Mediation Analysis for Count and Zero-Inflated Count Data without Sequential Ignorability

Zijian Guo*, Dylan S. Small
Department of Statistics, University of Pennsylvania Wharton School, Philadelphia, PA 19104, USA

Stuart A. Gansky and Jing Cheng**
Division of Oral Epidemiology & Dental Public Health, University of California, San Francisco, CA 94143, USA

*email: zijguo@wharton.upenn.edu
**email: jing.cheng@ucsf.edu

SUMMARY: Count or zero-inflated count data are common in many studies. Most existing mediation analysis approaches for count data assume sequential ignorability of the mediator. This is often not plausible because the mediator is not randomized so that there are unmeasured confounders associated with the mediator and the outcome. In this paper, we consider causal methods based on instrumental variable (IV) approaches for mediation analysis for count data possibly with a lot of zeros that do not require the assumption of sequential ignorability. We first define the direct and indirect effect ratios for those data, and then propose estimating equations and use empirical likelihood to estimate the direct and indirect effects consistently. A sensitivity analysis is proposed for the violations of the IV exclusion restriction assumption. Simulation studies demonstrate that our method works well for different types of outcomes under different settings. Our method is applied to a randomized caries prevention trial and a study of the effect of a massive flood in Bangladesh on children’s diarrhea.

KEY WORDS: Count; Direct Effect; Empirical Likelihood; Estimating Equation; Indirect Effect; Instrumental Variable; Poisson; Negative Binomial; Sensitivity Analysis; Type A Distribution.

This paper has been submitted for consideration for publication in Biometrics
1. Introduction

In many studies, the intervention is designed to change intermediate variables under the hypothesis that the change in those intermediate variables will lead to improvement in the final outcomes (MacKinnon and Luecken, 2011). In these studies, in addition to the overall effect of the intervention on the outcome in the end, researchers would like to know whether and how much the intervention affects the outcome through the measured intermediate variables (mediators) as designed (indirect effect) vs. “direct” intervention effects on the outcome not through the proposed mediators but involving other unknown pathways. Knowing those effects helps us to understand the working mechanism of an intervention and to tailor specific intervention components for future research and applications in specific populations.

Standard mediation approaches since Baron and Kenny (Baron and Kenny, 1986; Cole and Maxwell 2003; MacKinnon 2008), such as regression, path and structural equation model (SEM) among others, assume sequential ignorability of the intervention and mediator (randomization of the intervention and mediators) for a causal interpretation on the direct and mediation effects of the intervention on the outcome. Although the ignorability assumption for the intervention is usually reasonable because of randomization, the ignorability assumption for mediators can be questionable because mediators cannot be randomized by researchers so that there may be unmeasured confounders between the mediators and the outcome. Recently developed causal methods (Robins and Greenland 1992; Pearl 2001; Rubin 2004; Ten Have et al. 2007; van der Laan and Petersen 2008; Sobel 2008; VanderWeele and Vansteelandt 2009, 2010; Elliott et al. 2010; Imai et al. 2010; Jo et al. 2011; Daniels et al. 2012; Stayer et al. 2014) adopt the potential outcome framework and make different assumptions about the intervention and mediator to achieve a causal interpretation of the indirect (mediation) and direct effect of the intervention on the outcome. Those causal methods may replace the ignorability assumption of mediators with other assumptions,
such as the no interaction assumption between intervention and mediator among others. Most standard and causal approaches focus on continuous and/or binary mediators and outcomes. However, the outcome variable in many studies is often a count following a Poisson or Negative Binomial distribution, or a zero-inflated count that has a higher probability of being zero than expected under a Poisson or Negative Binomial. Examples include the number of healthcare visits, length of hospital stay, number of decayed, missing and filled tooth surfaces, and many quality of life scale scores.

In this paper, we propose to develop a new mediation analysis based on the instrumental variable (IV) approach for count and zero-inflated count data, when there is a concern about unmeasured confounding such that the assumption of sequential ignorability might fail. When confounding is of concern in a study, IV methods are very helpful for obtaining accurate estimates for treatment effects by adjusting for both unmeasured and measured confounders when a valid IV can be found (Angrist et al., 1996). Angrist and Krueger (1991) provide a good review of applications of the IV method. Methods based on the IV approach have been proposed by investigators (Ten Have et al., 2007; Albert, 2008; Dunn and Bentall, 2007; Small, 2012) using the randomization interacted with baseline covariates as IVs, but those methods only consider linear models for continuous outcomes. When the outcome model is linear, two stage least squares (2SLS) and two stage residual inclusion (2SRI) can estimate the direct and indirect effects well when there is a valid IV. We will show that the two stage method can give a biased estimate when the mediator is binary and the outcome model is a count or zero-inflated count model; we will first define the direct and indirect treatment effects in our context and then develop a consistent estimator based on estimating equations and empirical likelihood. We will use the random assignment interacted with baseline covariates as IVs to account for both measured and unmeasured confounding. Since the randomized treatment itself is not used as the IV, we are able to estimate the
direct and indirect effect of the treatment on the outcome of interest. Although we focus on count and zero-inflated count outcomes in this paper, the method can be generalized to other types of outcomes.

The paper is organized as follows. In Section 2, we introduce notation, framework, and the direct and indirect treatment effects of interest for count and zero-inflated count data. In Section 3, we introduce the IV and two stage approaches, and our new method, and provide a sensitivity analysis method. In Section 4, we present simulation results. In Section 5, we apply our method to two real studies. In Section 6, we discuss future research directions. The proofs and further simulation studies are provided in the supplementary materials.

2. Notation, Framework and Causal Effects of Interest

Notation We adopt the potential (counterfactual) outcome framework (Neyman, 1923; Rubin, 1974) and use $Z_i = z$ ($z = 0$ or 1) for the randomly assigned treatment for subject $i$; let $M_i^z$ denote the potential value of a mediator under treatment $z$; use $Y_i(z, m)$ to denote the potential outcome subject $i$ would have under the treatment $z$ and mediator $m$, and $Y_i(z, M_i^z)$ for potential outcome subject $i$ would have under $Z_i = z$ ($M_i^z$ would be at its “natural” level under $z$). We let $U_i$ denote unobserved confounders, $X_i$ denote observed baseline covariates, and $X_i^{IV}$ denote a subset of baseline covariates to construct IVs.

Direct and Indirect Effects of Interest For count and zero-inflated count outcomes, we are particularly interested in the controlled and natural direct and indirect ratios for comparing average potential outcomes at different levels of randomization and mediator.

\[
\text{Controlled effect ratio:} \quad \begin{align*}
\text{direct} & (z \text{ vs. } z^*; m, x, u) & \frac{\mathbb{E}(Y(z, m) | x, u)}{\mathbb{E}(Y(z^*, m) | x, u)}, \\
\text{indirect} & (m \text{ vs. } m^*; z, x, u) & \frac{\mathbb{E}(Y(z, m) | x, u)}{\mathbb{E}(Y(z, m^*) | x, u)} \cdot \frac{\mathbb{E}(Y(z^*, m) | x, u)}{\mathbb{E}(Y(z^*, m^*) | x, u)};
\end{align*}
\]

\[
\text{Natural effect ratio:} \quad \begin{align*}
\text{direct} & (z \text{ vs. } z^*; M^z, x, u) & \frac{\mathbb{E}(Y(z, M^z) | x, u)}{\mathbb{E}(Y(z^*, M^z) | x, u)}, \\
\text{indirect} & (M^z \text{ vs. } M^{z*}; z, x, u) & \frac{\mathbb{E}(Y(z, M^z) | x, u)}{\mathbb{E}(Y(z, M^{z*}) | x, u)} \cdot \frac{\mathbb{E}(Y(z^*, M^z) | x, u)}{\mathbb{E}(Y(z^*, M^{z*}) | x, u)};
\end{align*}
\]
A ratio of 1 indicates no effect. The controlled direct effect sets the mediator at a fixed value ($m$) and the natural direct effect sets the mediator at its “natural” level that would be achieved under treatment assignment $z$ ($M^z$). The natural indirect effect ratio tells us the ratio of average outcomes under treatment $z$ that would be observed if the mediator would change from the value under a treatment $z$ ($M^z$) to the value under another treatment $z^*$ ($M^{z^*}$). The expectations in (1) and (2) are taken over the conditional distribution of the conditional expectation of the potential outcome. However, in the natural effect ratio, since the mediator is random, the expectations in (3) and (4) are taken over the conditional joint distribution of the mediator and the potential outcome corresponding to the mediator. We will discuss below the settings in which the controlled and natural effects are identified.

**The Model Setting** We consider the following general potential outcome model:
\[
f\{ \mathbb{E}(Y(z, m) | x, u) \} = \beta_0 + \beta_z z + \beta_m m + \beta_x x + \beta_u u, \tag{5}\]

where $f$ is a link function. The generalization of Model (5) to include the interaction terms $z \times x, z \times m$ and $x \times m$ will be discussed in Section 6. For a Poisson or negative binomial count outcome, a log link function will be used in the model; For a zero-inflated count outcome, we consider a Neyman Type A distributed outcome (Dobbie and Welsh, 2001), $y = \sum_{k=1}^{N} y_k$, where
\[
N | x, z, m, u \sim \text{Poisson} (\exp (\gamma_0 + \gamma_z z + \gamma_m m + \gamma_x x + \gamma_u u));
\]
\[
y_k | x, z, m, u \sim \text{Poisson} (\exp (\lambda_0 + \lambda_z z + \lambda_m m + \lambda_x x + \lambda_u u)).
\]

Without loss of generality, we will assume that $\mathbb{E}(\exp(\beta_u u)) = 1$ for the count and zero-inflated count with a log link function.

**Controlled Effect Ratios** Given Model (5) with a log link function, we have:
\[
\frac{\mathbb{E}(Y(z, m) | x, u)}{\mathbb{E}(Y(z^*, m) | x, u)} = \exp(\beta_z (z - z^*)), \quad \frac{\mathbb{E}(Y(z, m) | x, u)}{\mathbb{E}(Y(z, m^*) | x, u)} = \exp(\beta_m (m - m^*)). \tag{6}\]

Therefore, estimating the controlled effect ratios is equivalent to estimating $\beta_z$ and $\beta_m$. 
Natural Effect Ratios

For natural effect ratios, Model (5) becomes
\[
f \{ \mathbb{E} (Y(z, M^*)|x, u) \} = \beta_0 + \beta_z z + \beta_m M^* + \beta_x x + \beta_u u. \tag{7}
\]
and we further consider a model for the mediator:
\[
h (\mathbb{E} (M^*|x, u)) = \alpha_0 + \alpha_z z^* + \alpha_x x + \alpha_{IV} z^* \cdot x^{IV} + \alpha_u u, \tag{8}
\]
where \( h \) is a link function, e.g., identity and logit functions for continuous and binary mediators respectively. Then the natural indirect effect ratio with a continuous mediator will be
\[
\frac{\mathbb{E} (Y(z, M)|x, u)}{\mathbb{E} (Y(z, M^*)|x, u)} = \exp \left( \beta_m \alpha_z (z - z^*) + \beta_m \alpha_{IV} x^{IV} (z - z^*) \right), \tag{9}
\]
and with a binary mediator:
\[
\frac{\mathbb{E} (Y(z, M)|x, u)}{\mathbb{E} (Y(z, M^*)|x, u)} = \frac{P(M = 1|x, u) \exp(\beta_m) + P(M = 0|x, u)}{P(M^* = 1|x, u) \exp(\beta_m) + P(M^* = 0|x, u)}. \tag{10}
\]
The natural direct effect ratio for both continuous and binary mediators will be the same as the controlled direct effect ratio:
\[
\frac{\mathbb{E} (Y(z, M)|x, u)}{\mathbb{E} (Y^*(z, M^*)|x, u)} = \exp (\beta_z (z - z^*)). \tag{11}
\]
The proofs of (9), (10) and (11) are provided in subsection 1.1 of the web-based supplementary materials. For a continuous mediator, the natural direct and indirect effect ratios (9) and (11) are identifiable given that the parameters \( \beta_z, \beta_m, \alpha_z \) and \( \alpha_{IV} \) can be estimated consistently. However, the natural indirect effect ratio for a binary mediator depends on the values of unmeasured \( U \) in (10) and is not identifiable.

3. The Instrumental Variable Approach

When there is a concern on unmeasured confounder \( U \), the instrumental variable (IV) approach is a popular technique for dealing with unmeasured confounding not addressed by regular regression and propensity score methods. In the context of mediation analysis, a valid IV is a variable that, given the measured baseline variables: (1) affects the value of the
mediator; (2) is independent of the unmeasured confounders; and (3) does not have a direct
effect on the outcome other than through its effect on the mediator.

Mediation methods based on the IV approach have been proposed by investigators (Ten
Have et al., 2007; Albert, 2008; Dunn and Bentall, 2007; Small, 2012) for continuous out-
comes, where baseline covariates interacted with random assignment are used as instrumental
variables in linear models. In this study, we will also use baseline covariates interacted
with random assignment \(Z \times X^{IV}\) as instrumental variables for mediation analysis but in
nonlinear models. In the setting of a randomized trial with noncompliance, the randomization
\(Z\), is often used as an instrument (e.g., Sommer and Zeger, 1991; Greevy et al., 2004) under
the assumption that the randomization has no direct effect on the outcome. In our setting,
we assume that the randomization is complied with and allow for the treatment itself to
have a direct effect, but use the IV \(Z \times X^{IV}\) instead of \(Z\) to enable us to estimate both the
direct and indirect effects of the treatment on the outcome of interest.

To consistently estimate the direct and indirect (mediation) effects discussed in Section 2
without the commonly used sequential ignorability assumption, we will use \(Z \times X^{IV}\) as an
IV and assume:

(1) The treatment \(Z\) is randomized.

(2) The conditional distribution of the unmeasured confounding \(P(U|X)\) is the same for all
\(X\), which implies that \(U\) and \(X\) are independent.

