (2010a). Assuming that the two error terms jointly follow a bivariate normal distribution with mean zero and covariance $\rho\sigma_3$ then the correlation between the two error terms, ρ, is the sensitivity parameter. Under these assumptions the causal mediation effects can be written in terms of consistently estimated model parameters and a fixed value of ρ. Uncertainty estimates are computed based on the quasi-Bayesian approach. Mathematical details are provided in Imai et al. (2010a). As shown below, graphing the ACME as a function of ρ is straightforward.

4.3 Continuous mediator and binary outcome variables

In situations with a binary outcome and continuous mediator a sensitivity analysis is also available in the mediation package. Here the outcome model is assumed to be a probit regression which allows us to assume the error terms are jointly normal with a possibly non-zero correlation ρ. The ACME can once again be written as a function of identifiable parameters and confidence intervals are once again approximate with the quasi-Bayesian approach discussed previously. Mathematical details are provided in Imai et al. (2010a).

4.4 Alternative Interpretations Based on R^2

Expressing the ACME as a function of ρ is simple. However, interpretation of the magnitude of this correlation coefficient may be difficult. An alternative approach is to express the ACME as a function of R^2s which will capture how important a confounder is for explaining the mediator or outcome variable. If there is an omitted confounder, U_i, then the error term will be a function of this confounder, yielding a decomposition of our error term $\epsilon_{ij} = \lambda_j U_i + \epsilon'_{ij}$ for $j = 2, 3$ (i.e., for the mediator model and outcome model). With this set-up, ρ can be expressed as a function of the *proportions of previously unex-*

plained variances in the mediator and outcome regressions[5], or based on the *proportions of original variances* that are explained by the unobserved confounder in the mediator and outcome regressions[6]. The relationship between the ACME and R^2 parameters can then be expressed as the product of the R^2 parameters for the mediator and outcome variables. For the case of previously unexplained variances this is $\rho = \text{sgn}(\lambda_2\lambda_3)R^*_M R^*_Y$ and for original variances this is $\text{sgn}(\lambda_2\lambda_3)\widetilde{R}_M\widetilde{R}_Y/\sqrt{(1-R^2_M)(1-R^2_Y)}$. In both cases ρ is a function of the product of unexplained variance measures. Below we show how the medsens function reports the values of $R^{*2}_M R^{*2}_Y$ or $\widetilde{R}^2_M\widetilde{R}^2_Y$ such that the ACME is 0. Because these are products, this critical point can occur across a range of values.[7]

When the mediator or the outcome variable is binary, we use the pseudo-R^2 of McKelvey and Zavoina (1975). For example, in the binary mediator case, we redefine $\widetilde{R}^2_M = \{1 - \text{Var}(\epsilon'_{i2})\}/\{\text{Var}(\widehat{M^*_i}) + 1\}$ and $R^2_M = \text{Var}(\widehat{M^*_i})/\{\text{Var}(\widehat{M^*_i}) + 1\}$ in the above formula where $\widehat{M^*_i}$ represents the predicted value of the latent mediator variable for the probit regression. Thus, in all cases considered here, we can interpret ρ using two alternative coefficients of determination. This value can then be used to compare across studies or evaluated in reference to subject specific knowledge about the likely magnitude of effect from the confounding variable.

5 The `mediate` command

5.1 Syntax

`mediate (`*`model depvar varlist`*` ) (  `*`model depvar varlist`*`)` [*if*] [*in*] ,
`mediate(`*varname*`) treat(`*varname*`##)` [`vce(vcetype) sims(#) seed(#)`
]

5.2 Structure and Options

In the first set of parentheses the user specifies the model for the mediator variable and in the second set the model for the outcome variable. Available model types are OLS regression (regress), probit (probit) and logit (logit). If there is to be a restriction on observations, this will apply to both models and is done with the standard *if* or *in* syntax. mediate(*varname*) is required and specifies the mediating variable to be used in the analysis. treat(*varname*##) is required and specifies the treatment variable used in the analysis, where the numbers following the treatment name are values to be used for the control and treatment conditions, respectively. By default these are set to 0 and

5. $R^{*2}_M \equiv 1 - \text{Var}(\epsilon'_{i2})/\text{Var}(\epsilon_{i2})$ and $R^{*2}_Y \equiv 1 - \text{Var}(\epsilon'_{i3})/\text{Var}(\epsilon_{i3})$
6. $\widetilde{R}^2_M \equiv \{\text{Var}(\epsilon_{i2}) - \text{Var}(\epsilon'_{i2})\}/\text{Var}(M_i)$ and $\widetilde{R}^2_Y \equiv \{\text{Var}(\epsilon_{i3}) - \text{Var}(\epsilon'_{i3})\}/\text{Var}(Y_i)$
7. $\text{sgn}(\lambda_2\lambda_3)$ captures whether the coefficient on the omitted variable is similar or different for the mediator and outcome equations. Hence linking these R^2 measures directly back to the ACME (which can be written as a function of ρ) requires that researchers specify the direction (positive/negative) of confounding for both models. R^2_M and R^2_Y are the coefficients of determination for the mediator and outcome regressions.

