Generalised method of moments estimation of mediation models and structural mean models

Tom Palmer

Division of Health Sciences, Warwick Medical School, University of Warwick, UK

March 2014

Outline

- Introduction to GMM
- I. Mediation models
  - Joint estimation of mediator and outcome models - delta method SE
  - Example
- II. Structural mean models
  - Multiplicative SMM
  - Logistic SMM
- Summary

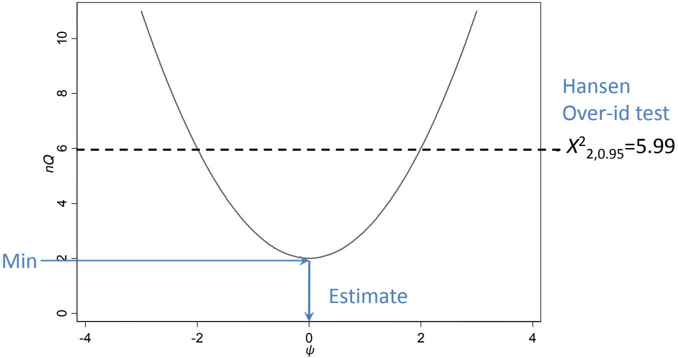
Introduction to Generalised Method of Moments (GMM) I

- Jointly solve system of moment conditions (equations)
- System: exactly identified, # instruments = # parameters
- System: over-identified, # instruments > # parameters
- m vector of moment conditions
- Minimises quadratic form w.r.t parameters (ψ)

$$Q = m' W^{-1} m = \left(\frac{1}{n} \sum_{i=1}^{n} m_i(\psi) \right)' W^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} m_i(\psi) \right)$$

Introduction to Generalised Method of Moments (GMM) II

- Profiling over parameter of interest



- Over-identification test Hansen 1982: $nQ \sim \chi^2_q$
- In quadratic form: W affects efficiency (SEs) rather than consistency

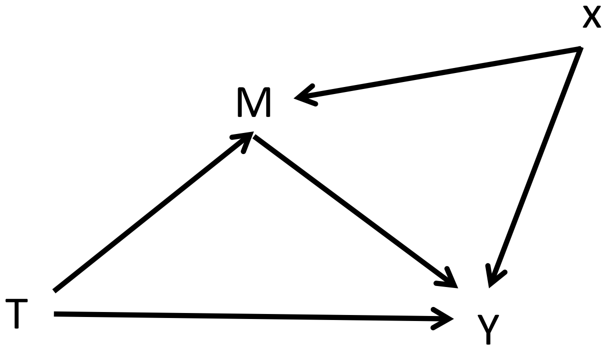
I. Mediation models

Mediation models – common implementation

Baron and Kenny (1986) type models

- Implementations:
  - Stata - `medeff`, `medsens`, `paramed`
  - R - `mediation`
  - bootstrapping for SEs of mediation parameters
- Fit mediator model
- Fit outcome model
- Mediation parameters – function of estimated parameters
- problem: fit models separately – missing covariance between parameters from different models

Mediation model with a single confounder



$$M = \alpha_0 + \alpha_1 T + \alpha_2 x + \epsilon_1$$
$$Y = \beta_0 + \beta_1 T + \beta_2 M + \beta_3 x + \epsilon_2$$

$$\widehat{\mathbf{V}} = \begin{bmatrix} \widehat{\sigma}^2_{\alpha 00} & \widehat{\sigma}_{\alpha 01} & \widehat{\sigma}_{\alpha 02} & . & . & . & . \\ \widehat{\sigma}_{\alpha 10} & \widehat{\sigma}^2_{\alpha 11} & \widehat{\sigma}_{\alpha 12} & . & . & . & . \\ \widehat{\sigma}_{\alpha 20} & \widehat{\sigma}_{\alpha 21} & \widehat{\sigma}^2_{\alpha 22} & . & . & . & . \\ . & . & . & \widehat{\sigma}^2_{\beta 00} & \widehat{\sigma}_{\beta 01} & \widehat{\sigma}_{\beta 02} & \widehat{\sigma}_{\beta 03} \\ . & . & . & \widehat{\sigma}_{\beta 10} & \widehat{\sigma}^2_{\beta 11} & \widehat{\sigma}_{\beta 12} & \widehat{\sigma}_{\beta 13} \\ . & . & . & \widehat{\sigma}_{\beta 20} & \widehat{\sigma}_{\beta 21} & \widehat{\sigma}^2_{\beta 22} & \widehat{\sigma}_{\beta 23} \\ . & . & . & \widehat{\sigma}_{\beta 30} & \widehat{\sigma}_{\beta 31} & \widehat{\sigma}_{\beta 32} & \widehat{\sigma}^2_{\beta 33} \end{bmatrix}$$

Imai et al. (2010)

$$\text{Natural indirect effect} = \beta_2\alpha_1$$

$$\text{V(NIE)} = \alpha_1^2\sigma^2_{\beta 2} + \beta_2^2\sigma^2_{\alpha 1} + 2\beta_2\alpha_1 \text{cov}(\beta_2, \alpha_1)$$