(3) There is an interaction between randomized treatment \(Z\) and baseline covariate \(X^{IV}\)
predicting mediator \(M\) conditional on \(Z\) and \(X\).

(4) The interaction \(Z \times X^{IV}\) affects the outcome only through its effect on the mediator
\(M\), conditional on \(X\) and \(Z\). This assumption is often called the exclusion restriction
assumption and cannot be formally tested. A sensitivity analysis is proposed in Section
3.4 to see how the method behaves when the assumption does not hold.
Assumptions (1) and (2) imply that the instrument $Z \times X^{IV}$ is independent of the unmeasured confounder $U$; Assumption (3) states that $Z \times X^{IV}$ affects the value of the mediator $M$; Assumption (4) says that $Z \times X^{IV}$ does not have a direct effect on the outcome other than through its effect on the mediator. Hence, under these assumptions, the interaction $Z \times X^{IV}$ is a valid IV for the mediation analysis. Note that although the treatment $Z$ and instrumental variable $Z \times X^{IV}$ affect the mediator $M$, conditioning on $Z$, $M$ and $X$, the instrumental variable $Z \times X^{IV}$ has no direct effect on the outcome $Y$. Similar model assumptions are discussed in equation (11) in Jo (2002).

3.1 Two Stage Approach

When a linear model is a good fit for a continuous outcome, the two stage least square (2SLS) estimator provides consistent estimates for the parameters of interest when a valid IV is available. When the outcome is not continuous such that a linear model does not fit well, two-stage predictor substitution (2SPS) and two-stage residual inclusion (2SRI) (Nagelkerke et al., 2000; Terza et al., 2008) have been proposed to evaluate the treatment effect with an IV. For a binary treatment and outcome, the 2SPS approach fits a logistic regression of the treatment on the IV and covariates in the first stage; and then in the second stage uses the predicted treatment from the first stage to fit a logistic regression for the outcome with covariates. Because of the use of predicted treatment in the second stage, this approach is called two-stage predictor substitution. The first stage of 2SRI approach is the same as that of 2SPS, but in the second stage, instead of using predicted treatment, 2SRI uses the residual from the first stage regression along with observed treatment and covariates to model the outcome in a logistic regression. Cai et al. (2011) showed that 2SPS and 2SRI estimators are asymptotically biased for the complier average causal effect (CACE) when there is unmeasured confounding. In the supplement, we compare the 2SRI and 2SPS estimators with simulation studies. When the second stage linear model is a good fit for the
outcome, both 2SPS and 2SRI estimates are consistent (Wooldridge, 2010). When the second stage outcome model is non-linear, 2SRI is approximately unbiased while 2SPS is slightly biased for the Neymann Type A outcome given a linear first stage model for a continuous mediator (see Web Table 8); but neither 2SPS and 2SRI estimate is consistent given a non-linear first stage model for a binary mediator (see Web Table 7).

Because 2SRI performs better than 2SPS, we will examine the performance of 2SRI in the mediation analysis in the paper. With the IV \( Z \times X^{IV} \), the 2SRI fits the models in two stages. Stage I: 

\[
h\{E(M|z, x, z \times x^{IV})\} = \alpha_0 + \alpha_z z + \alpha_m m + \alpha_{IV} z \times x^{IV} + \alpha_x x,
\]

where \( h \) is a link function; Stage II, 

\[
f\{E(Y|m, z, x, \hat{r})\} = \beta_0 + \beta_z z + \beta_m m + \beta_x x + \beta_r \hat{r},
\]

where \( \hat{r} \) is the residual \( m - \hat{m} \) from Stage I and \( f \) is a link function such as log link for count data.

We can decompose \( U \) into two parts 

\[
U = \tau R + \delta,
\]

where \( R \) is the residual from the first stage, \( \delta \) is the population residual and \( E(\delta|R) = 0 \). For continuous outcomes, we have:

\[
Y(Z, M|X, U) = \beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u U,
\]

(12)

\[
= \beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R + \beta_u \delta
\]

For count outcomes, we have

\[
E(Y(Z, M)|X, U) = \exp(\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u U) \sim \exp(\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R + \beta_u \delta)
\]

(13)

\[
= \int \exp(\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R + \beta_u \delta) \, dP(\delta|Z, M, X, R)
\]

\[
= \exp(\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R) \int \exp(\beta_u \delta) \, dP(\delta|Z, M, X, R).
\]

Now consider a linear model and logit model for a normal and binary mediator respectively:

The normal mediator model

\[
M = \alpha_0 + \alpha_z Z + \alpha_x X + \alpha_{IV} Z \times X^{IV} + \alpha_u U + V.
\]

(14)

where \( V \) is random error and \( U \) is the unmeasured confounder with \((\alpha_u U + V, U)\) following bivariate normal distribution and is independent of \((X, Z, Z \times X^{IV})\). 2SRI fits a linear model
for \( M \) on \( Z, X, Z \times X^{IV} \), and the probability limit of estimator \( \alpha_j^* = \alpha_j \). Then the residual

\[
R = M - \left( \alpha_0 + \alpha_z Z + \alpha_x X + \alpha_{IV} Z \times X^{IV} \right) = \alpha_u U + V.
\]

Since \((\alpha_u U + V, U)\) is independent of \((X, Z, Z \times X^{IV})\), \( \delta \) is independent of \((X, Z, Z \times X^{IV})\).

Since \( \delta \) is independent of \( R \) and \( M \) is linear combination of \((X, Z, Z \times X^{IV})\) and \( R \), \( \delta \) is independent of \( R \) and \( M \), and hence \( \int \exp(\beta \delta) dP(\delta|Z, M, X, R) \) is a constant and the 2SRI estimator is consistent.

The binary mediator model

\[
M|X, Z, U \sim \text{Ber}\left( \frac{\exp(\alpha_0 + \alpha_z Z + \alpha_x X + \alpha_{IV} Z \times X^{IV} + \alpha_u U)}{1 + \exp(\alpha_0 + \alpha_z Z + \alpha_x X + \alpha_{IV} Z \times X^{IV} + \alpha_u U)} \right),
\]

(15)

2SRI fits a logit model for \( M \) on \( Z, X, Z \times X^{IV} \), and the probability limit of estimator \( \alpha_j^* \neq \alpha_j \). Then the residual

\[
R = M - \frac{\exp(\alpha_0^* + \alpha_z^* Z + \alpha_x^* X + \alpha_{IV}^* Z \times X^{IV})}{1 + \exp(\alpha_0^* + \alpha_z^* Z + \alpha_x^* X + \alpha_{IV}^* Z \times X^{IV})}.
\]

Now \( \int \exp(\beta \delta) dP(\delta|Z, M, X, R) \) is generally not a constant but a function depending on \( Z, M, X, R \), so the estimate will probably be biased.

However, if either stage is a linear model with normal error, 2SRI is a consistent estimator under some regularity conditions. Please see the proof in the supplementary materials.

**Proposition 1:** Under regularity conditions, the 2SRI estimator is consistent for (1) a count outcome (14) and normal mediator (14); (2) a normal outcome (13) and binary mediator (15); and (3) a normal outcome (13) and normal mediator (14).

3.2 Estimating Equations and Empirical Likelihood Approach (EE-EL)

In this section, we consider a different approach to consistently estimate the parameters of interest in nonlinear models with unmeasured confounding. We let \( g(w, \theta) = (g_1(w, \theta), \ldots, g_r(w, \theta))^T \) be estimating functions such that \( E\{g(w, \theta)\} = 0 \), where \( w = (z, x, m, y) \) and \( \theta = (\beta_0, \beta_z, \beta_m, \beta_x) \) are the parameters associated with the outcome model. We consider a set of estimating functions to combine information about the parameters and distribution.
Under Assumptions (1)-(4), we have \( E\{g(w, \theta)\} = 0 \). The proof is provided in the subsection 1.2 of the web-based supplementary materials. Equations in (17) include more estimating equations than parameters and consequently there will not typically be a solution that satisfies all the estimating equations. Qin and Lawless (1994) proposed to use Owens (1988, 1990) empirical likelihood approach when there are more estimating equations than parameters and showed that empirical likelihood provides asymptotically efficient estimates of the parameters (in the sense of Van der Vaart(1988) and Bickel, Klaassen, Ritov and Wellner(1993)) under the semiparametric model given by the estimating equations. Following their approach, we let \( p_i \) be the probability of data \( (Z_i, X_i, M_i, Y_i) \) being observed and maximize \( \prod_{i=1}^{n} p_i \) subject to the restrictions

\[
p_i \geq 0, \quad \sum_{i=1}^{n} p_i = 1, \quad \sum_{i=1}^{n} p_i g(w_i, \theta) = 0.
\]

It is equivalent to minimize

\[
I_E(\theta) = \sum_{i=1}^{n} \log \left( 1 + t^\top(\theta)g(w_i, \theta) \right),
\]

where \( t = (t_1, ..., t_r)^\top \) are Lagrange multipliers and are determined by \( \frac{1}{n} \sum_{i=1}^{n} \frac{g(w_i, \theta)}{1 + t^\top g(w_i, \theta)} = 0 \).

With the first four estimating equations \( g_1(w, \theta), ..., g_4(w, \theta) \) for \( \theta = (\beta_0, \beta_z, \beta_m, \beta_x)^\top \), the maximized empirical likelihood estimate (MELE) will be the solution to the estimating equations \( \sum_{i=1}^{n} g_j(w_i, \theta) = 0, \ j = 1, ..., 4 \) that minimize (17).

Qin and Lawless (1994) provided a way to combine estimating equations in the empirical likelihood. We carry out the computation of maximizing the empirical likelihood subject to the estimating equations being satisfied in two steps: (1) Fix \( \theta \), we will minimize (17) with
respect to $t$; and (2) Given $t$ from the first step, minimize (17) with respect to $\theta$. Qin and Lawless(1994) showed that the MELE for the estimating equations is consistent under some regularization conditions. Proposition 2 provides the theory that the EE-EL estimator is consistent under some mild regularity conditions.

**Proposition 2:** Assume that $\mathbb{E} \left( g(w, \theta_0) g^T(w, \theta_0) \right)$ is positive definite and the rank of $\mathbb{E} \left( \frac{\partial g(w, \theta)}{\partial \theta} \right)$ is of rank $p$ and $\left\| \frac{\partial^2 g(w, \theta)}{\partial \theta \partial \theta^T} \right\|$ can be bounded by some integrable function $G(w)$ in the neighborhood $\|\theta - \theta_0\|_2 \leq 1$ of the true value $\theta_0$. Let $\hat{\theta}$ denote the minimizer of (17), then
\[
\sqrt{n} (\hat{\theta} - \theta_0) \rightarrow N(0, V), \quad \text{where } V = \left( \mathbb{E} \left( \frac{\partial g}{\partial \theta} \right)^T \mathbb{E} (gg^T)^{-1} \mathbb{E} \left( \frac{\partial g}{\partial \theta} \right) \right)^{-1}.
\]

**Proof:** It suffices to verify the conditions in Theorem 1 in Qin and Lawless(1994). By the expression of $g(w, \theta)$, $g(w, \theta)$ and $\frac{\partial g(w, \theta)}{\partial \theta}$ are continuous in a compact neighborhood $\|\theta - \theta_0\|_2 \leq 1$ of the true value $\theta_0$. Hence $\|g(w, \theta)\|_3$ and $\left\| \frac{\partial g(w, \theta)}{\partial \theta} \right\|_2$ are bounded in this compact neighborhood $\|\theta - \theta_0\|_2 \leq 1$. $\frac{\partial^2 g(w, \theta)}{\partial \theta \partial \theta^T}$ is continuous in $\theta$ in a neighborhood $\|\theta - \theta_0\|_2 \leq 1$ of the true value $\theta_0$.

3.3 Sensitivity Analysis

When Assumptions (1)-(4) hold, $Z \times X^{IV}$ is a valid IV. When the IV $Z \times X^{IV}$ actually affects the outcome directly, then the exclusion restriction (ER) assumption (Assumption (4)) fails so that the estimators will be biased. Unfortunately there is no formal test on the ER assumption. However, note that besides the pathway through the mediator of interest, the treatment effect on the outcome through other pathways (intermediate variables) are included in the direct effect on the outcome. Examining the effects of the IVs on the other intermediate variables allows us to partially assess the plausibility of the ER assumption of the IVs (please see the web supplement for more detailed discussion). In this section, we
propose a sensitivity analysis to allow $Z \times X^{IV}$ to affect the outcome directly by a specified magnitude and then examine how the results will change. Specifically, we consider

$$g(E(Y_{z,m}|z,m,x,u)) = \beta_0 + \beta_z z + \beta_m m + \beta_x x + \beta_u u + \eta z \times x^{IV},$$

(18)

where $\eta$ is the sensitivity parameter for the direct effect of the IV on the outcome. When $\eta = 0$, the ER assumption holds. When $\eta \neq 0$, the ER assumption fails and $Z \times X^{IV}$ is not a valid IV. Higher values of $|\eta|$ mean more severe violation of the ER assumption. For a zero-inflated count following Neyman Type A distribution, we have $Y = \sum_{k=1}^{N_k} y_k$; where

$N|x,z,m,u \sim \text{Poisson} \left( \exp(\gamma_0 + \gamma_z z + \gamma_m m + \gamma_x x + \gamma_u u + \eta_1 z \times x^{IV}) \right)$;

$y_k|x,z,m,u \sim \text{Poisson} \left( \exp(\lambda_0 + \lambda_z z + \lambda_m m + \lambda_x x + \lambda_u u + \eta_2 z \times x^{IV}) \right)$,

we can represent this model in the form of (18) with $\eta = \eta_1 + \eta_2$. With a log link function in Model (18), we will adjust the outcome as:

$$y^{adj} = \frac{y}{\exp(\eta z \times x^{IV})},$$

and have a model for the adjusted outcome:

$$E\left(y^{adj}|x,z,m,u\right) = \exp(\beta_0 + \beta_z z + \beta_m m + \beta_x x + \beta_u u).$$

Then we can construct the estimating equations (17) by replacing $y$ with $y^{adj}$ in (17) and (17) and obtain the MELE estimator. We will examine how the sensitivity analysis works in the simulation study.