1. In addition there are the following options:

sims(#) specifies the number of simulations to run for the quasi-Bayesian approximation of parameter uncertainty. The default value is 1000. Higher values will increase the computational time.

seed(#) sets the random number seed for precise replicability though with sufficient sims results will be very similar. The default value is the random seed Stata draws when starting a session.

vcetype(vcetype) allows users to specify how the standard errors will be calculated. Clustering is available.

Sampling weights may be used by specifying [pweight=weight] after the models.

5.3 Saved results

r(delta0) r(delta1) - point estimates for average causal mediation effects under the control and treatment conditions.

r(delta0hi) r(delta0lo) r(delta1hi) r(delta1lo) - confidence intervals for average causal mediation effects.

r(tau) - point estimate for total effect

r(tauhi), r(taulo) - confidence interval for total effect

r(zeta0) r(zeta1) - point estimates for average direct effect under the control and treatment conditions

r(zeta0hi), r(zeta0lo), r(zeta1hi), r(zeta1lo) - confidence intervals for average direct effects.

5.4 Example

To illustrate the use of the mediate function we produce 2000 observations of simulated data. To do this we utilize a system of linear structural equations (LSEMs) given in Equations 5 and 6, fixing the structural parameters $\alpha_2, \alpha_3, \beta_2, \beta_3, \gamma, \xi_2, \xi_3$. For simplicity all are set to .25. Below we implement the case for a continuous mediator and outcome variable using OLS regression for both models, in which case the ACME is equivalent to $\beta_2\gamma$.

```
. ***********
. *Create simulated data
. ***********
. clear all
. set seed 312789
. local n 2000
. set obs  `n´
obs was 0, now 2000
```

```
. *Population Values
. local alpha_2  .25
. local alpha_3  .25
. local beta_2  .25
. local beta_3  .25
. local gamma  .25
. local x_beta .25
. *Draw realizations of error terms and pre-treatment covariate x assuming no c
> orrelation
. matrix m = (0,0,0)
. matrix sd = (1,1,1)
. drawnorm e1 e2 x, n(`n´) means(m) sds(sd)
.
. *Generate realizations of treatment (T), mediator (M), and outcome (Y) variab
> les
. gen T = round(uniform(), 1)
. gen M = `alpha_2´ + `beta_2´*T + `x_beta´*x + e1
. gen Y =  `alpha_3´ + `beta_3´*T + `gamma´*M + `x_beta´*x + e2
.
. *Conduct mediation analysis
. mediate (regress M T x) (regress Y T M x) , treat(T) mediate(M) sims(1000)
Using 0 and 1 as treatment values
```

Source	SS	df	MS		
Model	157.797161	2	78.8985804		
Residual	2081.76275	1997	1.04244504		
Total	2239.55991	1999	1.12034013		

Number of obs = 2000
F(2, 1997) = 75.69
Prob > F = 0.0000
R-squared = 0.0705
Adj R-squared = 0.0695
Root MSE = 1.021

M	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	.2620606	.0456963	5.73	0.000	.1724431	.3516781
x	.2468286	.0231428	10.67	0.000	.201442	.2922152
_cons	.2428821	.0320858	7.57	0.000	.179957	.3058072

Source	SS	df	MS		
Model	376.062154	3	125.354051		
Residual	1998.18054	1996	1.00109245		
Total	2374.24269	1999	1.1877152		

Number of obs = 2000
F(3, 1996) = 125.22
Prob > F = 0.0000
R-squared = 0.1584
Adj R-squared = 0.1571
Root MSE = 1.0005

Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	.2632512	.045148	5.83	0.000	.174709	.3517934
M	.2372682	.0219291	10.82	0.000	.1942618	.2802746
x	.2604436	.0233161	11.17	0.000	.2147171	.30617
_cons	.2291991	.0318908	7.19	0.000	.1666563	.2917419

Effect	Mean	[95% Conf.	Interval]
ACME1	.0624004	.0410497	.0888912
ACME0	.0624004	.0410497	.0888912
Direct Effect 1	.2638839	.17911	.351345

```
Direct Effect 0 |   .2638839        .17911       .351345
Total Effect    |   .3262843      .2238124      .4363766
```

Here Stata reports the results from the two regression models and then the summary estimates of the mediation, direct, and total effects. Estimates for both $\delta(1)$, *ACME1*, and $\delta(0)$, *ACME0* are given.[8] The average effect of the treatment variable on the outcome that operates through the mediator is .067. The estimates of the direct effect, $\bar{\zeta}(t)$, *Direct Effect 1* and *Direct Effect 0*, are equal to .28. Finally, mediate also reports the average treatment effect, *Total Effect*. As expected, under the sequential ignorability assumption the estimate of the ACME is nearly identical to the product of coefficients method, even though the mediate function uses Algorithm 1. If an analyst had a binary mediator or outcome variable then instead of using the regress function a probit or logit model could be used instead. As mentioned above, the product of coefficients in this case will not correspond to the ACME. In such cases we recommend using a probit model because this permits sensitivity analyses which we discuss next.

