Using the Multiple-Site, Multiple-Mediator Instrumental Variables (MSMM-IV) Model When the Exclusion Restriction is Invalid

sean f. reardon

July 2016

Goal

Two challenges to using IV models to estimate mediator effects:

- The exclusion retraction is often not defensible.
- Finite sample bias can be large in the multiple-site IV case.

Nonetheless:

- We want to obtain unbiased estimates of the effect of a mediator when the exclusion restriction is invalid.
- We want to correct the finite-sample bias in the IV estimator in this case.

Outline

1. Stylized Example and Intuition
2. The MSMM-IV and MSMM-INT models
3. Sources of Bias in the MSMM-IV/INT models
4. A “de-biased” MSMM-INT estimator

A Motivating Example:

Suppose we offer families a voucher to pay for up to $5,000 for pre-school. We want to know if this improves children's school readiness.

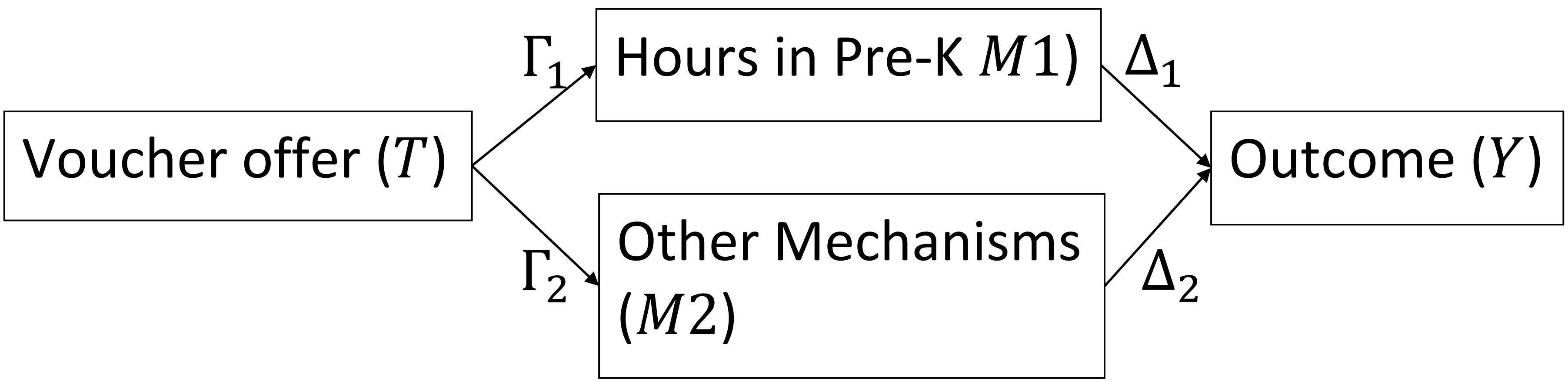
Under random assignment, it is straightforward to estimate the average effect of the voucher offer on school readiness.

In a multi-site trial, it may also be possible to estimate the variance in the effect of the voucher across sites.

But if we wish to know how the voucher affects readiness, **we need a theory, and a design that allows us to test that theory, and a statistical model that allows us to estimate the parameters of interest.**

A simple (simplistic) theory:

The voucher offer may induce a) more time in pre-school and/or b) affect other aspects of children's experiences that affect their reading skills.



How do we estimate the effects of Pre-K enrollment?

In particular, **how do we estimate it if we don't observe other potential mediators?**

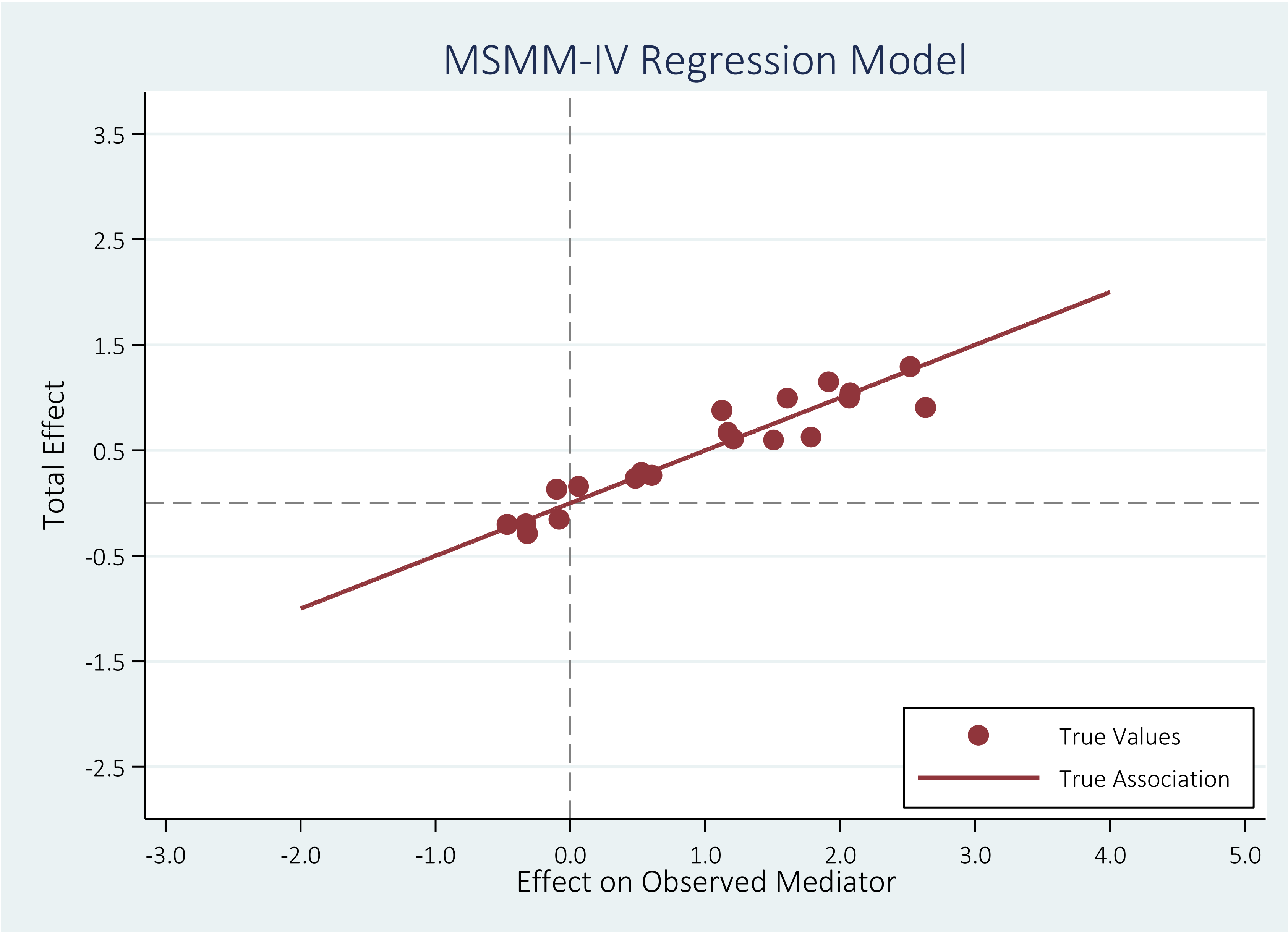
Suppose we conduct this experiment in a number of sites. In each site we can estimate both

a) The average effect of the HS offer on school readiness (β_s).
b) The average effects of the HS offer on reading instruction (γ_{1s}).

If we assume the exclusion restriction (that there is no pathway other than M1 through which T affects Y) is valid, then we can fit the MSMM-IV model:

$$\beta_s = \delta_1(\gamma_{1s}) + \omega_s$$

This is essentially an over-identified IV model using site-by-treatment interactions as instruments.



MSMM-IV Regression Model
Total Effect
3.5
2.5
1.5
0.5
-0.5
-1.5
-2.5
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
True Values
True Association

If, however, we do not wish to assume the exclusion restriction, we include an intercept in the model, and estimate the partial association between β_s and γ_{1s}, via the MSMM-INT regression model:

$$\beta_s = \lambda + \delta_1(\gamma_{1s}) + \omega_s$$

Here δ_1 describes the average difference in β between two sites that differ by one unit on γ_1, holding constant γ_2.

Under some assumptions, δ_1 is the average effect of $M1$ on Y.

Summary of MSMM-INT Assumptions

1) Stable unit treatment value assumptions
2) Person-specific linearity of the mediator(s) with respect to the treatment
3) Person-specific linearity of the outcome with respect to the mediator(s)
4) Parallel mediators (if multiple mediators)
5) Zero within-site compliance-effect covariance for each mediator
6) Within-site ignorable treatment assignment
7) Compliance matrix has sufficient rank (including intercept)
8) Between-site cross-mediator compliance-effect independence
9) Compliance-direct effect independence (no omitted mediator)

Sources of Bias in the MSMM-INT model

- Failure of one or more of the identifying assumptions
  - Cross-mediator compliance-effect covariance bias
  - Omitted mediator bias (compliance-direct effect covariance bias)
- Estimation error in the $\hat{\gamma}_s$'s
  - In the standard OLS context, measurement error in the X's leads to attenuation bias;
  - In the IV context, however, it leads to finite sample bias

Compare two models:

MSMM-IV Model

(no intercept, assumes the exclusion restriction is valid):

$$\beta_s = \delta\gamma_s + v_s$$

MSMM-IV Model with intercept

(includes an intercept, relaxes the exclusion restriction):

$$\beta_s = \lambda + \delta\gamma_s + v_s$$

Some notation

η: the OLS association between Y and M (the naïve estimate of δ)