### 3.4 Multiple Mediators

In addition to settings with one mediator discussed above, in this section, we consider settings with multiple conditionally independent mediators. That is, conditioning on $z,x,z \times x^{IV},u$, the mediators are independent. The outcome model (5) is generalized as

$$g\{E(Y(z,m)|x,u)\} = \beta_0 + \beta_z z + \beta_m^T m + \beta_x x + \beta_u u.$$

(19)

We need at least the same number of instrumental variables as the number of mediators to construct estimating equations and then use the same approaches discussed above for effect
estimations. We discuss the model and estimating equations for two conditionally independent mediators in the appendix, which can be easily generalized to multiple mediators.

4. Simulation Study

In this section, we will examine the performance of the methods discussed above in finite samples. We consider outcomes that follow Poisson, negative binomial and Neyman Type A distributions with a binary mediator, a normally distributed mediator, and multiple mediators respectively. The randomized treatment \( Z \) was generated with \( P(Z_i = 1) = 0.5 \). We consider one or two (standard normal and binary) covariates, and an unmeasured confounder \( U \) with \( \mathbb{E}(\exp(U)) = 1 \). Mediators and outcomes were generated based on Models (8) and (5) respectively. In the mediator model (8), \( \alpha_{IV} \) represents the strength of the IV \( Z \times X^{IV} \) and \( \alpha_u \) represents the strength of endogeneous variable \( U \). The following array (20) shows the true values of parameters in the outcome models. When there are two mediators, we have two IVs in the models. We consider sample sizes of 500, 1000 and 5000. For each setting, 1000 Monte Carlo replications were performed.

\[
\begin{pmatrix}
\text{Setting} & \Theta & \text{value} & \Theta & \text{value} & \Theta & \text{value} & \Theta & \text{value} \\
(A) & \beta_0 & 1 & \beta_z & 0.5 & \beta_m & 0.5 & \beta_x & 0.5 & \beta_u & 1 \\
(B) & \gamma_0 & 0.6 & \gamma_z & 0.25 & \gamma_m & 0.25 & \gamma_x & 0.25 & \gamma_u & 0.5 \\
& \lambda_0 & -1 & \lambda_z & 0.25 & \lambda_m & 0.25 & \lambda_x & 0.25 & \lambda_u & 0.5 \\
(C) & \beta_0 & 1 & \beta_z & 0.5 & \beta_{m1} & 1 & \beta_{x1} & 0.5 & \beta_u & 1 \\
& & & \beta_{m2} & 0.5 & \beta_{x2} & 0.5 \\
(D) & \gamma_0 & 1 & \gamma_z & 0.25 & \gamma_{m1} & 0.75 & \gamma_{x1} & 0.5 & \gamma_u & 0.5 \\
& & & \gamma_{m2} & 0.25 & \gamma_{x2} & 0.5 & \gamma_u & 0.5 \\
& \lambda_0 & -0.5 & \lambda_z & 0.25 & \lambda_{m1} & 0.25 & \lambda_{x1} & 0.5 & \lambda_u & 0.5 \\
& & & \lambda_{m2} & 0.25 & \lambda_{x2} & 0.5 \\
\end{pmatrix}
\] (20)

Single Binary Mediator A binary mediator is generated as

\[
m|x, z, u \sim Ber\left(\frac{\exp(-0.5 + 0.5z + 0.5x + \alpha_{IV} z \times x^{IV} + 0.5u)}{1 + \exp(-0.5 + 0.5z + 0.5x + \alpha_{IV} z \times x^{IV} + 0.5u)}\right),
\] (21)
where $\alpha_{IV} = 1$ for Strong IV (S) and $\alpha_{IV} = 0.5$ for Weak IV (W). Unless otherwise noted, $\alpha_{IV}$ is set as 1. The outcome variable is generated with setting (A) in matrix (20) for Poisson and negative binomial and with setting (B) in matrix (20) for Neyman Type A outcome. There were 21.6% zeros in the Poisson outcome, 25.6% zeros in the generated negative binomial outcome and 52.7% zeros in the Neyman Type A distribution outcome. We considered a standard normal baseline covariate $X$ and the corresponding IV $Z \times X^{IV}$. Then we had eight estimating equations in (17) for four parameters. Two computational methods are proposed to obtain estimates for the parameters:

(1) EE-EL1: The first four estimating functions $g_1, \ldots, g_4$ in (17) are incorporated into the empirical likelihood for estimates and the next four estimating equations $g_5, \ldots, g_8$ are used to evaluate the goodness of fit of the estimates.

(2) EE-EL2: All eight estimating functions $g_1, \ldots, g_8$ in (17) are incorporated into the empirical likelihood for estimates.

The comparisons in Web Table 6 show that both EE-EL1 and EE-EL2 work well. EE-EL1 is fast in terms of computation while EE-EL2 has better performance in terms of median absolute deviation (MAD). Simulation studies reported in this paper were performed with computationally efficient EE-EL1 while real data analyses were conducted with more stable EE-EL2. Table 1 shows the median and MAD of the estimates from the Estimating Equations and Empirical Likelihood (EE-EL), Two Stage Residual Inclusion (2SRI), and ordinary regression (Reg) for direct and indirect effect parameters ($\beta_z$ and $\beta_m$). The ordinary regression fits a Poisson, negative binomial, and zero-inflated Poisson model respectively for the outcome on treatment, mediator and covariates. The 2SRI fits a logistic regression first stage model for the binary mediator with the IV $Z \times X^{IV}$, and then fits a Poisson, negative binomial, and zero-inflated Poisson model for the outcome. The EE-EL does not assume any parametric assumption on the outcome distribution. As shown in Table 1, the ordinary
regression estimates (Reg) are generally biased while both 2SRI and EE-EL estimates have reduced bias with the use of the IV. 2SRI estimate has small bias on the direct effect parameter $\beta_z$ for the Poisson and negative binomial models, but can have a bias of greater than 15% for the Neyman Type A outcome with a big portion of zeroes. For the controlled indirect effect parameter $\beta_m$, 2SRI can have a bias of greater than 25% for the Poisson and negative binomial models and a bias of greater than 100% for the Neyman Type A model. The EE-EL estimator performed best for all the settings and the small bias diminished with increased sample size and stronger IV.

[Table 1 about here.]

Single Continuous Mediator As we discussed in Section 2.5, the natural indirect effect ratio (9) can be estimated for a continuous mediator. The mediator is generated as $m = -0.5 + 0.5z + 0.5x + \alpha_3 z \times x^{IV} + 0.5u + v$, where $v$ follows a standard normal distribution and the outcome variable is generated with setting $(A)$ in matrix (20) for Poisson and negative binomial and with setting $(A)$ in matrix (20) for Neyman Type A outcome. Table 2 shows the median estimates and MAD for natural direct and indirect effect ratios with a continuous mediator. The ordinary regression estimates are biased while both the 2SRI and EL estimates are approximately unbiased. The results for controlled direct and indirect effect ratios with a continuous mediator (see supplement Web Table 2) are similar to the results with a binary mediator in Table 1.

[Table 2 about here.]

Multiple Instrumental Variables We also considered settings with more than one instrumental variable. For example,

\[
m|x, z, u \sim Ber \left( \frac{\exp(-0.5 + 0.5z + 0.5x_1 + 0.5x_2 + z \times x_1^{IV} + z \times x_2^{IV} + 0.5u)}{1 + \exp(-0.5 + 0.5z + 0.5x_1 + 0.5x_2 + z \times x_1^{IV} + z \times x_2^{IV} + 0.5u)} \right),
\]

(22)
The results (see supplement Web Table 1) are similar as Table 1. The EE-EL estimates are consistent while the 2SRI estimates have large bias in some settings.

Multiple Binary Mediators

Considering settings with two independent mediators and two IVs $Z \times X_1^{IV}$ and $Z \times X_2^{IV}$, mediators $m_1$ and $m_2$ are generated independently as

$$m_1|x, z, u \sim \text{Ber} \left( \frac{\exp(-0.5 + 0.5z + 0.5x_1 + 0.5x_2 + z \times x_1^{IV} + z \times x_2^{IV} + 0.5u)}{1 + \exp(-0.5 + 0.5z + 0.5x_1 + 0.5x_2 + z \times x_1^{IV} + z \times x_2^{IV} + 0.5u)} \right),$$

(23)

$$m_2|x, z, u \sim \text{Ber} \left( \frac{\exp(1 + z - 0.5x_1 + x_2 + z \times x_1^{IV} + z \times x_2^{IV} + 0.5u)}{1 + \exp(1 + z - 0.5x_1 + x_2 + z \times x_1^{IV} + z \times x_2^{IV} + 0.5u)} \right).$$

(24)

The outcome variable is generated as

$$\mathbb{E}(Y(z, m_1, m_2)|z, m_1, m_2, x, u) = \exp(\beta_0 + \beta_z z + \beta_{m_1} m_1 + \beta_{m_2} m_2 + \beta_x x + \beta_u u).$$

(25)

with setting (C) in matrix (20) for Poisson and negative binomial and with setting (D) in matrix (20) for Neyman Type A outcome. There were 26.3% zeros in the Poisson outcome, 29.8% zeros in the generated Negative Binomial outcome and 38.6% zeros in the Neyman Type A distribution outcome. Similar to Table 1, Table 3 shows that the ordinary regression estimate is heavily biased. 2SRI can reduce the bias in some cases but has large bias in other cases when the sample size is small and/or percentage of zeros is relatively large. The EE-EL performed best in all the cases and produced consistent estimates with increased sample size.

[Table 3 about here.]

Sensitivity Analysis Method. Sensitivity analysis was examined as discussed in Section 3.3. The Web Supplement Table 3 shows that after we adjust the outcome and use the adjusted outcome in the methods, the results are similar to the results in Table 1 and the EE-EL estimates are approximately unbiased for the direct and indirect effect parameters when $\eta$ is known.

High Proportion of Zeros and Mis-specification in Outcome Distribution. Additional simu-
lation studies were conducted to examine the performance of methods with increased percentage of zeros (50% for Poisson and 55% for negative binomial), and when the outcome distribution is misspecified in 2SRI. The results with increased percentage of zeros (Web Supplement Table 4) are similar to the results in Table 1. When the outcome distribution is misspecified, the 2SRI produced biased estimates while the EE-EL does not rely on a parametric outcome distribution (Web Supplement Table 5).

In summary, the ordinary regression analysis produces biased estimates for the direct and indirect effect parameters when there is unmeasured confounding. The 2SRI and EE-EL reduce bias with the use of instrumental variables. However, 2SRI can have big bias when sample size is small or the percentage of zeros is big with a binary mediator or the outcome distribution is misspecified. The EE-EL generally performs well under different settings and is robust to the misspecification of outcome distribution. The sensitivity analysis we proposed performs well too.

5. Real Data Analysis

We analyzed two real studies with methods developed in this paper, one study with a binary mediator and another study with a continuous mediator.

Dental Study DDHP MI-DVD. In the Detroit Dental Health Project’s Motivational Interviewing DVD (DDHP MI-DVD) trial (Ismail et al. 2011), 790 families (0-5 years old children and their caregivers) were randomly assigned to one of two education groups (DVD only or MI+DVD). Both groups of families received a copy of a special 15-minute DVD for dental education. Additionally the families in the intervention group (MI+DVD) met a motivational interviewing (MI) interviewer, developed their own preventive goals, and received booster calls within 6 months of the intervention. The primary analyses of the study showed that caregivers in the MI+DVD group were more likely to make sure their child brushed at bedtime at 6 months and 2 years, but the intervention did not have a significant effect on
children’s dental outcomes at 2 years. In this study, we are interested in whether there was a direct effect of the intervention on children’s dental outcomes that cancelled out a mediation effect in the opposite direction so that no significant total effect of the intervention was found. We assessed if the intervention had an effect on the outcome (the number of new untreated lesions at 2 years) through a binary mediator (whether or not caregivers made sure their child brushed at bedtime), where more than 60% of children had zero new untreated lesions (Figure 1(a)). The instrumental variables used were the interactions between intervention and three baseline covariates: number of times child brushed, whether or not caregivers made sure their child brushed at baseline, and whether or not caregivers provided the child healthy meals at baseline. A logistic regression was fitted for the binary mediator, showing significant effect of both the treatment and IV.

Table 4 summarizes the EE-EL estimates of the direct and indirect effect ratios and the bootstrap confidence intervals. Note that a ratio of 1 indicates no effect. The result shows that the intervention did not have much direct effect on the number of new untreated lesions (controlled direct effect ratio 1.081), and parent behavior in making sure their child brushed at bedtime tended to decrease the number of new untreated lesions (controlled indirect effect ratio 0.595) but the effect was not statistically significant with a 90% CI (0.0524, 9.735).

In this paper we are particularly interested in whether or not the MI intervention affected the children’s oral health through its effect on the mediator whether or not caregivers made sure their child brushed at bedtime. Besides the pathway through the mediator, the effects of MI intervention on the children’s oral health through other pathways are included in the direct effect. Although we are not able to formally test the ER assumption of the IVs, we can examine the effects of the IVs through other pathways and therefore partially assess the plausibility of the ER assumption of the IVs. Specifically we examined whether the IVs were associated with other intermediate variables such as whether the caregiver provided the
child with nonsugared snacks, whether the caregiver gave the child healthy meals, whether
the caregiver checked the child for early non-cavitated demineralized enamel, whether the
caregiver made sure the child saw a dentist every 6 months given the intervention and baseline
covariates. None the IVs were significantly associated with other intermediate variables,
indicating no evidence of the violation of the ER assumption of instrumental variables.