6 The `medsens` command

6.1 Syntax

`medsens ( model` *depvar varlist*`) (  model` *depvar varlist* `)` [*if*] `,`
`mediate(`*varname*`) treat(`*varname*`)` [`graph sims(#) seed(#) eps(#)`]

6.2 Structure and Options

The first part of the `medsens` function follows the format of the `mediate` function in that it gives the required regression models. Similarly, mediate(*varname*) specifies the mediating variable to be used in the analysis and treat(*varname*) specifies the treatment variable used in the analysis. Values of 0 and 1 are used. In addition there are several options:

sims(#)- specifies the number of simulations to run. The default value is 100. For final production runs this should be set higher (500) but note this will take longer, especially for models with a binary mediator.

seed(#)- sets the random number seed for precise replicability though with sufficient sims results will be very similar. The default value is the random seed Stata draws when starting a session.

eps(#)- convergence tolerance parameter for the iterative FGLS. Only used when both the mediator and outcome models are linear. The default value is .01. Typically

8. In the case with all linear models, and no treatment/mediator interaction (an option not currently supported by the Stata mediate command), these estimates will be identical. With a binary outcome the estimates can differ due to the non-linear link functions.

users will not change this, and if so only will decrease it.

graph- if specified, produces a graph of the results with the confidence intervals. Alternatively users can use saved results to produce a graph of their own.

6.3 Saved results

r(errcr)- the ρ (the correlation in error terms) at which the ACME= 0

r(r2s_thresh)- proportions of residual variance in the mediator and outcome explained by the hypothesized unobserved confounder

r(r2t_thresh)- the proportions of total variance in the mediator and outcome explained by the hypothesized unobserved confounder

6.4 Example

We conduct sensitivity analyses based on the previous empirical example. In each case the `medsens` function is used, which automatically detects which type of sensitivity analysis should be conducted. The value of rho where the ACME is 0, as well as the sensitivity to both types of R^2 expressions, is provided. In addition, information required to graph the ACME as a function ρ is provided.

```
. **Run Sensitivity Analysis
. medsens (regress M T x) (regress Y T M x) , treat(T) mediate(M) sims(100)
```

Source	SS	df	MS		
Model	157.797161	2	78.8985804	Number of obs	= 2000
				F(2, 1997)	= 75.69
Residual	2081.76275	1997	1.04244504	Prob > F	= 0.0000
				R-squared	= 0.0705
				Adj R-squared	= 0.0695
Total	2239.55991	1999	1.12034013	Root MSE	= 1.021

M	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
T	.2620606	.0456963	5.73	0.000	.1724431	.3516781
x	.2468286	.0231428	10.67	0.000	.201442	.2922152
_cons	.2428821	.0320858	7.57	0.000	.179957	.3058072

Source	SS	df	MS		
Model	376.062154	3	125.354051	Number of obs	= 2000
				F(3, 1996)	= 125.22
Residual	1998.18054	1996	1.00109245	Prob > F	= 0.0000
				R-squared	= 0.1584
				Adj R-squared	= 0.1571
Total	2374.24269	1999	1.1877152	Root MSE	= 1.0005

Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
M	.2372682	.0219291	10.82	0.000	.1942618	.2802746
T	.2632512	.045148	5.83	0.000	.174709	.3517934
x	.2604436	.0233161	11.17	0.000	.2147171	.30617
_cons	.2291991	.0318908	7.19	0.000	.1666563	.2917419

```
Sensitivity results
------------------------------------------------------------------------------
        Rho at which ACME = 0                 |       .2354
        R^2_M*R^2_Y* at which ACME = 0:       |       .0554
        R^2_M~R^2_Y~ at which ACME = 0:       |       .0434
------------------------------------------------------------------------------
95% Confidence interval

.
. set scheme sj
. twoway rarea  _med_updelta0 _med_lodelta0 _med_rho, bcolor(gs14) || line  _me
> d_delta0 _med_rho , lcolor(black) ytitle("ACME") xtitle("Sensitivity paramete
> r: {&rho}") legend(off) title("ACME({&rho})")
```

The results show that for the point estimate of the ACME to be 0 the correlation between ϵ_{i2} and ϵ_{i3} must be approximately .26. Alternatively, the product of R^2's measures of sensitivity for the mediator and outcome models, for either the residual and total variance, may be examined. For example, an omitted confounder must explain 35% of remaining variance in the mediator and 19% of the remaining variance in the outcome, $.35 \times .19 \approx .068$, in order for the ACME to be 0. Similar calculations can be done for sensitivity with respect to total variation, where the product of R^2's is .05.