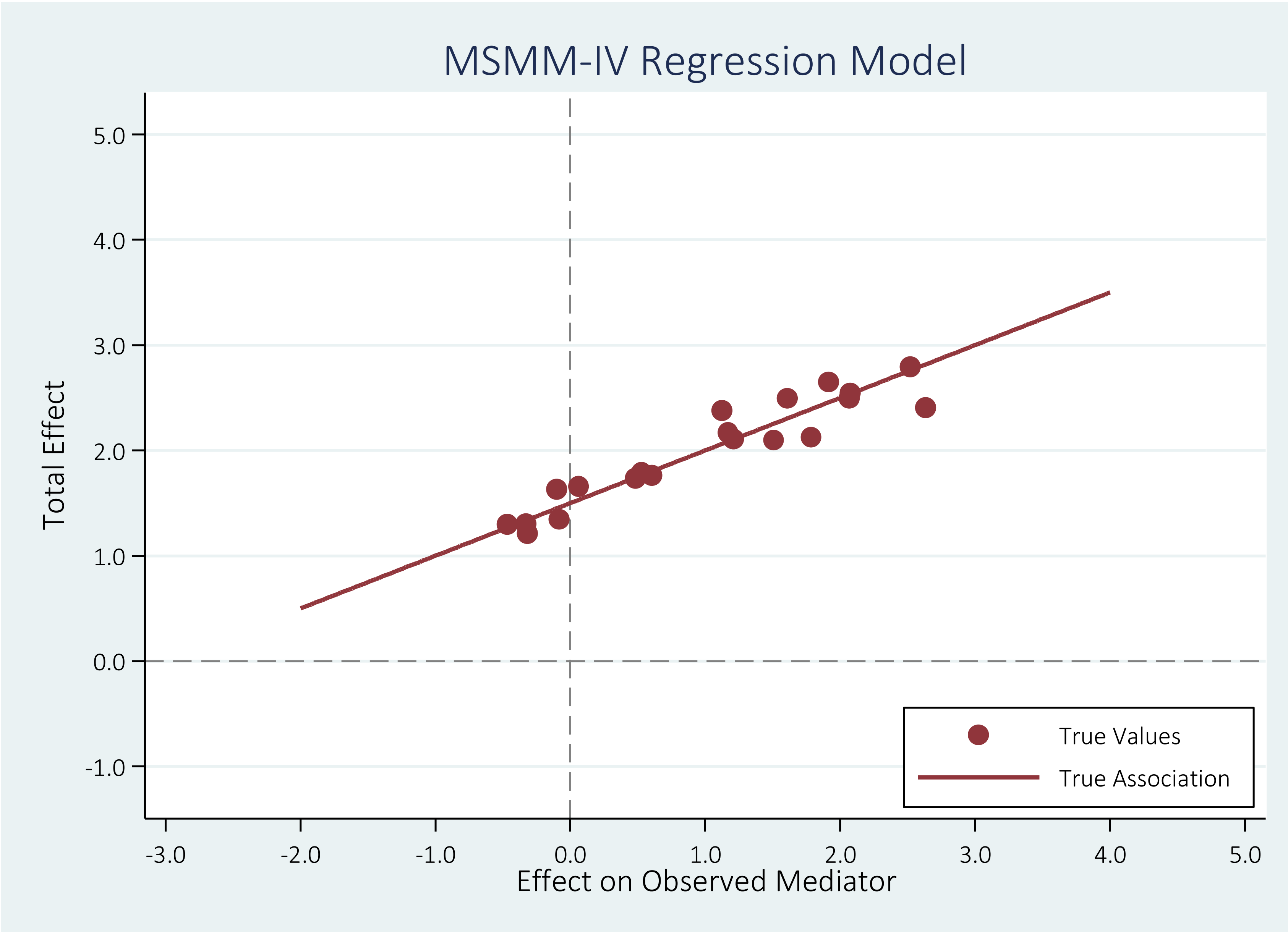
$\eta - \delta$: selection bias in the OLS association between Y and M

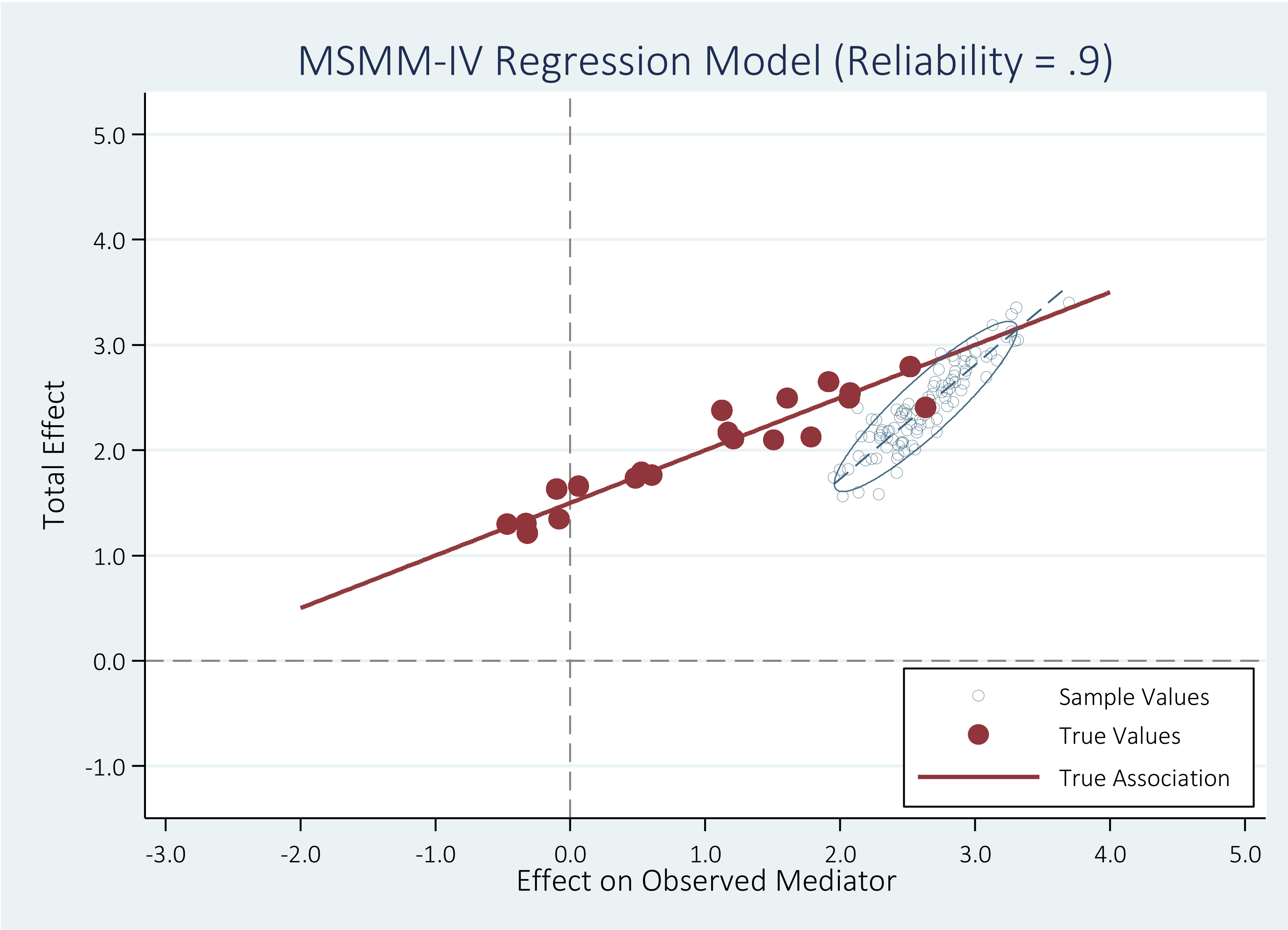
τ_γ: variance of the (true) γ_s's

$\tau_{\hat{\gamma}}$: variance of the (estimated) γ_s's (i.e., variance of $\hat{\gamma}_s$'s)

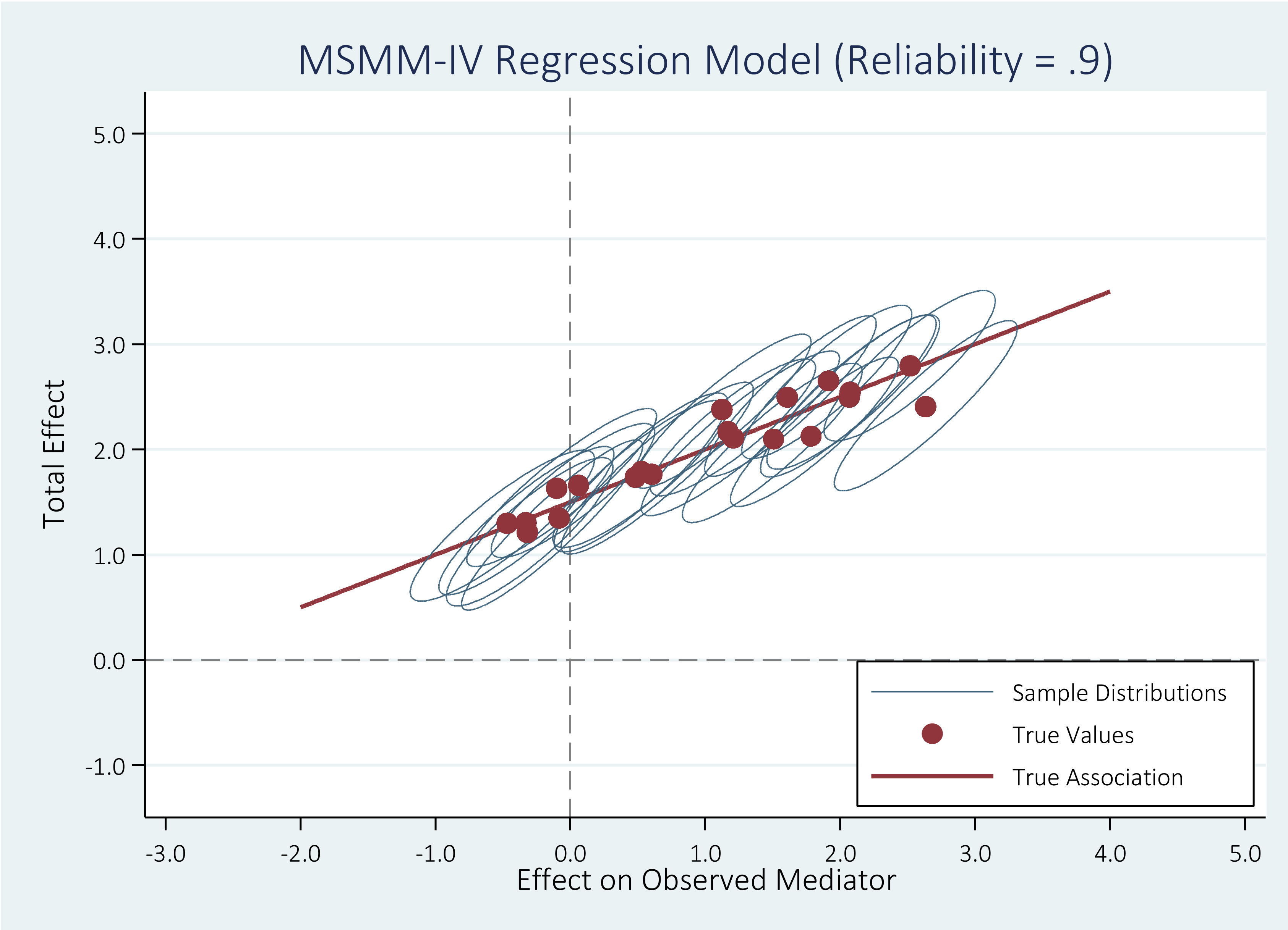
$r_\gamma = \frac{\tau_\gamma}{\tau_{\hat{\gamma}}} = \frac{\tau_\gamma}{\tau_\gamma + \tau_g}$: reliability of the $\hat{\gamma}_s$'s