The sensitivity analysis (Figure 2) shows that with increased amount of violation of the
ER assumption, that is, with increased direct effect of the IV \((Z \times X^{IV})\) on children’s oral
health \((\eta\) increases from 0 to 0.15), the direct and indirect effects of the MI intervention on
children’s oral health decreased. Specifically, the direct effect changed from a small increase
to a small reduction in the number of new untreated lesions (the direct effect ratio drops
below 1 from 1.081 but neither was significant), and indirect effect through parents making
sure their child brushed at bedtime showed additional reduction in the number of untreated
lesions (the controlled indirect effect ratio decreases from 0.595 to 0.147).

Flood data analysis. In 1998, two-thirds of Bangladesh suffered from massive floods. del
Ninno et al. (2001) conducted a study of the effects of flooding on health outcomes. We will
use our method to see whether a household being severely affected by the flood (treatment)
influenced the number of days a child had diarrhea in the three month period after the flood
(outcome) through its effect on the per capita calorie consumption of the household (mediator).
We assume strong ignobility for the treatment conditioning on baseline covariates: sex, 
age, the size of the household, mother’s education, father’s education, indicator of missing
values for mother’s education and father’s education, mother’s age, father’s age, indicator
of missing values for mother’s age and father’s age. Because the mediator is continuous, we
are able to evaluate the natural effects as discussed in Section 2.5. The instrumental variable
is interaction of the flooding and a baseline covariate, whether the household has a low or
large amount of farmland available. A primary analysis has shown that the treatment and
the instrumental variable have significant effects on the mediator. The outcome histogram Figure 1(b) shows that more than 70% children had zero days of diarrhea.

Table 4 illustrates the EE-EL estimates of the controlled (natural) direct and indirect effect and the bootstrap confidence intervals. The flooding tended to increase but did not have a significant direct effect on the number of days of diarrhea (direct effect ratio 1.229, 90% CI: 0.300, 2.409). A larger per capita calorie consumption of a household led to a significant decrease in the number of days a child had diarrhea over the three month period after the flood (controlled indirect effect ratio 0.040, 90% CI: 2.524 × 10⁻⁵, 0.822); and decreased per capita calorie consumption of a household due to flooding led to a significant increase in the number of days a child had diarrhea after the flood (natural indirect effect ratio 1.685, 90% CI: 1.010, 6.239).

In addition to the flood effect on the children’s number of days of diarrhea through its effect on the per capita calorie consumption of the household, the flood could also affect children’s number of days of diarrhea through other pathways, which were not the interest in this paper and therefore included in direct effects. One possible pathway is mother’s health. del Ninno et al. (2001) used the variable, whether a mother had chronic energy deficiency (CED) as a measure of mother’s health. Specifically a mother was classified as being CED if her body mass index was less than 18.5. Non-significant association of the IV with CED (p=0.1501) given flood and baseline covariates indicated no evidence of violation of the ER assumption.

We further conducted a sensitivity analysis to see how the results would change with the violation of the ER assumption. Figure 2 shows that with increased direct effect of $Z \times X^{IV}$ on children’s diarrhea, the direct effect of flood changed from a small increase 1.229 to a small reduction 0.934 in the number of days of diarrhea with neither significant, and the indirect effect through higher per capita calorie consumption showed a greater reduction in the number of days of diarrhea (from 0.040 to 0.013). With increased amount of the violation
of the ER assumption, the increased indirect effect indicates a greater rise in the number of days of diarrhea from decreased per capita calorie consumption of a household due to flooding.

[Figure 1 about here.]

[Figure 2 about here.]

[Table 4 about here.]

6. Conclusion and discussion

In this paper, we consider mediation analysis and define the direct and indirect effect ratios for count or zero inflated count outcome when there is concern about unmeasured confounding between the mediator and outcome. Our method uses the interaction of treatment and baseline covariates as instrumental variables, constructs estimating equations and use the empirical likelihood approach to combine the information in estimating equations. Our method relaxes the assumption of sequential ignorability with reasonable assumptions and does not rely on parametric outcome distribution assumptions. A sensitivity analysis is proposed for the violation of the ER assumption. Simulation studies show that the two stage approach (2SRI) reduces bias with the use of IV compared to ordinary regression, but can produce biased estimates when the mediator is binary. The estimating equations empirical likelihood (EE-EL) method generally provides approximately unbiased estimates for the direct and indirect effects for different types of outcomes and under different settings (binary mediator, continuous mediator, multiple independent mediators, multiple instrumental variables) and is robust to the outcome distribution.

The model (5) considered in this paper can be generalized to include the interaction terms \( z \times x \), \( z \times m \) and \( x \times m \).

\[
f\{E(Y(z,m)|x,u)\} = \beta_0 + \beta_z z + \beta_m m + \beta_x^T x + \beta_u u + \beta_{zx}^T x \times z + \beta_{zm}^T x \times m + \beta_{zm} z \times m, \quad (26)
\]
where $f$ is a link function and $x$ is a random vector of length $l$. In model (26), we have $l + 2$ endogenous variables $m$, $x \times m$ and $z \times m$. Hence, it requires at least $l + 2$ valid instruments $z \times x^{IV}$ such that conditioning on $z$, $x$ and $z \times x$, (1) $z \times x^{IV}$ predict the endogenous variables $m$, $x \times m$ and $z \times m$; (2) $z \times x^{IV}$ is independent of the unmeasured confounding $u$; and (3) $z \times x^{IV}$ does not have a direct effect on the outcome other than through its effects on $m$, $x \times m$ and $z \times m$. However, the effects are not the same as (9)-(11), but should be re-derived from the generalized model (26) and (1)-(4). This is beyond the scope of this paper and we will leave it for a further study.

Supplementary Materials

Web Appendices, Tables, and proofs are available at the Biometrics website on Wiley Online Library.

Acknowledgement

The authors thank Amid Ismail and Sungwoo Lim for providing us the DDHP MI-DVD data. This paper was made possible by Grant Number U54 DE 019285 from the NIDCR, a component of the National Institutes of Health (NIH).

References

Albert, J.M. (2008). Mediation Analysis via potential outcomes models. *Statistics in Medicine*, **27**, 1282-1304.

Angrist, J. and Krueger, A. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics* **106**, 979-1014.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* **91**, 444-455.

Baron, R.M and Kenny, D.A. (1986). The moderator-mediator variable distinction in social
psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*, **51**, 1173-1182.

Cai, B., Small, D., and Ten Have, T. (2011). Two-stage instrumental variable methods for estimating the causal odds ratio: analysis of bias. *Statistics in Medicine*, **30**, 1809-1824.

Cole, D.A., and Maxwell, S.E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology* **112**, 558577.

Daniels, M.J., Roy, J., Kim, C., Hogan, J.W., and Perri, M.G. (2012) Bayesian Inference for the Causal Effect of Mediation. *Biometrics*, **68**, 1028-1036.

Dobbie, M.J. and Welsh, A.H. (2001). Models for zero-inflated count data using the Neyman type A distribution. *Statistical Modeling*, **1**, 65-80.

Dunn, G. and Bentall, R. (2007). Modeling treatment effect heterogeneity in randomised controlled trials of complex interventions (psychological treatments). *Statistics in Medicine* **26**, 4719-4745.

Elliott, M.R., Raghunathan, T.E., and Li, Y. (2010). Bayesian Inference for Causal Mediation Effects Using Principal Stratification with Dichotomous Mediators and Outcomes. *Biostatistics* **11**, 353-372.

Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods* **15**, 309-334.

Ismail, A.I., Ondersma, S., Willem Jedele, J.M., Little, R.J., Lepkowski J.M. (2011). Evaluation of a brief tailored motivational intervention to prevent early childhood caries. *Community Dentistry and Oral Epidemiology* **39**, 433-448.

Jo, B. (2002). Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics* **27**(4), 385-409.

Jo, B., Stuart, E.A., MacKinnon, D.P., and Vinokur, A.D. (2011). The use of propensity
scores in mediation analysis. *Multivariate Behavioral Research* 46, 425-452.

MacKinnon, D.P. (2008). *Introduction to Statistical Mediation Analysis*. New York: Erlbaum.

MacKinnon, D.P. and Luecen, L.J. (2011). Statistical analysis for identifying mediating variables in public health dentistry interventions. *Journal of Public Health Dentistry*, 71 Suppl 1, S37-46.

Mullahy, J. (1997). Instrumental-variable estimation of count data models: Applications to models of cigarette smoking behavior. *Review of Economics and Statistics*, 79, 586–593.

Nagelkerke, N., Fidler, V., Bernsen, R., Borgdorff, M. (2000). Estimating treatment effects in randomized clinical trials in the presence of non-compliance. *Statistics in Medicine*, 19, 1849-1864.

Neyman, J. (1923). On the Application of Probability Theory to Agricultural Experiments. Trans. D. Dabrowska. *Statistical Science*, 1990, 5, 463-480.

del Ninno, C., Dorosh, P.A., Smith, L.C. and Roy, D.K (2001). The 1998 Floods in Bangladesh: Disaster impacts, household coping strategies and response. Research Report 122. Washington D.C.: *International Food Policy Research Institute (IFPRI)*.

Owen, A.B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75, 237-249.

Owen, A.B. (1990). Empirical likelihood confidence regions. *Annals of Statistics*, 18, 90-120.

Pearl, J. (2001). Direct and indirect effects. In J. Breese & D. Koller (Eds.), Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence (pp. 411420). San Francisco, CA: Morgan Kaufmann.

Qin, J. and Lawless, J. (1994) Empirical likelihood and general estimating equations. *The Annals of Statistics*, 22, 300–325.

Robins, J.M. and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 3, 143-155.
Rubin, D.B. (1974). Estimating Causal Effects of Treatments in Randomized and Non-
normalized Studies. *Journal of Educational Psychology* **66**, 688-701.

Small, D. (2012). Mediation analysis without sequential ignorability: using baseline covariates 
interacted with random assignment as instrumental variables. *Journal of Statistical 
Research*, **46**, 91–103.

Sobel, M.E. (2008). Identification of causal parameters in randomized studies with mediating 
variables. *Journal of Educational and Behavioral Statistics* **33**, 230251.

Steyer, R., Mayer, A., and Fiege, C. (2014). Causal inference on total, direct, and indirect 
effects. In A.C. Michalos (Ed.), Encyclopedia of quality of life and well-being research 
(pp. 606631). Dordrecht, The Netherlands: Springer.

TenHave, T.R., Joffe, M., Lynch, K., Maisto, S., Brown, G. and Beck, A. (2007). Causal 
mediation analyses with rank preserving models. *Biometrics* **63**, 926-934.

Terza, J., Basu, A., Rathouz, P. (2008). Two-stage residual inclusion estimation: addressing 
edgenessity in health econometric modeling. *Health Economics*, **27**, 527-543.

van der Laan, M. and Petersen, M. (2008). Direct effect models. *International Journal of 
Biostatistics* **4**, Article 23.

VanderWeele, T.J. and Vansteelandt, S. (2009). Conceptual issues concerning mediation, 
interventions and composition. *Statistics in Its Interface* **2**, 457-468.

VanderWeele, T.J. and Vansteelandt, S. (2010). Odds ratios for mediation analysis for a 
dichotomous outcome. *American Journal of Epidemiology* **172**, 1339-1348.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data, MIT press.
Figure 1: (a) Dental Study; (b) Flood data;
Figure 2: Sensitivity analysis of the real data. On the left plot, Direct represents the controlled and natural direct effect and Indirect only represents the controlled indirect effect; on the right plot, Direct represents the controlled and natural direct effect, ConInd represents the controlled indirect effect and NatInd represents the natural indirect effect.
Table 1: Estimates for the direct effect parameter ($\beta_z = 0.5$) and the indirect effect parameter ($\beta_m = 0.5$) with one instrumental variable and one binary mediator.