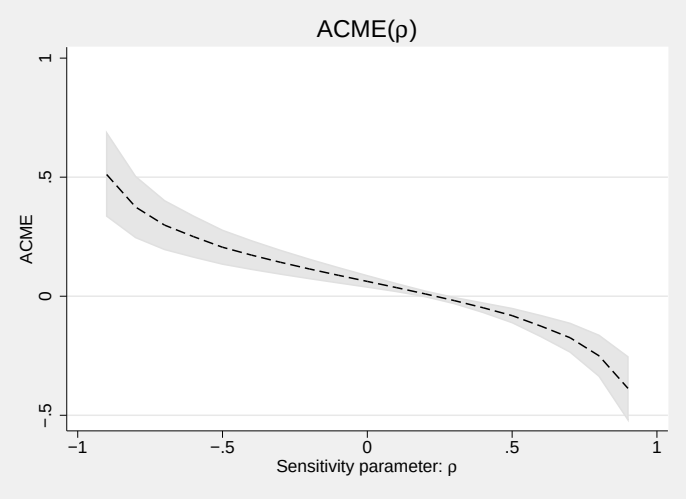

Figure 1: Average causal mediation effect as a function of degree of violation of sequential ignorability assumption.

7 References

Baron, R. M., and D. A. Kenny. 1986. The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology* 51(6): 1173–1182.

Buis, M. 2011. Direct and indirect effects in a logit model. *The Stata Journal* 10(1): 11–29.

Bullock, J., D. Green, and S. Ha. 2010. Yes, But What's the Mechanism? (Don't Expect an Easy Answer). *Journal of Personality and Social Psychology* 98(4): 550–558.

Imai, K., L. Keele, and D. Tingley. 2010a. A General Approach to Causal Mediation Analysis. *Psychological Methods* 15(4): 309–334.

Imai, K., L. Keele, D. Tingley, and T. Yamamoto. 2010b. *Advances in Social Science Research Using R (ed. H. D. Vinod)*, chap. Causal Mediation Analysis Using R, 129–154. Lecture Notes in Statistics, New York: Springer.

———. 2011. Unpacking the Black Box: Learning about Causal Mechanisms from Experimental and Observational Studies. *American Political Science Review* .

Imai, K., L. Keele, and T. Yamamoto. 2010c. Identification, Inference, and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science* 25(1): 51–71.

Imai, K., D. Tingley, and T. Yamamoto. 2009. Experimental Designs for Identifying Causal Mechanisms. Technical report, Department of Politics, Princeton University. Available at http://imai.princeton.edu/research/Design.html.

Kenny, D. 2008. Mediation with Dichotomous Outcomes . http://www.davidakenny.net/doc/dichmed.doc.

King, G., M. Tomz, and J. Wittenberg. 2000. Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science* 44: 341–355.

Kohler, U., K. Karlson, and A. Holm. 2011. Comparing coefficients of nested nonlinear probability models. *The Stata Journal* .

Li, Y., J. Schneider, and D. Bennett. 2007. Estimation of the mediation effect with a binary mediator. *Statistics in Medicine* 26: 3398–3414.

MacKinnon, D., C. Lockwood, C. Brown, W. Wang, and J. Hoffman. 2007. The intermediate endpoint effect in logistic and probit regression. *Clinical Trials* 4: 499–513.

MacKinnon, D., G. Warsi, and J. Dwyer. 1995. A simulation study of mediated effect measures. *Multivariate Behavioral Research* 30: 41–62.

McKelvey, R. D., and W. Zavoina. 1975. A Statistical Model for the Analysis of Ordinal Level Dependent Variable. *Journal of Mathematical Sociology* 4: 103–120.

Pearl, J. 2001. Direct and Indirect Effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, ed. M. Kaufmann, 411–420. San Francisco, CA.

———. 2011. The Causal Mediation Formula – A Guide to the Assessment of Pathways and Mechanisms. *Prevention Science* forthcoming.

Robins, J., and S. Greenland. 1992. Identifiability and Exchangeability For Direct and Indirect Effects. *Epidemiology* 3(2): 143–155.

Rosenbaum, P. R. 2002. *Observational Studies*. 2nd ed. New York: Springer-Verlag.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 57: 348–368.