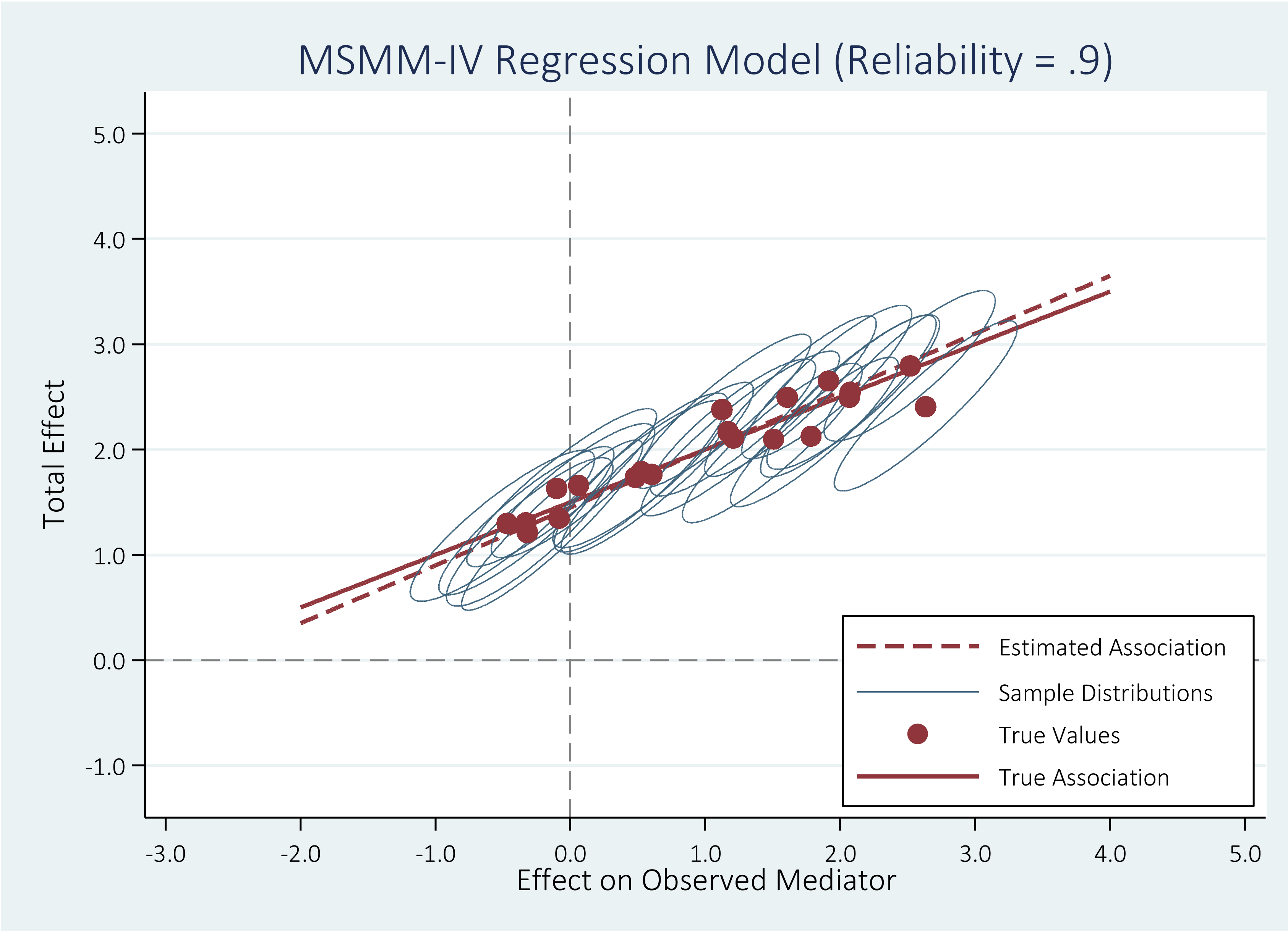
$F = \frac{(\gamma^2 + \tau_\gamma)}{\tau_g} + 1$: first-stage F-statistic

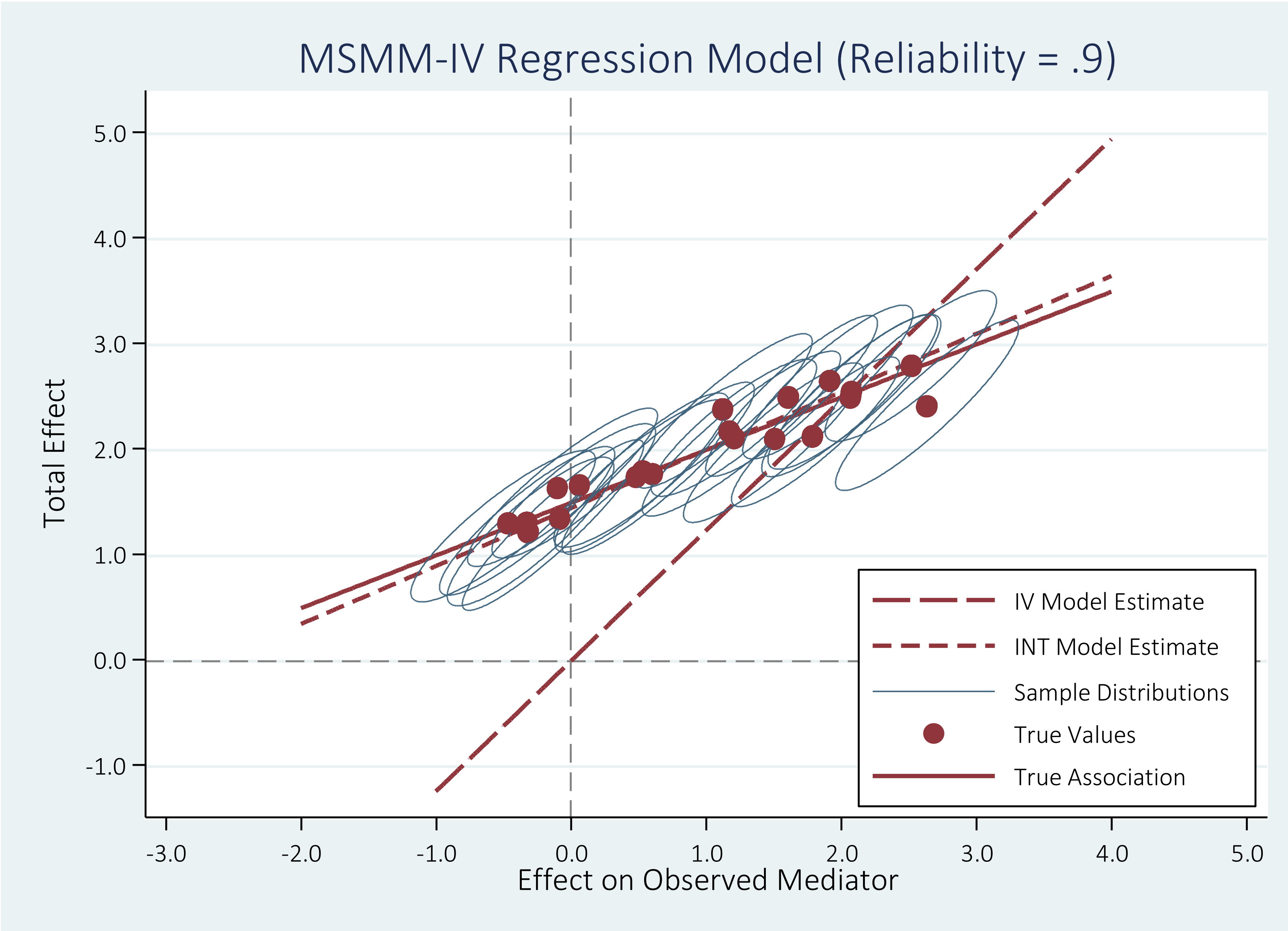

MSMM-IV Regression Model
Total Effect
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
True Values
True Association