| Outcome | IV | n   | Direct | Indirect |
|---------|----|-----|--------|----------|
|         |    |     | EE-EL Med. (MAD) | 2SRI Med. (MAD) | Reg EE-EL Med. (MAD) | 2SRI Med. (MAD) | Reg EE-EL Med. (MAD) | 2SRI Med. (MAD) |
| Poi S   | 500|     | 0.496 (0.126) | 0.505 (0.150) | 0.416 (0.117) | 0.522 (0.691) | 0.559 (0.722) | 0.955 (0.102) |
| Poi S   | 1000|    | 0.497 (0.094) | 0.497 (0.111) | 0.419 (0.080) | 0.522 (0.489) | 0.561 (0.495) | 0.948 (0.082) |
| Poi S   | 5000|    | 0.497 (0.039) | 0.500 (0.047) | 0.415 (0.034) | 0.510 (0.225) | 0.528 (0.224) | 0.948 (0.036) |
| Poi W   | 500|     | 0.489 (0.159) | 0.500 (0.198) | 0.426 (0.122) | 0.530 (1.117) | 0.524 (1.266) | 0.966 (0.104) |
| Poi W   | 1000|    | 0.486 (0.120) | 0.483 (0.133) | 0.430 (0.078) | 0.575 (0.793) | 0.643 (0.852) | 0.962 (0.082) |
| Poi W   | 5000|    | 0.500 (0.054) | 0.496 (0.058) | 0.432 (0.035) | 0.499 (0.389) | 0.576 (0.388) | 0.959 (0.036) |
| NB S    | 500|     | 0.494 (0.164) | 0.503 (0.155) | 0.443 (0.127) | 0.448 (0.820) | 0.494 (0.736) | 0.944 (0.131) |
| NB S    | 1000|    | 0.499 (0.114) | 0.503 (0.116) | 0.447 (0.086) | 0.555 (0.615) | 0.528 (0.515) | 0.942 (0.089) |
| NB S    | 5000|    | 0.498 (0.050) | 0.499 (0.051) | 0.442 (0.037) | 0.486 (0.262) | 0.502 (0.225) | 0.945 (0.041) |
| NB W    | 500|     | 0.493 (0.177) | 0.500 (0.194) | 0.447 (0.123) | 0.543 (1.261) | 0.526 (1.295) | 0.952 (0.136) |
| NB W    | 1000|    | 0.489 (0.134) | 0.493 (0.130) | 0.439 (0.084) | 0.525 (0.983) | 0.515 (0.849) | 0.955 (0.094) |
| NB W    | 5000|     | 0.499 (0.068) | 0.499 (0.060) | 0.449 (0.039) | 0.524 (0.468) | 0.547 (0.395) | 0.951 (0.038) |
| NTA S   | 500|     | 0.465 (0.316) | 0.413 (0.478) | 0.386 (0.228) | 0.589 (1.759) | 0.841 (1.723) | 0.988 (0.206) |
| NTA S   | 1000|    | 0.504 (0.209) | 0.451 (0.395) | 0.361 (0.174) | 0.448 (1.256) | 0.692 (1.365) | 0.983 (0.157) |
| NTA S   | 5000|    | 0.499 (0.119) | 0.461 (0.242) | 0.369 (0.088) | 0.500 (0.639) | 0.674 (0.783) | 0.981 (0.073) |
| NTA W   | 500|     | 0.476 (0.293) | 0.407 (0.581) | 0.414 (0.239) | 0.515 (2.261) | 1.054 (2.896) | 1.012 (0.211) |
| NTA W   | 1000|    | 0.496 (0.248) | 0.412 (0.484) | 0.403 (0.176) | 0.436 (1.848) | 0.903 (2.458) | 0.988 (0.154) |
| NTA W   | 5000|    | 0.499 (0.160) | 0.448 (0.302) | 0.400 (0.084) | 0.500 (1.053) | 0.766 (1.317) | 0.982 (0.072) |

n: sample size; Med: median of 1000 Monte Carlo estimates; MAD: median absolute deviance; EE-EL: estimating equations and empirical likelihood; 2SRI: two stage residual inclusion; Reg: ordinary regression; Poi: Poisson; NB: negative binomial; NTA: Neyman Type A distribution outcome; S: stronger IV (setting 1); W: relatively weaker IV (setting 2).
Table 2: Estimates for the natural indirect rate ratio (9) with one instrumental variable and one continuous mediator.

| Outcome | IV | n  | True | (9) with \( x = 1 \) | (9) with \( x = -1 \) |
|---------|----|----|------|-----------------|-----------------|
|         |    |    | (EE-EL Med. (MAD), 2SRI Med. (MAD)) | (EE-EL Med. (MAD), 2SRI Med. (MAD)) |
| Poi     | S  | 500| 2.117 | 2.114 (0.487), 2.053 (0.619) | 2.088 (0.320), 3.163 (0.491) |
|         | S  | 1000| 2.117 | 2.088 (0.320), 2.100 (0.455) | 2.112 (0.152), 3.169 (0.388) |
|         | S  | 5000| 2.117 | 2.112 (0.152), 2.098 (0.219) | 2.117 (0.062), 2.030 (0.184) |
| Poi     | W  | 500| 1.649 | 1.644 (0.442), 1.635 (0.558) | 1.635 (0.305), 2.247 (0.214) |
|         | W  | 1000| 1.649 | 1.635 (0.305), 1.644 (0.417) | 1.635 (0.062), 2.223 (0.214) |
| Poi     | W  | 5000| 1.649 | 1.642 (0.150), 1.639 (0.193) | 1.642 (0.062), 2.227 (0.101) |
| NB      | S  | 500| 2.117 | 2.128 (0.534), 2.114 (0.513) | 2.114 (0.379), 3.428 (0.493) |
|         | S  | 1000| 2.117 | 2.114 (0.379), 2.101 (0.357) | 2.115 (0.166), 3.422 (0.154) |
|         | S  | 5000| 2.117 | 2.115 (0.166), 2.096 (0.163) | 2.115 (0.062), 3.421 (0.154) |
| NB      | W  | 500| 1.649 | 1.661 (0.564), 1.650 (0.537) | 1.661 (0.386), 2.376 (0.323) |
|         | W  | 1000| 1.649 | 1.681 (0.386), 1.670 (0.364) | 1.681 (0.051), 2.385 (0.229) |
|         | W  | 5000| 1.649 | 1.646 (0.156), 1.638 (0.158) | 1.646 (0.051), 2.365 (0.103) |
| NTA     | S  | 500| 2.117 | 2.057 (1.256), 2.079 (1.064) | 2.057 (1.006), 3.225 (0.877) |
|         | S  | 1000| 2.117 | 2.141 (1.006), 2.075 (0.951) | 2.136 (0.948), 3.247 (0.591) |
|         | S  | 5000| 2.117 | 2.136 (0.948), 2.082 (0.528) | 2.136 (0.305), 3.252 (0.354) |
| NTA     | W  | 500| 1.649 | 1.585 (1.092), 1.610 (1.050) | 1.585 (1.109), 2.237 (0.378) |
|         | W  | 1000| 1.649 | 1.654 (0.889), 1.609 (0.799) | 1.654 (0.808), 2.225 (0.280) |
|         | W  | 5000| 1.649 | 1.627 (0.455), 1.631 (0.522) | 1.627 (0.084), 2.233 (0.155) |

n: sample size; Med: median of 1000 Monte Carlo estimates; MAD: median absolute deviance; EE-EL: estimating equations and empirical likelihood; 2SRI: two stage residual inclusion; Reg: ordinary regression; Poi: Poisson; NB: negative binomial; NTA: Neyman Type A distribution outcome; S: stronger IV (setting 1); W: relatively weaker IV (setting 2).
Table 3: Estimates for the direct effect parameter ($\beta_z = 0.5$) and the indirect effect parameter ($\beta_{m_1} = 1$ and $\beta_{m_2} = 0.5$) with two IVs and two binary mediators.

| Dis. | Direct | Indirect1 | Indirect2 |
|------|--------|-----------|-----------|
|      | EE-EL  | Med. (MAD)| EE-EL  | Med. (MAD)| EE-EL  | Med. (MAD)| EE-EL  | Med. (MAD)| EE-EL  | Med. (MAD)|
|      | Med. (MAD)|         | Med. (MAD)|         | Med. (MAD)|         | Med. (MAD)|         | Med. (MAD)|         |
| Poi  |        |           |           |           |        |          |        |          |        |          |
| (500)| 0.448  | 0.464     | 0.319    | 0.932    | 1.156  | 1.425    | 0.661  | 0.644    | 0.911  |
| (1000)| 0.482  | 0.489     | 0.319    | 0.910    | 1.094  | 1.417    | 0.527  | 0.623    | 0.897  |
| (5000)| 0.490  | 0.502     | 0.321    | 0.990    | 1.080  | 1.416    | 0.521  | 0.511    | 0.902  |
| NB   |        |           |           |           |        |          |        |          |        |          |
| (500)| 0.442  | 0.481     | 0.342    | 0.979    | 1.184  | 1.432    | 0.592  | 0.514    | 0.950  |
| (1000)| 0.470  | 0.504     | 0.340    | 0.910    | 1.076  | 1.431    | 0.579  | 0.441    | 0.939  |
| (5000)| 0.492  | 0.508     | 0.340    | 0.972    | 1.088  | 1.427    | 0.528  | 0.437    | 0.944  |
| NTA  |        |           |           |           |        |          |        |          |        |          |
| (500)| 0.423  | 0.377     | 0.315    | 0.925    | 1.247  | 1.445    | 0.653  | 0.838    | 0.949  |
| (1000)| 0.473  | 0.440     | 0.316    | 0.915    | 1.185  | 1.440    | 0.546  | 0.616    | 0.939  |
| (5000)| 0.490  | 0.480     | 0.311    | 0.970    | 1.099  | 1.436    | 0.492  | 0.552    | 0.917  |

n: sample size; Med: median of 1000 Monte Carlo estimates; MAD: median absolute deviance; EE-EL: estimating equations and empirical likelihood; 2SRI: two stage residual inclusion; Reg: ordinary regression; Poi: Poisson; NB: negative binomial; NTA: Neyman Type A distribution outcome; S: stronger IV (setting 1); W: relatively weaker IV (setting 2).
Table 4: EL estimate and Bootstrap confidence interval for direct effect rate ratio and indirect effect rate ratio.

|                      | Dental study data |                             |                             |
|----------------------|-------------------|-----------------------------|-----------------------------|
|                      | EE-EL             | 90% confidence interval     |                             |
| Controlled/Natural Direct | 1.081             | (0.760, 1.442)              |                             |
| Controlled Indirect  | 0.595             | (0.052, 9.735)              |                             |

|                      | Flood data        |                             |                             |
|----------------------|-------------------|-----------------------------|-----------------------------|
|                      | EE-EL             | 90% confidence interval     |                             |
| Controlled/Natural Direct | 1.229             | (0.299, 2.409)              |                             |
| Controlled Indirect  | 0.040             | (2.524 × 10^{-5}, 0.822)   |                             |
| Natural Indirect     | 1.685             | (1.010, 6.239)              |                             |
Web-based Supplementary Materials for Mediation Analysis for Count and
Zero-Inflated Count Data Without Sequential Ignorability

Zijian Guo*, Dylan S. Small

Department of Statistics, University of Pennsylvania Wharton School, Philadelphia, PA 19104, USA

Stuart A. Gansky and Jing Cheng**
Division of Oral Epidemiology & Dental Public Health, University of California, San Francisco, CA 94143, USA

*email: zijguo@wharton.upenn.edu

**email: jing.cheng@ucsf.edu

SUMMARY: This web-based supplementary materials contain three sections. Web Appendix A presents the technical
proofs; Web Appendix B presents the estimating equations for two conditional independent mediators; Web Appendix
C presents more simulation studies.

This paper has been submitted for consideration for publication in *Biometrics*
1. Web Appendix A

1.1 Proof of natural effect ratio

We first show the proof for the natural effect rate ratio. When the mediator is continuous and \( M(z^*) = \alpha_0 + \alpha_z z^* + \alpha_x x + \alpha_{IV} x z^* + \alpha_u u + v \) with \( v \) independent of \( z^* \) and \( x \), we have the conditional expectation of the potential outcome \( Y(z, M(z^*)) \),

\[
\mathbb{E}_Y(Y(z, M(z^*)) | x, u) = \exp(\beta_0 + \beta_z z + \beta_m (\alpha_0 + \alpha_z z^* + \alpha_x x + \alpha_{IV} x z^* + \alpha_u u + v) + \beta_x x + \beta_u u).
\]

By integrating with respect to \( v \), we have

\[
\mathbb{E}_{M(z^*) | x, u} \mathbb{E}_Y(Y(z, M(z^*)) | x, u) = \exp(\log \mathbb{E}(\exp(v)) + \beta_0 + \beta_m \alpha_0 + \beta_z z + \beta_m \alpha_z z^* + \beta_m \alpha_{IV} x z^* + (\beta_x + \beta_m \alpha_x) x + (\beta_u + \beta_m \alpha_u) u).
\]

The natural direct effect rate ratio can be expressed as

\[
\frac{\mathbb{E}(Y(z, M^z) | x, u)}{\mathbb{E}(Y(z^*, M^z) | x, u)} = \exp(\beta_z (z - z^*)), \tag{2}
\]

and the natural indirect effect rate ratio is

\[
\frac{\mathbb{E}(Y(z, M^z) | x, u)}{\mathbb{E}(Y(z, M^z^*) | x, u)} = \exp(\beta_m \alpha_z (z - z^*) + \beta_m \alpha_{IV} x (z - z^*)). \tag{3}
\]

In the following, we derive the natural direct ratio for binary mediator.

\[
\frac{\mathbb{E}(Y(z, M^z) | x, u)}{\mathbb{E}(Y(z^*, M^z) | x, u)} = \frac{P(M(z^*) = 1 | x, u) \mathbb{E}(y(z, M(z^*) = 1) | x, u) + P(M(z^*) = 0 | x, u) \mathbb{E}(y(z, M(z^*) = 0) | x, u)}{P(M(z^*) = 1 | x, u) \mathbb{E}(y(z^*, M(z^*) = 1) | x, u) + P(M(z^*) = 0 | x, u) \mathbb{E}(y(z^*, M(z^*) = 0) | x, u)}
\]

\[
= \frac{\exp(\beta_z z) (P(M(z^*) = 1 | x, u) \exp(\beta_m) + P(M(z^*) = 0 | x, u)) \exp(\beta_0 + \beta_z x + \beta_u u)}{\exp(\beta_z z^*) (P(M(z^*) = 1 | x, u) \exp(\beta_m) + P(M(z^*) = 0 | x, u)) \exp(\beta_0 + \beta_z x + \beta_u u)}
\]

\[
= \exp(\beta_z (z - z^*)).
\]
The proof of the natural indirect ratio for binary mediator is as follows.

\[
\frac{E(Y(z, M^z) | x, u)}{E(Y(z, M^z^*) | x, u)} = \frac{P(M(z) = 1| x, u) E(y(z^*, M(z) = 1) | x, u) + P(M(z) = 0| x, u) E(y(z^*, M(z) = 0) | x, u)}{P(M(z^*) = 1| x, u) E(y(z^*, M^z = 1) | x, u)} + \frac{P(M(z) = 0| x, u) E(y(z^*, M(z^*) = 0) | x, u)}{P(M(z^*) = 0| x, u)}
\]

(5)

1.2 Proof of the estimating equations

In the following, we derive the estimation equations that we propose in the main paper.

\[
\frac{E\left(\frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m + \beta_x x) - 1}\right)}{E\left(\frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m + \beta_x x) - 1| z, x, m, u}\right)} = E\left(\frac{E(y| z, x, m, u)}{\exp(\beta_0 + \beta_z z + \beta_m m + \beta_x x) - 1}\right) - 1 = 0.
\]

(6)

\[
\frac{E\left(\frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m + \beta_x x) - 1} \times xz\right)}{E\left(\frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m + \beta_x x) - 1| z, x, m, u}\right) \times xz} = E\left(\frac{E(y| z, x, m, u)}{\exp(\beta_0 + \beta_z z + \beta_m m + \beta_x x) - 1| z, x, m, u}\right) \times xz - 1 = 0.
\]

(7)

\[
= E((\exp(\beta_u u) - 1) xz) = 0.
\]

where the last equation follows from the randomness of treatment and the assumption that 
\(P(u|x)\) has the same distribution across different \(x\). We also have the following estimating
equations,
\[
\mathbb{E}\left( \frac{y}{\exp(\beta_m m)} - \exp(\beta_0 + \beta_z z + \beta_x x) \right) \\
= \mathbb{E}\left( \mathbb{E}\left( \frac{y}{\exp(\beta_m m)} - \exp(\beta_0 + \beta_z z + \beta_x x) | z, x, m, u \right) \right) \\
= \mathbb{E}\left( \frac{\mathbb{E}(y | z, x, m, u)}{\exp(\beta_m m)} - \exp(\beta_0 + \beta_z z + \beta_x x) \right) \\
= \mathbb{E}\left((\exp(\beta_u u) - 1) \exp(\beta_0 + \beta_z z + \beta_x x) \right) = 0;
\]
where the last equation follows from the randomness of treatment and the assumption that \( P(u|x) \) has the same distribution across different \( x \). The proofs of other estimating equations are similar.