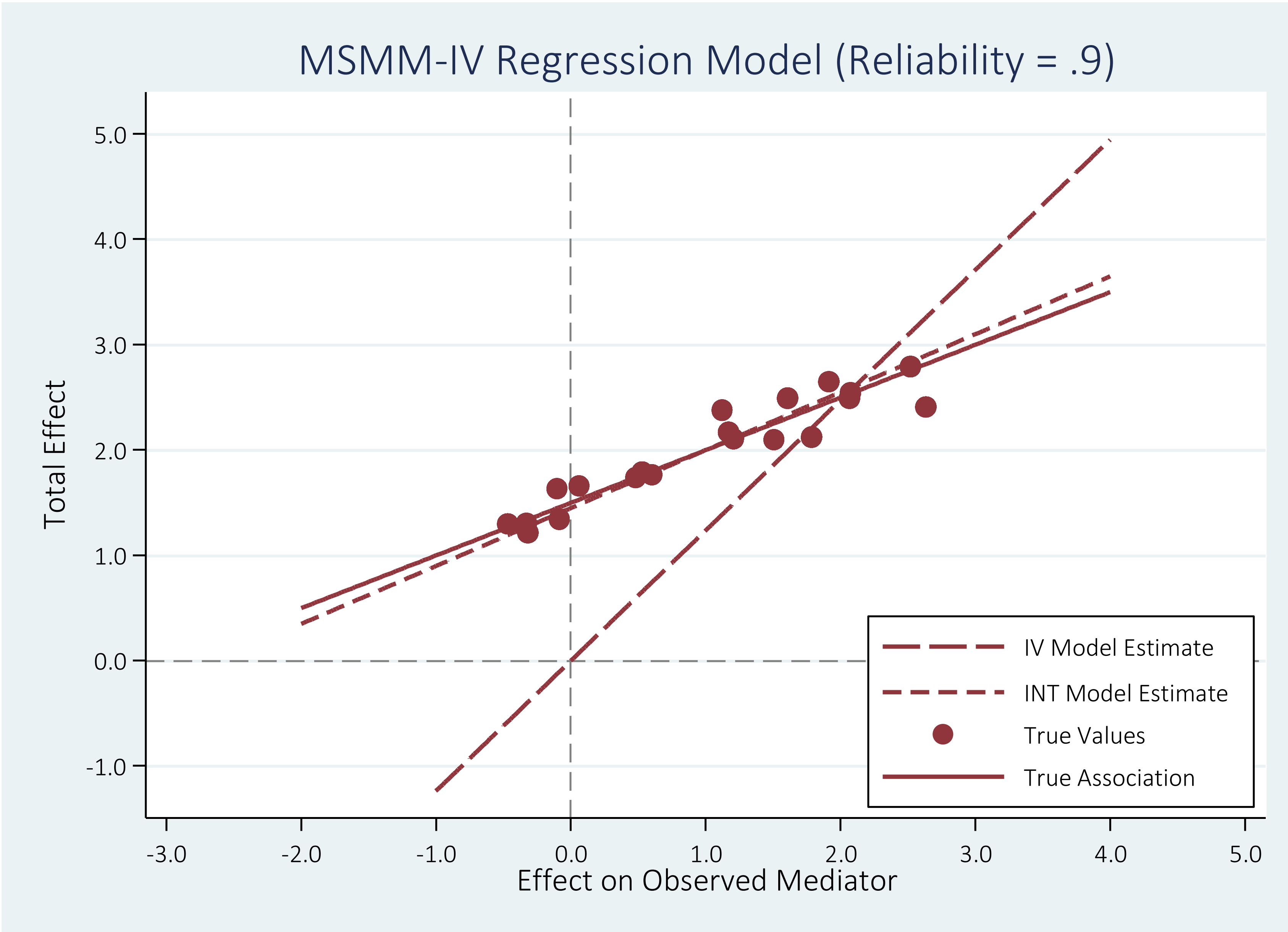


MSMM-IV Regression Model (Reliability = .9)
Total Effect
Effect on Observed Mediator
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Sample Values
True Values
True Association


MSMM-IV Regression Model (Reliability = .9)
Total Effect
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
Sample Distributions
True Values
True Association


MSMM-IV Regression Model (Reliability = .9)
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
Total Effect
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
Estimated Association
Sample Distributions
True Values
True Association


MSMM-IV Regression Model (Reliability = .9)
Total Effect
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association


MSMM-IV Regression Model (Reliability = .9)
Total Effect
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
True Values
True Association

Bias in $\hat{\delta}$:

MSMM-IV Model:

$$E\left[\hat{\delta}\right] = \delta + \frac{\eta - \delta}{F} + \frac{\lambda\gamma}{F\left(1 - r_\gamma\right)\tau_\gamma} + \frac{2\gamma Cov(\delta_s, \gamma_s)}{F\left(1 - r_\gamma\right)\tau_\gamma} + \frac{Cov(\lambda_s, \gamma_s)}{F\left(1 - r_\gamma\right)\tau_\gamma}$$

Bias in $\hat{\delta}$:

MSMM-IV Model:

$$E[\hat{\delta}] = \delta + \underbrace{\boxed{\frac{\eta - \delta}{F}}}_{\text{finite sample bias}} + \underbrace{\boxed{\frac{\lambda\gamma}{F(1 - r_\gamma)\tau_\gamma}}}_{\text{exclusion restriction violation bias}} + \underbrace{\boxed{\frac{2\gamma Cov(\delta_s, \gamma_s)}{F(1 - r_\gamma)\tau_\gamma}}}_{\text{compliance-effect covariance bias}} + \underbrace{\boxed{\frac{Cov(\lambda_s, \gamma_s)}{F(1 - r_\gamma)\tau_\gamma}}}_{\text{compliance-omitted pathway covariance bias}}$$

Note that $F(1 - r_\gamma)\tau_\gamma = \gamma^2 r + \tau_\gamma$, so the last three terms are small when $\gamma^2 r + \tau_\gamma$ is large and/or when the covariances are small.

Bias in $\hat{\delta}$:

MSMM-INT Model:

$$E[\hat{\delta}] = \delta + \underbrace{\boxed{(\eta - \delta)(1 - r_\gamma)}}_{\text{finite sample bias}} + \underbrace{\boxed{\frac{r_\gamma}{\tau_\gamma}\gamma \cdot Cov(\delta_s, \gamma_s)}}_{\text{compliance-effect covariance bias}} + \underbrace{\boxed{\frac{r_\gamma}{\tau_\gamma} Cov(\lambda_s, \gamma_s)}}_{\text{compliance-omitted pathway covariance bias}}$$

The final two bias terms are small when r_γ is small and/or τ_γ is large, and when the covariances are small.

Finite sample bias in $\hat{\delta}$:

MSMM-IV Model:

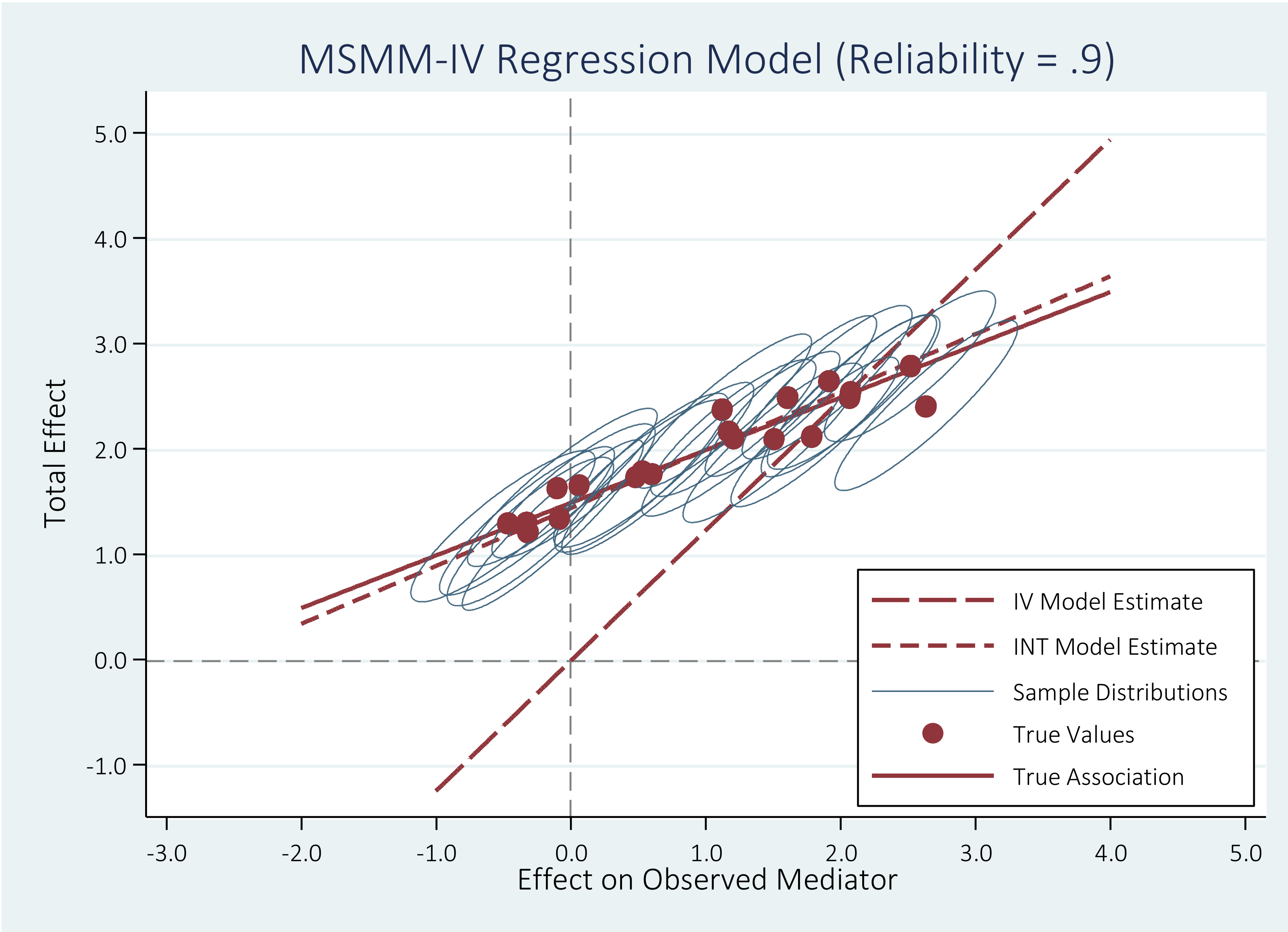
$$E[\hat{\delta}] = \delta + (\eta - \delta)\frac{1}{F}$$

MSMM-INT Model:

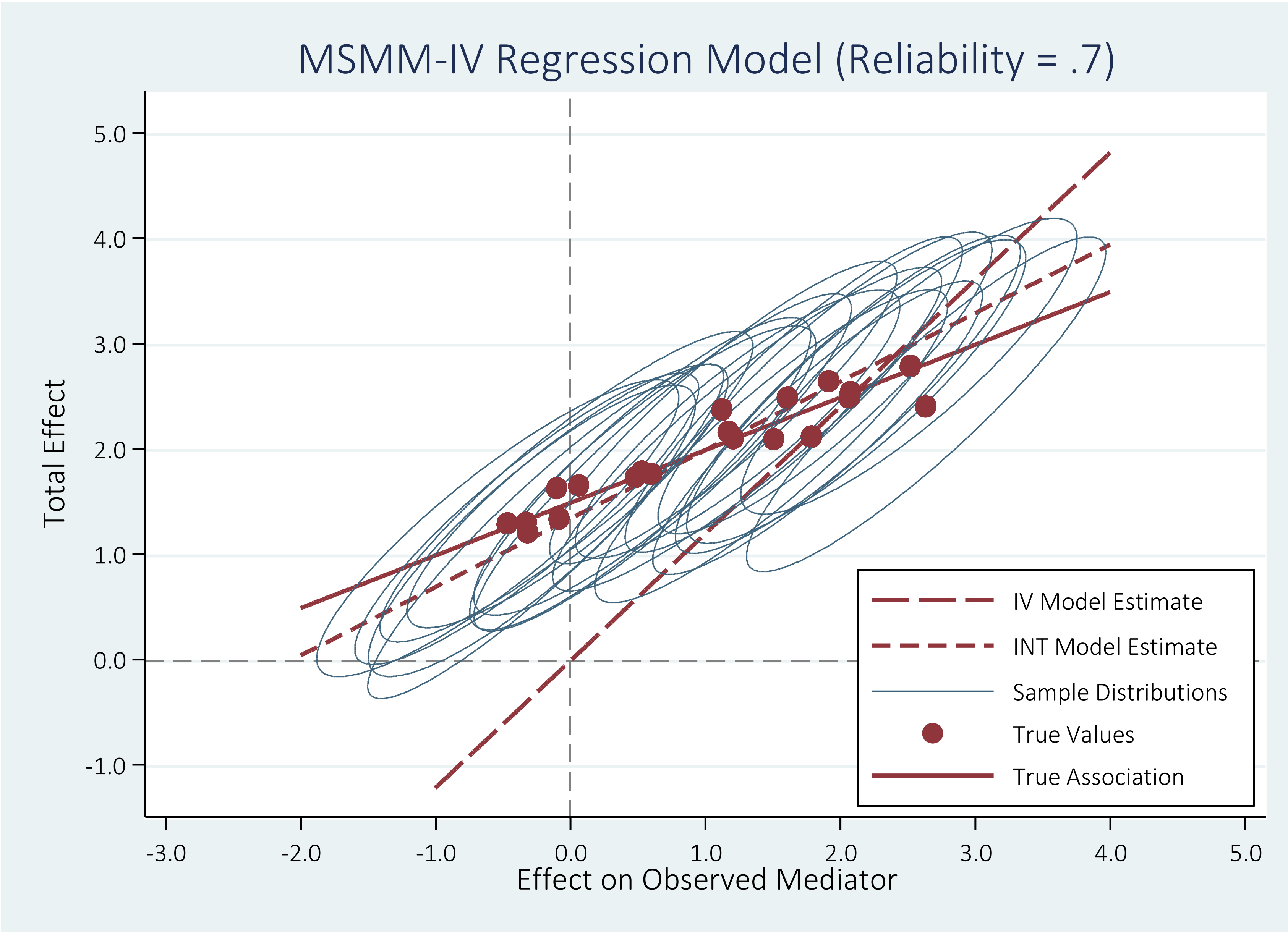
$$E[\hat{\delta}] = \delta + (\eta - \delta)(1 - r_\gamma)$$

In general, $\frac{1}{F} \leq 1 - r_\gamma$ (they are equal *iff* $\gamma = 0$), so finite sample bias is larger in model with an intercept.

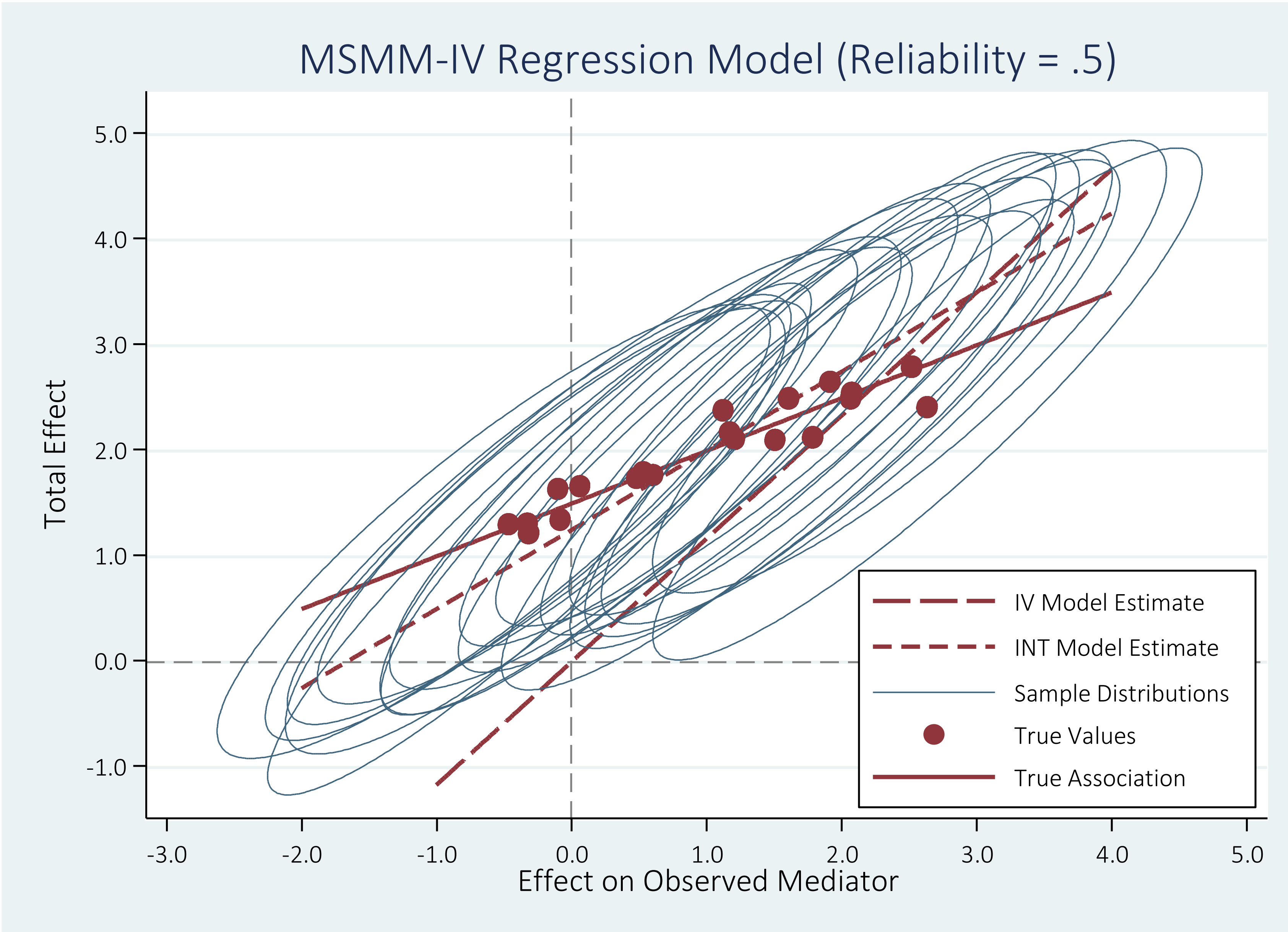
But if exclusion restriction is invalid, the no intercept (IV) model will have additional bias due to the failure of the exclusion restriction.