1.3 **The consistency of the 2SRI estimator when the first stage is linear**

Consider the following outcome model
\[
\mathbb{E}(Y(Z, M)|X,U) = \exp(\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u U),
\]
and the mediator model,
\[
M = \alpha_0 + \alpha_z Z + \alpha_x X + \alpha_{IV} Z \times X + (\alpha_u U + V),
\]
where \( \alpha_u U + V \) is the error and \( U \) is the unmeasured confounder with \( (\alpha_u U + V, U) \) following bivariate normal distribution and is independent of \((X, Z, Z \times X)\). Let \( \alpha^*_j \) denote the probability limit of the logistic regression estimator \( m \sim z + x + z \times x \) and \( \alpha^*_j = \alpha_j \) for \( j = 0, z, x, IV \). In the first stage, the residual \( R \) is defined as
\[
R = M - (\alpha_0 + \alpha_z Z + \alpha_x X + \alpha_{IV} Z \times X) = \alpha_u U + V;
\]
and decompose \( U \) into two parts
\[
U = \tau R + \delta,
\]
where \( \delta \) is the population residual of the OLS \( U \sim R \) and \( \delta \) is independent of \( R \). Since
\[
\mathbb{E}(Y(Z, M)|X, R, \delta) = \exp(\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R + \beta_u \delta),
\]
we have
\[ E(Y(Z, M)|X, R) = \int \exp(\beta_0 + \beta_zZ + \beta_mM + \beta_xX + \beta_u\tau R + \beta_u\delta) \, dP(\delta|Z, M, X, R), \]
\[ = \exp(\beta_0 + \beta_zZ + \beta_mM + \beta_xX + \beta_u\tau R) \int \exp(\beta_u\delta) \, dP(\delta|Z, M, X, R). \]  
(11)

Since \((\alpha_u U + V, U)\) is independent of \((X, Z, Z \times X)\), \(\delta\) is independent of \((X, Z, Z \times X)\). Since \(\delta\) is independent of \(R\) and \(M\) is linear combination of \((X, Z, Z \times X)\) and \(R\), \(\delta\) is independent of \(R\) and \(M\) and hence \(\int \exp(\beta_u\delta) \, dP(\delta|Z, M, X, R)\) is a constant and the 2SRI estimator is consistent. We just show that if we know the error \(R\). Since \(R\) is unknown, we need to estimate \(R\) by \(\hat{R}\) which is the residual of the first stage regression \(m \sim z + x + z \times x\). Under the identification assumption and the regularity conditions, the 2SRI estimator is consistent even when \(R\) is replaced by \(\hat{R}\) in the second stage. More detailed and rigorous discussion is referred to section 12.4.1 in Wooldridge(2010).

1.4 The consistency of the 2SRI estimator when the second stage is linear

Consider the first stage model
\[ M|X, Z, U \sim Ber\left(\frac{\exp(\alpha_0 + \alpha_zZ + \alpha_xX + \alpha_IVZ \times X + \alpha_uU)}{1 + \exp(\alpha_0 + \alpha_zZ + \alpha_xX + \alpha_IVZ \times X + \alpha_uU)}\right), \]  
(12)

and the following second stage model
\[ Y(Z, M) = \beta_0 + \beta_zZ + \beta_mM + \beta_xX + (\alpha_u U + V), \]  
(13)

where \(\alpha_u U + V\) is the error and \(U\) is the unmeasured confounder with \((\alpha_u U + V, U)\) following bivariate normal distribution and is independent of \((X, Z, Z \times X)\). When the second stage model is linear, then the 2SPS and 2SRI estimators are same. The validity of the instrumental variables will guarantee that the 2SPS and 2SRI estimators are consistent. However, to contrast the argument with the case where the second stage model is Poisson, Negative Binomial and Neyman Type A distribution, we show in the following that the 2SRI estimator is consistent if \(R\) is known. When we estimate \(R\) by a consistent estimator \(\hat{R}\), the 2SRI estimator is still consistent but the proof is omitted here. Let \(\alpha^*\) denote the probability limit
of the logistic regression estimator \( m \sim z + x + z \times x \). In the first stage, the residual \( R \) is defined as
\[
R = M - \frac{\exp (\alpha_0 + \alpha_z Z + \alpha_x^* X + \alpha_{IV} Z \times X)}{1 + \exp (\alpha_0 + \alpha_z Z + \alpha_x^* X + \alpha_{IV} Z \times X)},
\]
and decompose \( U \) into two parts
\[
U = \tau R + \delta
\]
where \( \delta \) is the population residue of the OLS \( U \sim R \) and \( \mathbb{E}(\delta|R) = 0 \).

Since
\[
\mathbb{E}(Y(Z,M)|X,R,\delta) = \beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R + \beta_u \delta,
\]
we have
\[
\mathbb{E}(Y(Z,M)|X,R) = \int (\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R) \, dP(\delta|Z,M,X,R),
= (\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R) + \int (\beta_u \delta) \, dP(\delta|Z,M,X,R),
= (\beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R) + \beta_u E(E(\delta|R)|Z,M,X,R),
= \beta_0 + \beta_z Z + \beta_m M + \beta_x X + \beta_u \tau R,
\]
(14)
hence the 2SRI estimator is consistent.

1.5 Testing the Exclusion Restriction assumption in the real data analysis

The Exclusion Restriction assumption states that the interaction \( Z \times X^{IV} \) affects the outcome only through its effect on the mediator \( M \), conditional on \( X \) and \( Z \). The Exclusion Restriction assumption cannot be formally tested. In the following, we will provide a partial test for this assumption. The direct effect of the treatment \( Z \) on the outcome \( Y \) can be actually through the pathway of some other intermediate variable, \( \bar{M} \). Such pathway can be visualized as
\[
Z \rightarrow \bar{M} \rightarrow Y
\]
(15)
For example, in the dental data, \( Z \) is the motivational interviewing and \( \bar{M} \) can be the kid’s dental visit and diet behavior other than the mediator of interest; in the flood data, \( Z \) is
the flood and $\bar{M}$ can be the mother’s health other than the mediator of interest. If the instrument $Z \times X_{IV}$ affects $\bar{M}$, conditioning on $Z, X$, then the instrument $Z \times X_{IV}$ can affect $Y$ through the mediator $\bar{M}$, which violates the Exclusion Restriction assumption. Therefore, conditioning on $Z$ and $X$, we can assess if $Z \times X_{IV}$ predicts $\bar{M}$ to evaluate if the Exclusion Restriction assumption of the instrument $Z \times X_{IV}$ is potentially violated. Even if $Z \times X_{IV}$ does not significantly affect $\bar{M}$, we could not conclude that the Exclusion Restriction assumption is verified. However, we are more confident in the plausibility of the Exclusion Restriction assumption for that it is not violated through the intermediate variable $\bar{M}$. As illustrated in Figure 1, we test whether the dotted arrow exists.

[Figure 1 about here.]

Let $h$ denote the link function for $E (\bar{M}^{z^*} | X = x)$ and define

$$
\gamma = h \left( E (\bar{M}^{z^*} | X = x) \right) - h \left( E (\bar{M}^z | X = x) \right).
$$

(16)

Formally, we test the following null hypothesis

$$
H_0 : \gamma \text{ is a function which does not depend on } z \times x_{IV}.
$$

(17)

When $\bar{M}$ is continuous and the link function $h$ is identity function with

$$
E (\bar{M}^{z^*} | X = x) = \nu_0 + \nu_z z^* + \nu_x x + \nu_{IV} z \times x_{IV}
$$

and hence

$$
E (\bar{M}^{z^*} - \bar{M}^z | X = x) = \nu_{IV} (z^* - z) x_{IV}.
$$

In this case, we are testing

$$
H_0 : \nu_{IV} = 0.
$$

(18)

When $\bar{M}$ is binary and the link function $h$ is logit function with

$$
\logit \left( E (\bar{M}^{z^*} | X = x) \right) = \nu_0 + \nu_z z^* + \nu_x x + \nu_{IV} z \times x_{IV}
$$
and hence
\[
\logit \left( E(\tilde{M}^* | X = x) \right) - \logit \left( E(\tilde{M}^* | X = x) \right) = \nu_{IV} (z^* - z) x^{IV}.
\]
In this case, we also test
\[
H_0 : \nu_{IV} = 0. \quad (19)
\]

2. Web Appendix B

In this section, we consider the mediators \((m_1, m_2)\), where \(m_1\) and \(m_2\) are independent conditioning on \(z, m, x, u\) and construct the estimating equations for these two conditional independent mediators. The outcome model can be written as
\[
g\{E(Y(z, m_1, m_2)|z, m_1, m_2, x, u)\} = \beta_0 + \beta_z z + \beta_{m_1} m_1 + \beta_{m_2} m_2 + \beta_x x + \beta_u u. \quad (20)
\]
We assume there exist two valid IVs \(z \times x_1\) and \(z \times x_2\). We will focus on the harder case, binary mediator in this section. With the log link function, the model (20) represents Poisson, Negative Binomial and Neyman Type A distribution,
\[
Y|x, z, m, u \sim \text{Poisson} \left( \exp \left( \beta_0 + \beta_z z + \beta_{m_1} m_1 + \beta_{m_2} m_2 + \beta_x x + \beta_u u \right) \right),
\]
\[
Y|x, z, m, u \sim \text{NegBin} \left( \exp \left( \beta_0 + \beta_z z + \beta_{m_1} m_1 + \beta_{m_2} m_2 + \beta_x x + \beta_u u \right) \right).
\]
The binary mediators \(m_1\) and \(m_2\) are independent generated as
\[
m_1|x, z, u \sim \text{Ber} \left( \frac{\exp \left( \tau_0 + \tau_z z + \tau_{x_1} x_1 + \tau_{x_2} x_2 + \tau_{IV_1} z x_1 + \tau_{IV_2} z x_2 + \tau_u u \right)}{1 + \exp \left( \tau_0 + \tau_z z + \tau_{x_1} x_1 + \tau_{x_2} x_2 + \tau_{IV_1} z x_1 + \tau_{IV_2} z x_2 + \tau_u u \right)} \right),
\]
\[
m_2|x, z, u \sim \text{Ber} \left( \frac{\exp \left( \tau_0 + \tau_z z + \tau_{x_1} x_1 + \tau_{x_2} x_2 + \tau_{IV_1} z x_1 + \tau_{IV_2} z x_2 + \tau_u u \right)}{1 + \exp \left( \tau_0 + \tau_z z + \tau_{x_1} x_1 + \tau_{x_2} x_2 + \tau_{IV_1} z x_1 + \tau_{IV_2} z x_2 + \tau_u u \right)} \right).
\]
We can establish the following estimating equations,

\[ h_1(w, \theta) = \left( \frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m_1 + \beta_m m_2 + \beta_x x_1 + \beta_x x_2)} - 1 \right) ; \]

\[ h_2(w, \theta) = \left( \frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m_1 + \beta_m m_2 + \beta_x x_1 + \beta_x x_2)} - 1 \right) z ; \]

\[ h_3(w, \theta) = \left( \frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m_1 + \beta_m m_2 + \beta_x x_1 + \beta_x x_2)} - 1 \right) x_1 ; \]

\[ h_4(w, \theta) = \left( \frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m_1 + \beta_m m_2 + \beta_x x_1 + \beta_x x_2)} - 1 \right) x_2 ; \]

\[ h_5(w, \theta) = \left( \frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m_1 + \beta_m m_2 + \beta_x x_1 + \beta_x x_2)} - 1 \right) z x_1 ; \]

\[ h_6(w, \theta) = \left( \frac{y}{\exp(\beta_0 + \beta_z z + \beta_m m_1 + \beta_m m_2 + \beta_x x_1 + \beta_x x_2)} - 1 \right) z x_2 ; \]

\[ h_7(w, \theta) = \left( \frac{y}{\exp(\beta_m m_1 + \beta_m m_2)} - \exp(\beta_0 + \beta_z z + \beta_x x_1 + \beta_x x_2) \right) ; \]

\[ h_8(w, \theta) = \left( \frac{y}{\exp(\beta_m m_1 + \beta_m m_2)} - \exp(\beta_0 + \beta_z z + \beta_x x_1 + \beta_x x_2) \right) z ; \]

\[ h_9(w, \theta) = \left( \frac{y}{\exp(\beta_m m_1 + \beta_m m_2)} - \exp(\beta_0 + \beta_z z + \beta_x x_1 + \beta_x x_2) \right) x_1 ; \]

\[ h_{10}(w, \theta) = \left( \frac{y}{\exp(\beta_m m_1 + \beta_m m_2)} - \exp(\beta_0 + \beta_z z + \beta_x x_1 + \beta_x x_2) \right) x_2 ; \]

\[ h_{11}(w, \theta) = \left( \frac{y}{\exp(\beta_m m_1 + \beta_m m_2)} - \exp(\beta_0 + \beta_z z + \beta_x x_1 + \beta_x x_2) \right) z x_1 ; \]

\[ h_{12}(w, \theta) = \left( \frac{y}{\exp(\beta_m m_1 + \beta_m m_2)} - \exp(\beta_0 + \beta_z z + \beta_x x_1 + \beta_x x_2) \right) z x_2 . \]

(21)

3. Web Appendix C

In this section, we discuss the more simulation results.

3.1 Single Mediator with Two Instrumental Variables

[Table 1 about here.]

3.2 More simulation results for Continuous Mediator

[Table 2 about here.]
3.3 *Simulation results for sensitivity analysis*

[Table 3 about here.]

3.4 *A larger proportion of zeros for Poisson and Negative Binomial*

We simulate the Poisson and Negative Binomial outcome model with a larger proportion of zero.

\[
y|x, z, m, u \sim \text{Poisson} \left( \exp \left( -0.5 + 0.5z + 0.5m + 0.5x + u \right) \right). \tag{22}
\]

\[
y|x, z, m, u \sim \text{NegBin} \left( \exp \left( -0.5 + 0.5z + 0.5m + 0.5x + u \right) \right). \tag{23}
\]

The proportion of zeros for poisson increases from 20% to 50% and for Negative Binomial from 25% to 55%.

[Table 4 about here.]

3.5 *Robust to the outcome distribution*

It is necessary to know the outcome model for 2SRI second stage regression, which is another challenge for applying 2SRI to real data analysis. In Table 5, we generate the data by Negative Binomial Outcome model while fitting the second stage with Poisson Outcome. Table 5 shows that the proposed estimating equation approach consistently estimates the treatment and mediation effects while 2SRI estimates have a large bias, which illustrates that our method does not rely on the distribution of outcome model.

[Table 5 about here.]

3.6 *Comparison of Two Estimating Equation Methods*

[Table 6 about here.]

[Table 7 about here.]

[Table 8 about here.]
References

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data, MIT press.
Figure 1: Causal pathway of testing the Exclusion Restriction assumption
Table 1: EL estimate (with Multi-starting values) and 2SRI estimate for the direct effect parameter ($\beta_1$) and the indirect effect parameter ($\beta_2$) with two instrumental variables.

| Dis. | n  | Direct                |              |               | Indirect               |              |
|------|----|-----------------------|--------------|--------------|-----------------------|--------------|
|      |    | EL Med. | 2SRI Med. | Reg Med. | EL Med. | 2SRI Med. | Reg Med. |
|      |    | (MAD)    | (MAD)      | (MAD)    | (MAD)    | (MAD)      | (MAD)    |
| Poi  | 500 | 0.484   | 0.478      | 0.388    | 0.527   | 0.592      | 0.971    |
|      |     | (0.169) | (0.198)    | (0.117)  | (0.868) | (0.671)    | (0.115)  |
| Poi  | 1000| 0.506   | 0.483      | 0.385    | 0.466   | 0.598      | 0.965    |
|      |     | (0.131) | (0.148)    | (0.079)  | (0.611) | (0.514)    | (0.078)  |
| Poi  | 5000| 0.506   | 0.502      | 0.389    | 0.494   | 0.548      | 0.962    |
|      |     | (0.050) | (0.064)    | (0.036)  | (0.265) | (0.253)    | (0.035)  |
| NB   | 500 | 0.497   | 0.485      | 0.415    | 0.525   | 0.569      | 0.952    |
|      |     | (0.213) | (0.187)    | (0.126)  | (1.019) | (0.685)    | (0.132)  |
| NB   | 1000| 0.491   | 0.494      | 0.413    | 0.491   | 0.524      | 0.951    |
|      |     | (0.153) | (0.122)    | (0.081)  | (0.773) | (0.510)    | (0.089)  |
| NB   | 5000| 0.499   | 0.493      | 0.414    | 0.542   | 0.541      | 0.955    |
|      |     | (0.064) | (0.058)    | (0.038)  | (0.309) | (0.235)    | (0.042)  |
| NTA  | 500 | 0.526   | 0.342      | 0.367    | 0.285   | 1.029      | 1.003    |
|      |     | (0.336) | (0.813)    | (0.260)  | (1.817) | (2.262)    | (0.236)  |
| NTA  | 1000| 0.512   | 0.337      | 0.358    | 0.463   | 1.112      | 1.010    |
|      |     | (0.280) | (0.616)    | (0.197)  | (1.451) | (1.643)    | (0.181)  |
| NTA  | 5000| 0.506   | 0.425      | 0.368    | 0.477   | 0.809      | 0.995    |
|      |     | (0.145) | (0.407)    | (0.092)  | (0.651) | (1.147)    | (0.091)  |

The median (out of parenthesis) and the MAD (inside parenthesis) are reported. EL denotes the Empirical Likelihood estimate, 2SRI denotes the 2SRI estimates and Reg denotes the ordinary (Poisson or Negative Binomial) regression estimate. n stands for sample size; Poi stands for Poisson distribution. NB stands for Negative Binomial outcome distribution and NTA stands for Neyman Type A distribution outcome. The simulation time is 1000 and the true coefficients are 0.5. The size of negative binomial model is 3.
Table 2: Continuous Mediator: EL estimate (with Multi-starting values), 2SRI estimate and Regression estimate without IV (Reg) for the direct effect parameter ($\beta_1$) and the indirect effect parameter ($\beta_2$) with one instrumental variable.

| Dist. | Str. | n   | EL Med. (MAD) | 2SRI Med. (MAD) | Reg Med. (MAD) | EL Med. (MAD) | 2SRI Med. (MAD) | Reg Med. (MAD) |
|-------|------|-----|---------------|-----------------|---------------|---------------|----------------|---------------|
| Poi   | S    | 500 | 0.498 (0.151) | 0.519 (0.212)  | 0.146 (0.143) | 0.502 (0.136) | 0.478 (0.190) | 0.771 (0.079)  |
| Poi   | S    | 1000| 0.505 (0.107) | 0.515 (0.154)  | 0.153 (0.107) | 0.494 (0.093) | 0.492 (0.140) | 0.772 (0.056)  |
| Poi   | S    | 5000| 0.499 (0.047) | 0.507 (0.083)  | 0.146 (0.050) | 0.500 (0.043) | 0.497 (0.067) | 0.778 (0.031)  |
| Poi   | W    | 500 | 0.498 (0.191) | 0.506 (0.249)  | 0.235 (0.131) | 0.498 (0.243) | 0.488 (0.328) | 0.803 (0.051)  |
| Poi   | W    | 1000| 0.498 (0.130) | 0.495 (0.192)  | 0.230 (0.093) | 0.503 (0.170) | 0.495 (0.242) | 0.803 (0.051)  |
| Poi   | W    | 5000| 0.498 (0.060) | 0.506 (0.090)  | 0.232 (0.044) | 0.497 (0.079) | 0.494 (0.118) | 0.800 (0.027)  |
| NB    | S    | 500 | 0.501 (0.168) | 0.504 (0.166)  | 0.261 (0.135) | 0.501 (0.151) | 0.499 (0.142) | 0.822 (0.052)  |
| NB    | S    | 1000| 0.493 (0.122) | 0.493 (0.110)  | 0.249 (0.094) | 0.505 (0.105) | 0.500 (0.096) | 0.820 (0.039)  |
| NB    | S    | 5000| 0.500 (0.053) | 0.501 (0.053)  | 0.258 (0.040) | 0.502 (0.048) | 0.496 (0.044) | 0.821 (0.016)  |
| NB    | W    | 500 | 0.508 (0.221) | 0.508 (0.214)  | 0.285 (0.131) | 0.510 (0.310) | 0.492 (0.304) | 0.863 (0.060)  |
| NB    | W    | 1000| 0.491 (0.156) | 0.494 (0.145)  | 0.276 (0.096) | 0.514 (0.208) | 0.503 (0.195) | 0.865 (0.042)  |
| NB    | W    | 5000| 0.500 (0.067) | 0.499 (0.063)  | 0.275 (0.041) | 0.500 (0.093) | 0.495 (0.091) | 0.863 (0.017)  |
| NTA   | S    | 500 | 0.509 (0.464) | 0.535 (0.572)  | -0.086 (0.330) | 0.482 (0.421) | 0.481 (0.345) | 0.786 (0.139)  |
| NTA   | S    | 1000| 0.495 (0.343) | 0.550 (0.514)  | -0.100 (0.254) | 0.510 (0.315) | 0.495 (0.311) | 0.786 (0.113)  |
| NTA   | S    | 5000| 0.489 (0.154) | 0.532 (0.315)  | -0.124 (0.172) | 0.503 (0.138) | 0.490 (0.168) | 0.785 (0.069)  |
| NTA   | W    | 500 | 0.521 (0.556) | 0.508 (0.707)  | 0.104 (0.268) | 0.473 (0.748) | 0.481 (0.690) | 0.808 (0.124)  |
| NTA   | W    | 1000| 0.493 (0.417) | 0.531 (0.544)  | 0.090 (0.195) | 0.511 (0.572) | 0.472 (0.513) | 0.803 (0.095)  |
| NTA   | W    | 5000| 0.517 (0.225) | 0.521 (0.341)  | 0.091 (0.118) | 0.484 (0.280) | 0.491 (0.313) | 0.804 (0.060)  |
Table 3: Sensitivity Analysis: EL estimate (with Multi-starting values), 2SRI estimate and Regression estimate without IV (Reg) for the direct effect parameter ($\beta_1$) and the indirect effect parameter ($\beta_2$) with one instrumental variable. Poisson with strong IV and NTA with weak IV.

| Dis. | $\theta$ | n  | Direct | Indirect |
|------|----------|----|--------|----------|
|      |          |    | EL     | 2SRI     | Reg      | EL     | 2SRI     | Reg      |
|      |          |    | Med. (MAD) | Med. (MAD) | Med. (MAD) | Med. (MAD) | Med. (MAD) | Med. (MAD) |
| Poi  | $\eta = +0.1$ | 1000 | 0.507 | 0.501 | 0.415 | 0.495 | 0.559 | 0.954 |
|      |          |    | (0.099) | (0.119) | (0.079) | (0.512) | (0.494) | (0.075) |
| Poi  | $\eta = -0.1$ | 1000 | 0.501 | 0.496 | 0.412 | 0.477 | 0.566 | 0.952 |
|      |          |    | (0.090) | (0.105) | (0.079) | (0.510) | (0.495) | (0.080) |
| Poi  | $\eta = +0.5$ | 1000 | 0.492 | 0.494 | 0.419 | 0.525 | 0.558 | 0.950 |
|      |          |    | (0.104) | (0.113) | (0.081) | (0.560) | (0.532) | (0.081) |
| Poi  | $\eta = -0.5$ | 1000 | 0.498 | 0.501 | 0.417 | 0.512 | 0.570 | 0.952 |
|      |          |    | (0.093) | (0.112) | (0.082) | (0.478) | (0.485) | (0.077) |
| Poi  | $\eta = +1.0$ | 1000 | 0.488 | 0.495 | 0.412 | 0.552 | 0.564 | 0.952 |
|      |          |    | (0.110) | (0.108) | (0.080) | (0.615) | (0.480) | (0.080) |
| Poi  | $\eta = -1.0$ | 1000 | 0.504 | 0.492 | 0.417 | 0.490 | 0.560 | 0.950 |
|      |          |    | (0.092) | (0.115) | (0.081) | (0.524) | (0.531) | (0.084) |
| NegB | $\eta = +0.1$ | 1000 | 0.488 | 0.500 | 0.446 | 0.541 | 0.564 | 0.949 |
|      |          |    | (0.119) | (0.121) | (0.091) | (0.574) | (0.521) | (0.096) |
| NegB | $\eta = -0.1$ | 1000 | 0.492 | 0.496 | 0.440 | 0.494 | 0.511 | 0.944 |
|      |          |    | (0.113) | (0.114) | (0.088) | (0.612) | (0.535) | (0.091) |
| NegB | $\eta = +0.5$ | 1000 | 0.494 | 0.500 | 0.442 | 0.510 | 0.504 | 0.943 |
|      |          |    | (0.117) | (0.113) | (0.088) | (0.609) | (0.515) | (0.098) |
| NegB | $\eta = -0.5$ | 1000 | 0.502 | 0.504 | 0.449 | 0.487 | 0.514 | 0.939 |
|      |          |    | (0.116) | (0.117) | (0.085) | (0.643) | (0.561) | (0.096) |
| NegB | $\eta = +1.0$ | 1000 | 0.488 | 0.500 | 0.448 | 0.562 | 0.517 | 0.947 |
|      |          |    | (0.133) | (0.120) | (0.094) | (0.749) | (0.558) | (0.098) |
| NegB | $\eta = -1.0$ | 1000 | 0.494 | 0.501 | 0.445 | 0.511 | 0.528 | 0.941 |
|      |          |    | (0.116) | (0.119) | (0.090) | (0.643) | (0.548) | (0.101) |
| NTA  | $\eta_1 = 0.05$ | 1000 | 0.486 | 0.415 | 0.362 | 0.505 | 0.843 | 0.991 |
|      | $\eta_2 = 0.05$ |    | (0.234) | (0.389) | (0.177) | (1.313) | (1.363) | (0.161) |
| NTA  | $\eta_1 = 0.05$ | 5000 | 0.503 | 0.457 | 0.368 | 0.497 | 0.684 | 0.984 |
|      | $\eta_2 = 0.05$ |    | (0.117) | (0.249) | (0.081) | (0.684) | (0.746) | (0.072) |
| NTA  | $\eta_1 = 0.10$ | 1000 | 0.495 | 0.459 | 0.370 | 0.489 | 0.619 | 0.983 |
|      | $\eta_2 = 0.10$ |    | (0.225) | (0.429) | (0.183) | (1.295) | (1.374) | (0.165) |
| NTA  | $\eta_1 = 0.10$ | 5000 | 0.502 | 0.468 | 0.366 | 0.493 | 0.665 | 0.983 |
|      | $\eta_2 = 0.10$ |    | (0.123) | (0.235) | (0.091) | (0.701) | (0.700) | (0.073) |
| NTA  | $\eta_1 = 0.25$ | 1000 | 0.482 | 0.422 | 0.370 | 0.616 | 0.832 | 0.987 |
|      | $\eta_2 = 0.25$ |    | (0.270) | (0.428) | (0.172) | (1.450) | (1.447) | (0.151) |
| NTA  | $\eta_1 = 0.25$ | 5000 | 0.491 | 0.470 | 0.378 | 0.571 | 0.653 | 0.979 |
|      | $\eta_2 = 0.25$ |    | (0.136) | (0.236) | (0.091) | (0.743) | (0.791) | (0.074) |
Table 4: EL estimate (with Multi-starting values) and 2SRI estimate for the direct effect parameter ($\beta_1$) and the indirect effect parameter ($\beta_2$) with one instrumental variable.