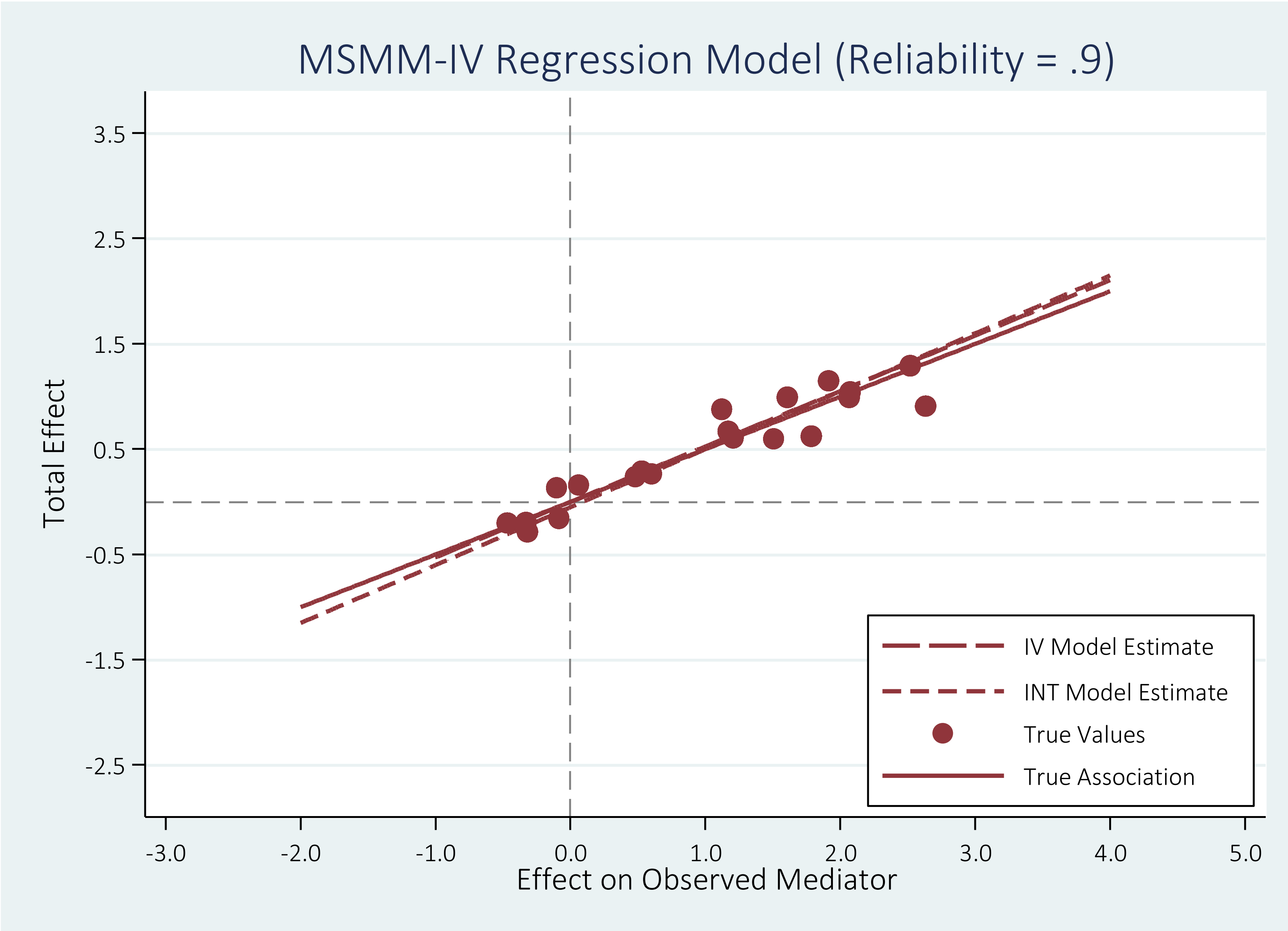


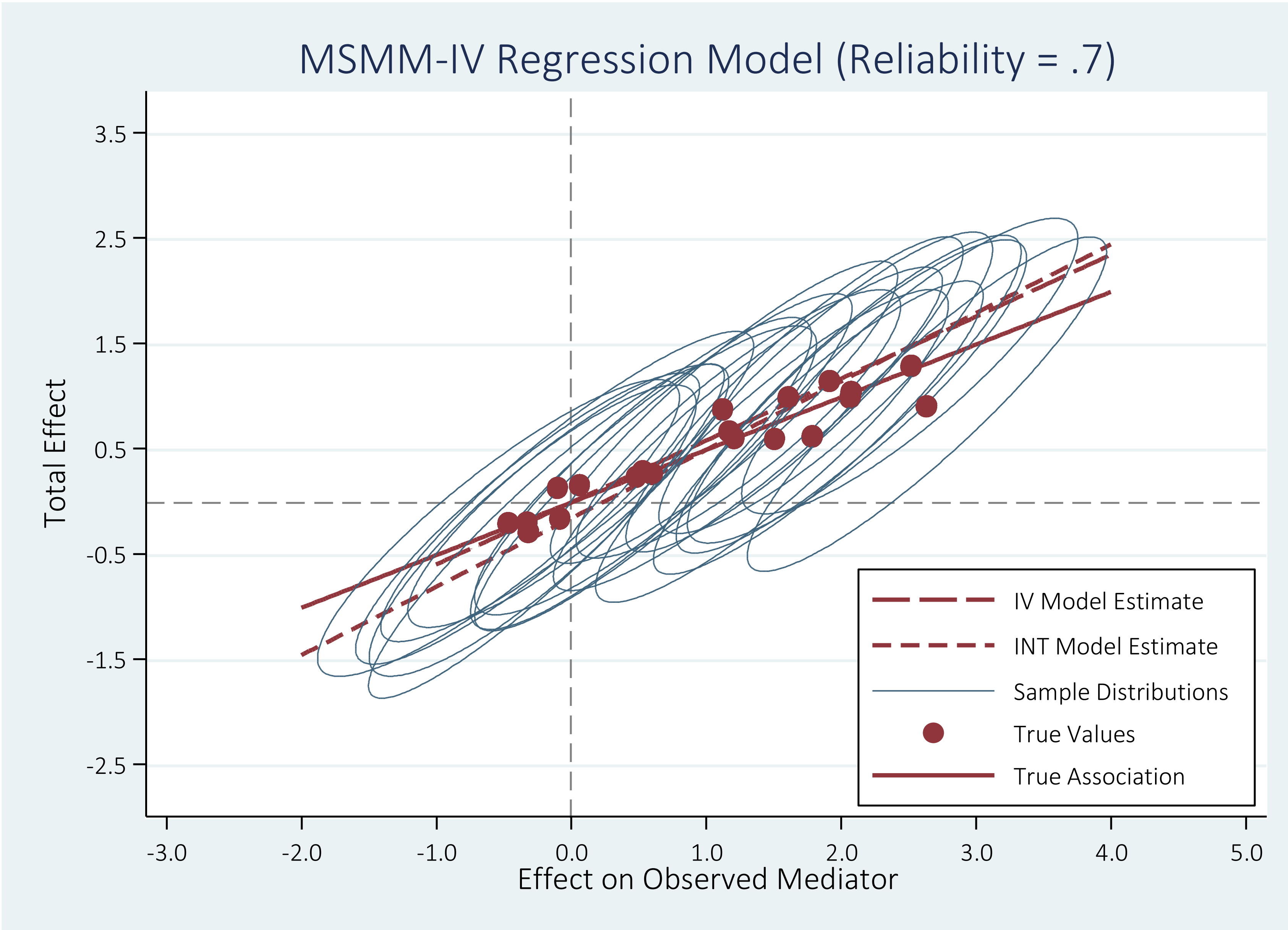
MSMM-IV Regression Model (Reliability = .9)
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
Total Effect
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association

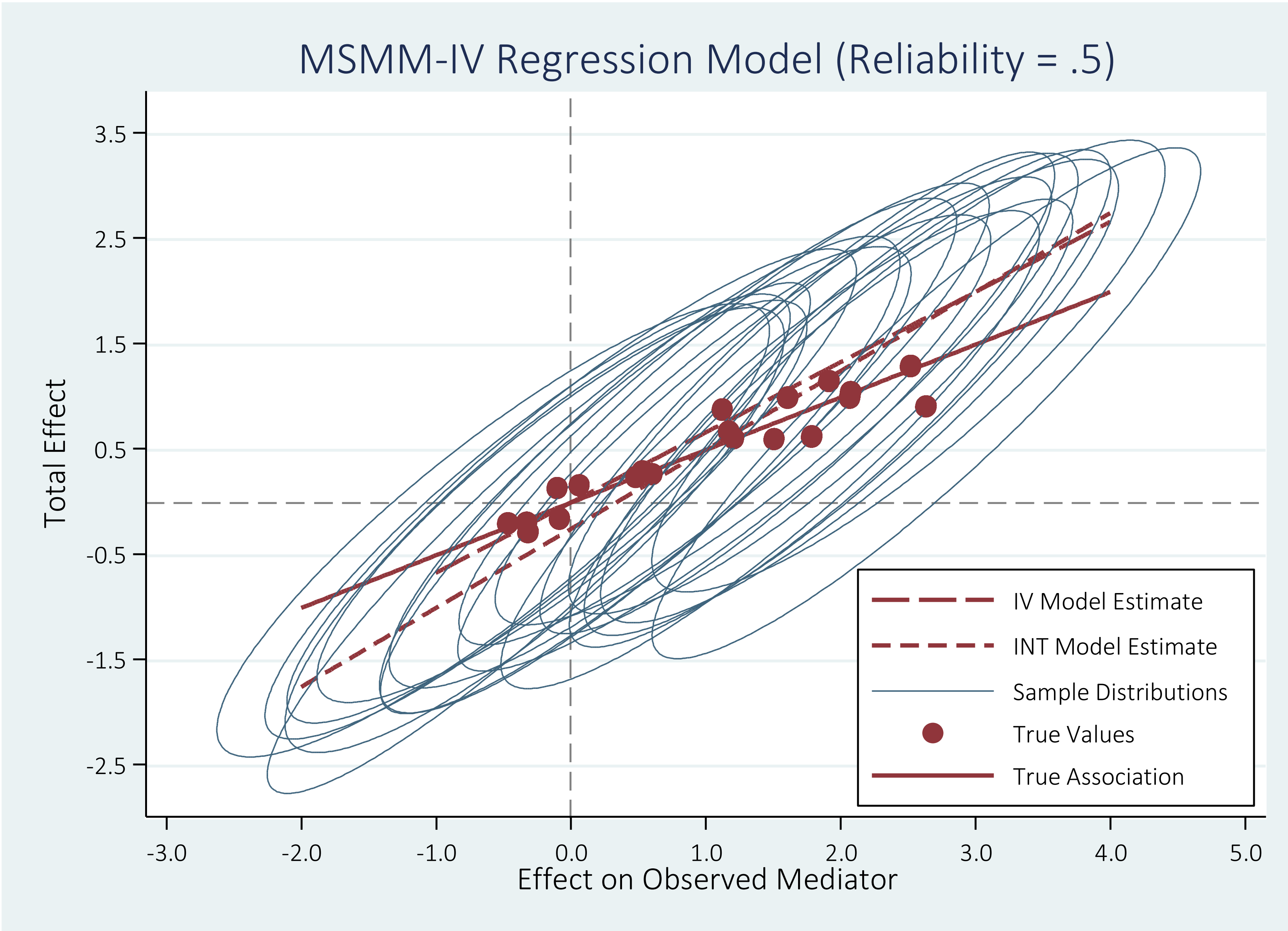


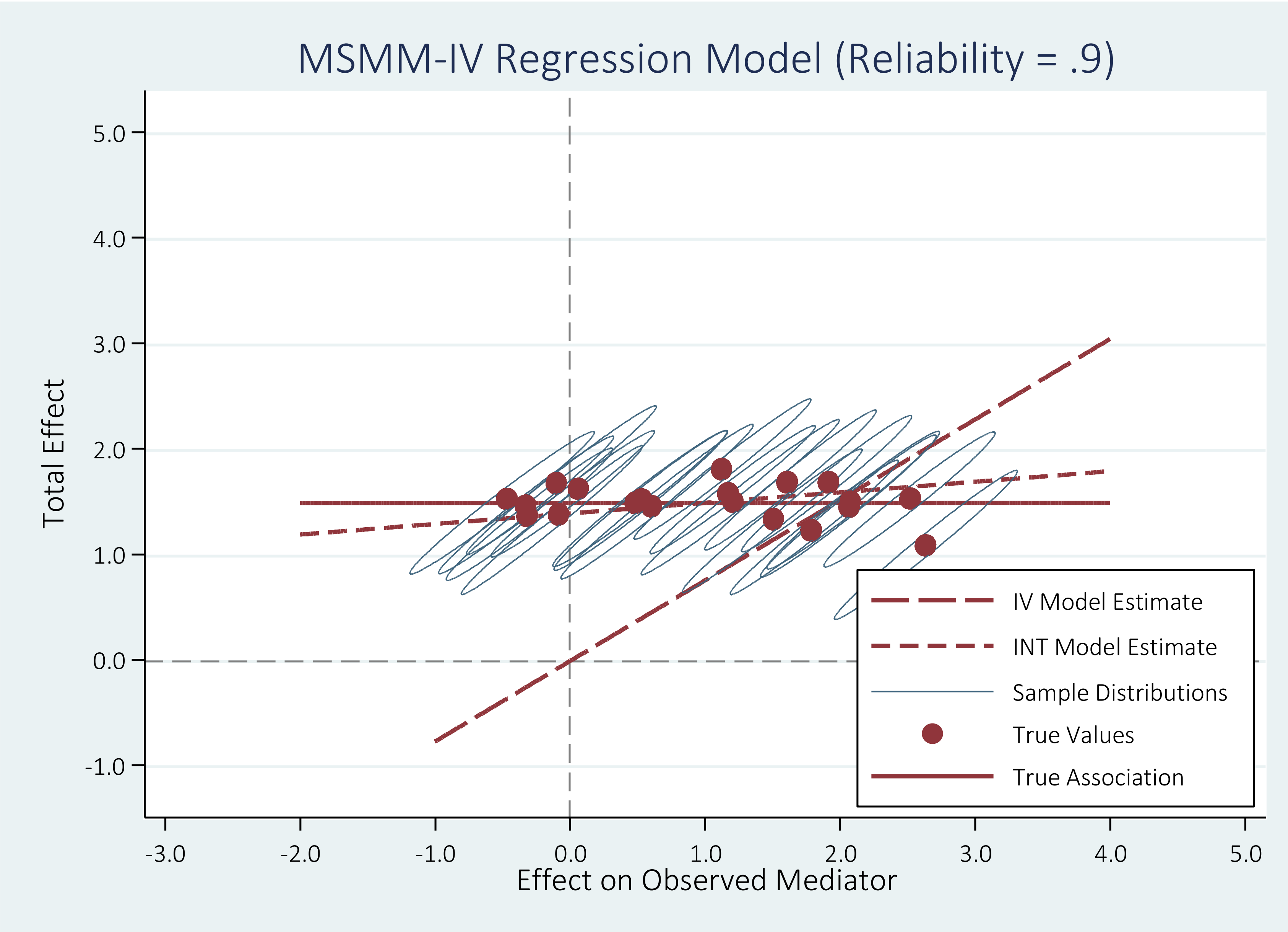
MSMM-IV Regression Model (Reliability = .7)
Total Effect
Effect on Observed Mediator
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association

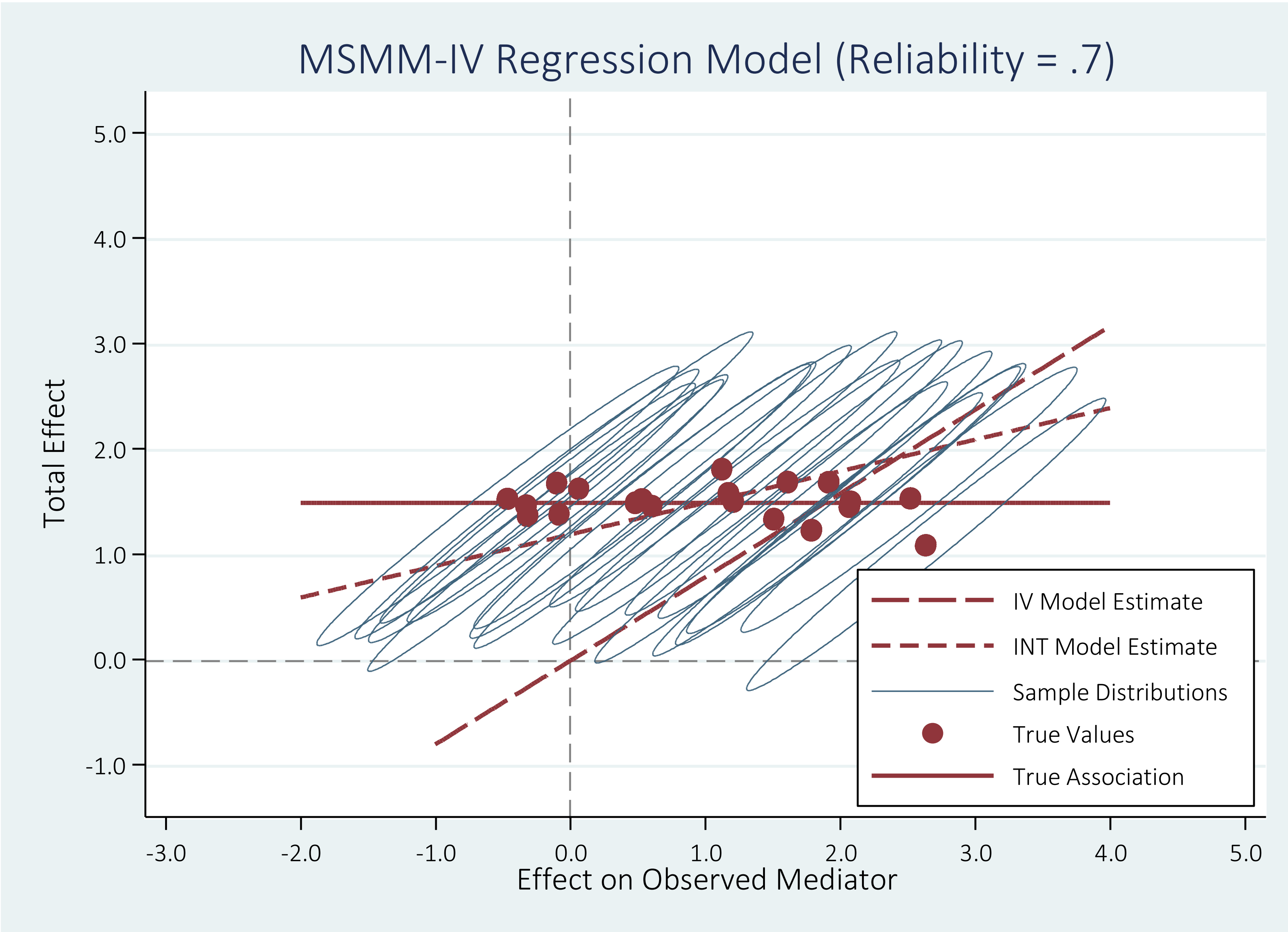

MSMM-IV Regression Model (Reliability = .5)
Total Effect
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association