| Dis. | Str. | n  | EL Med. | EL MAD | 2SRI Med. | 2SRI MAD | Indirect EL Med. | Indirect EL MAD | Indirect 2SRI Med. | Indirect 2SRI MAD |
|------|------|----|---------|--------|------------|----------|-----------------|----------------|-------------------|------------------|
| Poi  | S    | 500| 0.489   | (0.188)| 0.489      | (0.185)  | 0.492           | (1.009)        | 0.566             | (0.830)          |
| Poi  | S    | 1000| 0.494  | (0.132)| 0.494      | (0.126)  | 0.500           | (0.692)        | 0.540             | (0.597)          |
| Poi  | S    | 5000| 0.501  | (0.062)| 0.504      | (0.062)  | 0.501           | (0.328)        | 0.537             | (0.270)          |
| Poi  | W    | 500| 0.487  | (0.205)| 0.485      | (0.234)  | 0.642           | (1.397)        | 0.686             | (1.403)          |
| Poi  | W    | 1000| 0.495  | (0.170)| 0.488      | (0.162)  | 0.514           | (1.145)        | 0.650             | (1.010)          |
| Poi  | W    | 5000| 0.500  | (0.081)| 0.497      | (0.077)  | 0.487           | (0.553)        | 0.547             | (0.478)          |
| NB   | S    | 500| 0.497  | (0.200)| 0.519      | (0.187)  | 0.503           | (1.116)        | 0.521             | (0.841)          |
| NB   | S    | 1000| 0.487 | (0.151)| 0.493      | (0.142)  | 0.530           | (0.780)        | 0.579             | (0.635)          |
| NB   | S    | 5000| 0.500  | (0.066)| 0.504      | (0.059)  | 0.495           | (0.363)        | 0.519             | (0.277)          |
| NB   | W    | 500| 0.477  | (0.229)| 0.480      | (0.251)  | 0.539           | (1.579)        | 0.595             | (1.427)          |
| NB   | W    | 1000| 0.485 | (0.165)| 0.489      | (0.180)  | 0.470           | (1.200)        | 0.535             | (1.054)          |
| NB   | W    | 5000| 0.491  | (0.089)| 0.498      | (0.081)  | 0.547           | (0.604)        | 0.538             | (0.487)          |

We report the median and the MAD of the estimates. The column indexed with EL denotes the corresponding Empirical Likelihood estimate while the column indexed with 2S denotes the corresponding 2SRI estimates. The column indexed with $n$ stands for sample size; the column indexed with Dis. represents the conditional distribution of the outcome, where Poi stands for Poisson distribution, NB stands for Negative Binomial outcome distribution and NTA stands for Neyman Type A distribution outcome. The column indexed with Str. represents the strength of instrumental variables, where $S$ stands for stronger IV (setting 1) while $W$ stands for relatively weaker IV (setting 2). The simulation time is 1000 and the true coefficients are 0.5. The size of negative binomial model is 3.
Table 5: Fit the second stage with a wrong model.

|       | Direct |               |               | Indirect |               |               |
|-------|--------|---------------|---------------|----------|---------------|---------------|
|       |        | EL            | 2SRI          | EL       | 2SRI          |               |
| Dis.  | Str.   | n  | Med. | MAD | Med. | MAD | Med. | MAD | Med. | MAD |
| NB    | S      | 500 | 0.487 (0.163) | 0.491 (0.181) | 0.539 (0.939) | 0.574 (0.940) |
| NB    | S      | 1000 | 0.493 (0.112) | 0.494 (0.127) | 0.491 (0.655) | 0.532 (0.630) |
| NB    | S      | 5000 | 0.500 (0.049) | 0.498 (0.059) | 0.492 (0.257) | 0.548 (0.284) |
| NB    | W      | 500 | 0.481 (0.177) | 0.494 (0.225) | 0.585 (1.224) | 0.545 (1.485) |
| NB    | W      | 1000 | 0.484 (0.146) | 0.487 (0.177) | 0.560 (1.003) | 0.594 (1.131) |
| NB    | W      | 5000 | 0.500 (0.064) | 0.496 (0.073) | 0.490 (0.445) | 0.564 (0.476) |

The outcome follows Negative Binomial outcome while we fit the second stage model with Poisson outcome. We report the median and MAD of the estimates. The sample size n is 500, 1000 or 5000.
Table 6: Two EL estimators (EE-EL1 and EE-EL2), 2SRI estimate and Regression estimate without IV (Reg) for the direct effect parameter ($\beta_1$) and the indirect effect parameter ($\beta_2$) with one instrumental variable.

|                | EE-EL1 | EE-EL2 | 2SRI | Reg | EE-EL1 | EE-EL2 | 2SRI | Reg |
|----------------|--------|--------|------|-----|--------|--------|------|-----|
| **Poisson**    |        |        |      |     |        |        |      |     |
| Median         | 0.501  | 0.499  | 0.499| 0.416| 0.497  | 0.508  | 0.544| 0.950|
| MAD            | 0.042  | 0.040  | 0.047| 0.037| 0.231  | 0.213  | 0.241| 0.031|
| **NB**         |        |        |      |     |        |        |      |     |
| Median         | 0.497  | 0.496  | 0.500| 0.445| 0.498  | 0.516  | 0.514| 0.942|
| MAD            | 0.049  | 0.046  | 0.050| 0.038| 0.274  | 0.253  | 0.233| 0.041|
| **NTA**        |        |        |      |     |        |        |      |     |
| Median         | 0.496  | 0.495  | 0.462| 0.375| 0.521  | 0.545  | 0.678| 0.981|
| MAD            | 0.116  | 0.110  | 0.245| 0.088| 0.683  | 0.510  | 0.787| 0.074|
Table 7: Count outcome and binary mediator: Comparison of 2SPS estimate and 2SRI estimate for the direct effect parameter ($\beta_1$) and the indirect effect parameter ($\beta_2$) with one instrumental variable.

| Outcome | IV | n  | Direct 2SRI Med. (MAD) | Direct 2SPS Med. (MAD) | Indirect 2SRI Med. (MAD) | Indirect 2SPS Med. (MAD) |
|---------|----|----|------------------------|------------------------|--------------------------|--------------------------|
| Poi     | S  | 500| 0.487 (0.149)          | 0.478 (0.151)          | 0.609 (0.660)            | 0.463 (0.694)            |
|         |    |    |                        |                        |                          |                          |
|         | S  | 1000| 0.492 (0.099)          | 0.484 (0.100)          | 0.570 (0.500)            | 0.458 (0.498)            |
| Poi     | S  | 5000| 0.500 (0.047)          | 0.493 (0.047)          | 0.552 (0.229)            | 0.431 (0.226)            |
|         |    |    |                        |                        |                          |                          |
|         | W  | 500| 0.483 (0.181)          | 0.490 (0.181)          | 0.692 (1.146)            | 0.528 (1.138)            |
| Poi     | W  | 1000| 0.495 (0.137)          | 0.501 (0.141)          | 0.582 (0.897)            | 0.425 (0.888)            |
|         |    | 5000| 0.497 (0.064)          | 0.509 (0.066)          | 0.539 (0.395)            | 0.362 (0.404)            |
| NB      | S  | 500| 0.500 (0.161)          | 0.500 (0.160)          | 0.543 (0.735)            | 0.460 (0.771)            |
|         |    |    |                        |                        |                          |                          |
|         | S  | 1000| 0.496 (0.105)          | 0.489 (0.109)          | 0.570 (0.504)            | 0.488 (0.506)            |
| NB      | S  | 5000| 0.503 (0.051)          | 0.494 (0.053)          | 0.499 (0.236)            | 0.432 (0.239)            |
|         |    |    |                        |                        |                          |                          |
|         | W  | 500| 0.496 (0.197)          | 0.513 (0.211)          | 0.423 (1.256)            | 0.355 (1.228)            |
| NB      | W  | 1000| 0.489 (0.130)          | 0.493 (0.134)          | 0.591 (0.840)            | 0.443 (0.849)            |
|         |    | 5000| 0.492 (0.059)          | 0.505 (0.059)          | 0.551 (0.378)            | 0.438 (0.391)            |
| NTA     | S  | 500| 0.393 (0.501)          | 0.394 (0.497)          | 0.786 (1.789)            | 0.619 (1.817)            |
|         |    |    |                        |                        |                          |                          |
|         | S  | 1000| 0.413 (0.398)          | 0.404 (0.413)          | 0.812 (1.308)            | 0.641 (1.355)            |
| NTA     | S  | 5000| 0.465 (0.260)          | 0.455 (0.261)          | 0.653 (0.821)            | 0.500 (0.800)            |
|         |    |    |                        |                        |                          |                          |
|         | W  | 500| 0.411 (0.599)          | 0.390 (0.614)          | 1.028 (3.045)            | 0.918 (3.314)            |
| NTA     | W  | 1000| 0.422 (0.495)          | 0.421 (0.504)          | 0.899 (2.338)            | 0.793 (2.377)            |
|         |    | 5000| 0.442 (0.292)          | 0.431 (0.298)          | 0.834 (1.349)            | 0.687 (1.371)            |
Table 8: Count outcome and normal mediator: Comparison of 2SPS estimate and 2SRI estimate for the direct effect parameter ($\beta_1$) and the indirect effect parameter ($\beta_2$) with one instrumental variable.

| Outcome | IV | n   | Direct 2SRI Med. (MAD) | Direct 2SPS Med. (MAD) | Indirect 2SRI Med. (MAD) | Indirect 2SPS Med. (MAD) |
|---------|----|-----|-------------------------|------------------------|--------------------------|--------------------------|
| Poi     | S  | 500 | 0.508 (0.207)           | 0.523 (0.272)          | 0.496 (0.171)            | 0.477 (0.243)            |
| Poi     | S  | 1000| 0.515 (0.157)           | 0.533 (0.202)          | 0.497 (0.131)            | 0.476 (0.168)            |
| Poi     | S  | 5000| 0.507 (0.086)           | 0.516 (0.113)          | 0.498 (0.070)            | 0.493 (0.095)            |
| Poi     | W  | 500 | 0.511 (0.249)           | 0.509 (0.314)          | 0.497 (0.321)            | 0.469 (0.377)            |
| Poi     | W  | 1000| 0.521 (0.185)           | 0.532 (0.233)          | 0.484 (0.238)            | 0.479 (0.300)            |
| Poi     | W  | 5000| 0.500 (0.088)           | 0.500 (0.118)          | 0.501 (0.115)            | 0.501 (0.154)            |
| NB      | S  | 500 | 0.505 (0.166)           | 0.493 (0.196)          | 0.497 (0.143)            | 0.498 (0.163)            |
| NB      | S  | 1000| 0.503 (0.115)           | 0.496 (0.132)          | 0.496 (0.105)            | 0.508 (0.112)            |
| NB      | S  | 5000| 0.499 (0.056)           | 0.499 (0.062)          | 0.498 (0.043)            | 0.500 (0.052)            |
| NB      | W  | 500 | 0.505 (0.224)           | 0.515 (0.262)          | 0.469 (0.273)            | 0.493 (0.324)            |
| NB      | W  | 1000| 0.496 (0.146)           | 0.499 (0.163)          | 0.508 (0.188)            | 0.505 (0.235)            |
| NB      | W  | 5000| 0.499 (0.067)           | 0.503 (0.079)          | 0.496 (0.090)            | 0.504 (0.103)            |
| NTA     | S  | 500 | 0.535 (0.540)           | 0.654 (0.771)          | 0.475 (0.340)            | 0.405 (0.480)            |
| NTA     | S  | 1000| 0.558 (0.459)           | 0.640 (0.707)          | 0.479 (0.284)            | 0.445 (0.411)            |
| NTA     | S  | 5000| 0.546 (0.341)           | 0.609 (0.401)          | 0.480 (0.175)            | 0.453 (0.226)            |
| NTA     | W  | 500 | 0.550 (0.669)           | 0.584 (0.881)          | 0.463 (0.641)            | 0.475 (0.835)            |
| NTA     | W  | 1000| 0.513 (0.553)           | 0.599 (0.800)          | 0.495 (0.516)            | 0.410 (0.761)            |
| NTA     | W  | 5000| 0.484 (0.348)           | 0.547 (0.482)          | 0.518 (0.308)            | 0.466 (0.448)            |