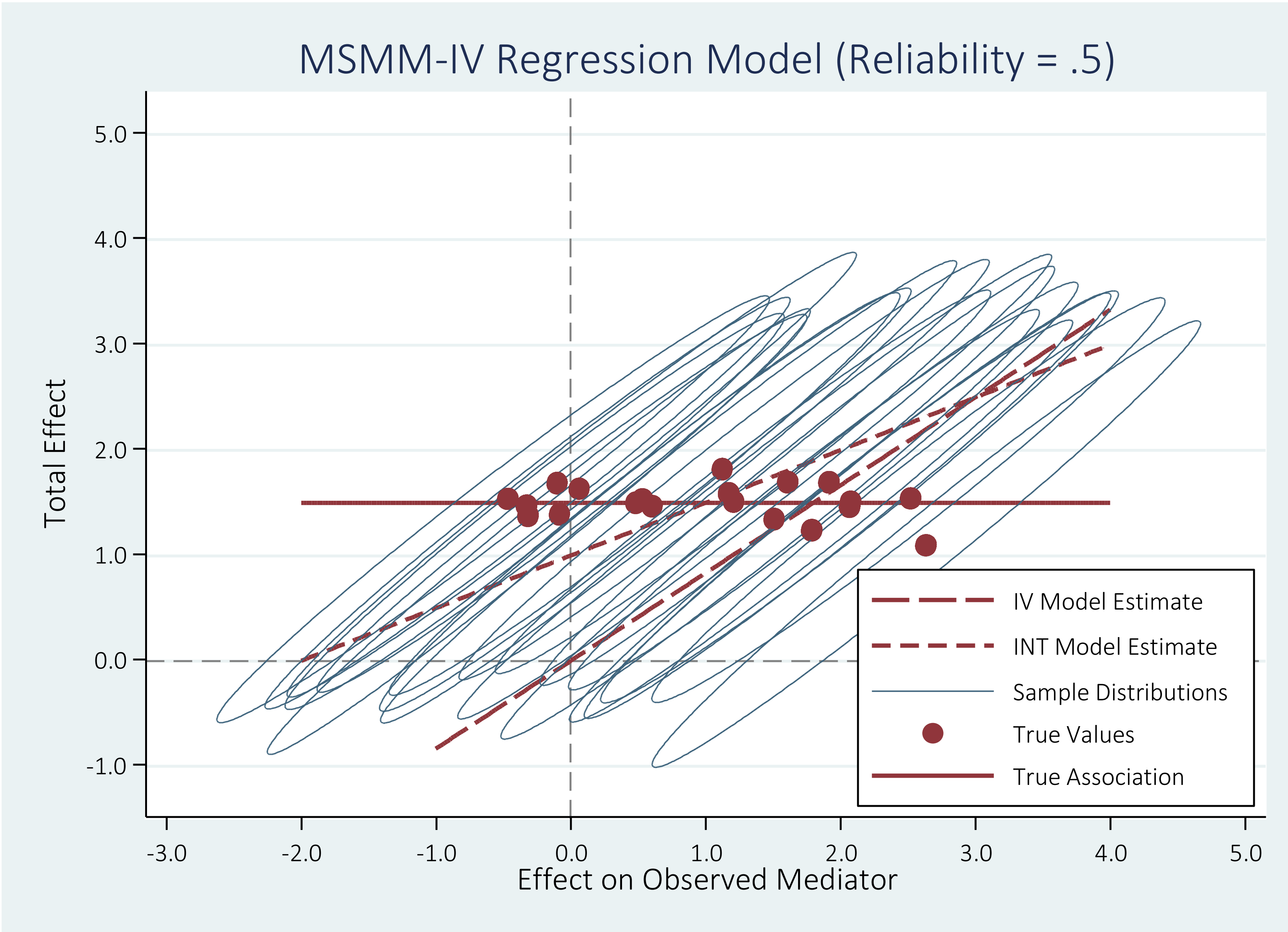

MSMM-IV Regression Model (Reliability = .9)
Total Effect
3.5
2.5
1.5
0.5
-0.5
-1.5
-2.5
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
True Values
True Association


MSMM-IV Regression Model (Reliability = .7)
Total Effect
3.5
2.5
1.5
0.5
-0.5
-1.5
-2.5
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association


MSMM-IV Regression Model (Reliability = .5)
Total Effect
3.5
2.5
1.5
0.5
-0.5
-1.5
-2.5
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association


MSMM-IV Regression Model (Reliability = .9)
Total Effect
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association


MSMM-IV Regression Model (Reliability = .7)
Total Effect
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association


MSMM-IV Regression Model (Reliability = .5)
Total Effect
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
Effect on Observed Mediator
IV Model Estimate
INT Model Estimate
Sample Distributions
True Values
True Association

Can we eliminate finite sample bias?

Recall that the expected value of the MSMM-INT estimator is

$$E\left[\hat{\delta}^{INT}\right] = \delta + (\eta - \delta)(1 - r_\gamma)$$

We can rewrite this as

$$\delta = \frac{\left[E\left[\hat{\delta}^{INT}\right] - \eta\left(1 - r_\gamma\right)\right]}{r_\gamma}$$

This suggests a “de-biased” estimator of δ:

$$\hat{\delta}^{DB} = \frac{\left[\hat{\delta}^{INT} - \hat{\eta}\left(1 - \hat{r}_\gamma\right)\right]}{\hat{r}_\gamma}$$

$$\hat{\delta}^{DB} = \frac{\left[\hat{\delta}^{INT} - \hat{\eta}\left(1 - \hat{r}_{\gamma}\right)\right]}{\hat{r}_{\gamma}}$$

We get $\hat{\delta}^{INT}$ from the MSMM-INT model;

We can estimate $\hat{r}_{\gamma}$ from the first-stage estimates;

And we can get $\hat{\eta}$ by fitting the fixed effects model (using only observations of individuals in the control condition)

$$Y_{is} = \theta_s + \eta M_{is} + e_{is}$$

or, using all observations:

$$Y_{is} = \theta_s + \eta M_{is} + \zeta_s T_{is} + e_{is}.$$

Under what conditions will this approach yield enough precision to be useful?

The sampling variance of

$$\hat{\delta}^{DB} = \frac{\left[\hat{\delta}^{INT} - \hat{\eta}\left(1 - \hat{r}_{\gamma}\right)\right]}{\hat{r}_{\gamma}}$$

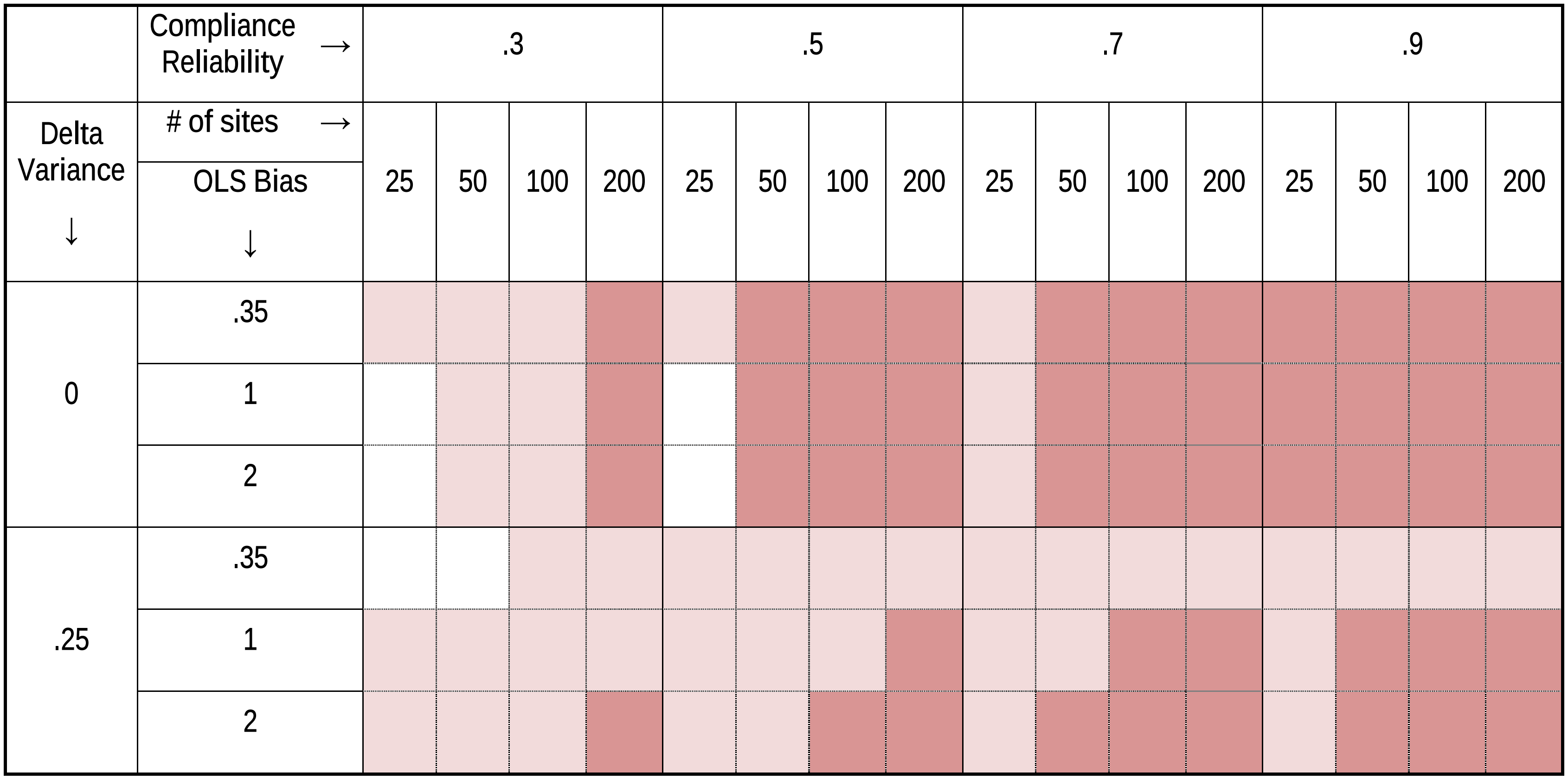
will depend on the sampling variance of $\hat{\delta}^{INT}$, $\hat{\eta}$, and $\hat{r}_{\gamma}$.

Although $\hat{\delta}^{DB}$ may be less biased than $\hat{\delta}^{INT}$, it may have greater RMSE, and so be less useful.

The reliability $\hat{r}_{\gamma}$, in particular, may be noisily estimated, particularly if there are relatively few sites and/or $\hat{r}_{\gamma}$ is small.

This will add considerable noise to $\hat{\delta}^{DB}$.

Simulation-Based Comparisons of $\widehat{\boldsymbol{\delta}}^{DB}$ and $\widehat{\boldsymbol{\delta}}^{INT}$:



Legend: $|\boldsymbol{Bias}(\boldsymbol{\delta_{db}})| \leq |\boldsymbol{Bias}(\boldsymbol{\delta_{IV}})|$ $|\boldsymbol{Bias}(\boldsymbol{\delta_{db}})| \leq |\boldsymbol{Bias}(\boldsymbol{\delta_{IV}})|$ & $\boldsymbol{RMSE}(\boldsymbol{\delta_{db}}) \leq \boldsymbol{RMSE}(\boldsymbol{\delta_{IV}})$

Notes: $\boldsymbol{\delta_{true}} = \mathbf{1}$ **and** $\boldsymbol{site\ size\ (n) = 100}$