- bootstrap to estimate SE of mediation parameter
- GMM – *joint estimation of mediation and outcome models*
  - full var-covar matrix then delta-method SE

GMM estimation of GLMs

Estimating equation for GLMs - solve wrt β:

$$\sum_{i=1}^{n} x_i(y_i - g^{-1}(X\beta)) = 0$$

Model	Link fn $g(\mu)$	Inverse link $g^{-1}(X\beta)$
Linear	$I(\mu)$	$I(X\beta)$
Poisson	$\log(\mu)$	$\exp(X\beta)$
Logistic	$\text{logit}(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$	$\text{expit}(X\beta) = \frac{\exp(X\beta)}{1+\exp(X\beta)}$
Probit	$\Phi^{-1}(\mu)$	$\Phi(X\beta)$

In (exactly identified) GMM use covariates as *instruments* for themselves

Moment conditions

$$E[(M - \alpha_0 - \alpha_1 T - \alpha_2 x)1] = 0$$
$$E[(M - \alpha_0 - \alpha_1 T - \alpha_2 x)T] = 0$$
$$E[(M - \alpha_0 - \alpha_1 T - \alpha_2 x)x] = 0$$
$$E[(Y - \beta_0 - \beta_1 T - \beta_2 M - \beta_3 x)1] = 0$$
$$E[(Y - \beta_0 - \beta_1 T - \beta_2 M - \beta_3 x)T] = 0$$
$$E[(Y - \beta_0 - \beta_1 T - \beta_2 M - \beta_3 x)M] = 0$$
$$E[(Y - \beta_0 - \beta_1 T - \beta_2 M - \beta_3 x)x] = 0$$

- Stata: `gmm`/`sem` then `nlcom`
- Stata: `ml` – joint ML estimation of the 2 models
- Stata: 2 `regress` commands then `suest` then `nlcom`
- Stata: `medeff` (Hicks and Tingley)
- Stata: `paramed` (Liu and Emsley)
- R: `gmm` package then `deltamethod()` from MSM package

Example Stata code – gmm, nlcom

```
gmm (M - {a1}*T - {a2}*x - {a0}) ///
   (Y - {b1}*T - {b2}*M - {b3}*x - {b0}) ///
   , instruments(1:T x) ///
   instruments(2:T M x) ///
   winitial(unadjusted,independent)

* NIE
nlcom [b2]_cons*[a1]_cons
```

Example Stata code – suest, nlcom

```
regress M T x
estimates store medmodel

regress Y T M x
estimates store outmodel

suest medmodel outmodel

* NIE
nlcom [outmodel_mean]M*[medmodel_mean]T
```

Example Stata code – sem, nlcom

```
rename M m
rename T t
rename Y y

sem (m <- t x) ///
(y <- t m x)

// NIE
nlcom [y]m*[m]t
```

Example from `medsens` helpfile I

$$M = \alpha_0 + \alpha_1 T + \alpha_2 x + \epsilon_1$$
$$Y = \beta_0 + \beta_1 T + \beta_2 M + \beta_3 x + \epsilon_2$$

	Estimate (95% CI)
α_0	0.28 (0.19, 0.36)
α_1	0.17 (0.04, 0.29)
α_2	0.27 (0.21, 0.33)
β_0	0.32 (0.24, 0.41)
β_1	-0.58 (-0.70, -0.46)
β_2	0.71 (0.65, 0.77)
β_3	0.27 (0.21, 0.34)

Estimated variance-covariance matrix:
2 separate regressions

$$\begin{bmatrix} 0.00178 & & & & & & \\ -0.00178 & 0.00398 & & & & & \\ -0.00001 & 0.00008 & 0.0009 & & & & \\ & & & 0.00193 & & & \\ & & & -0.00181 & 0.0037 & & \\ & & & -0.0003 & -0.00007 & 0.0009 & \\ & & & 0.0002 & -0.00009 & -0.0002 & 0.0010 \end{bmatrix}$$

Estimated variance-covariance matrix:
with covariance from GMM

$$\begin{bmatrix} 0.00178 & & & & & & \\ -0.00178 & 0.00398 & & & & & \\ -0.00001 & 0.00008 & 0.0009 & & & & \\ 0.00009 & -\ 0.00009 & 0.00005 & 0.00193 & & & \\ -\ 0.00007 & -\ 0.00003 & -\ 0.000002 & -0.00181 & 0.0037 & & \\ -\ 0.00005 & \color{red}{0.00007} & -\ 0.00002 & -0.0003 & -0.00007 & 0.0009 & \\ 0.00006 & -\ 0.00004 & -\ 0.00003 & 0.0002 & -0.00009 & -0.0002 & 0.0010 \end{bmatrix}$$

Example from `medsens` helpfile II

Estimates of mediation parameters

	Estimate	95% CI	
		Bootstrap	DM (using GMM)
NIE $= \beta_2\alpha_1$	0.12	(0.035, 0.211)	(0.031, 0.209)
CDE $= \beta_1$	-0.58	(-0.697, -0.451)	(-0.697, -0.457)
Total effect	-0.46	(-0.610, -0.304)	(-0.605, -0.309)
% TE mediated	-0.26	(-0.397, -0.198)	(-0.517, -0.008)

DM: delta-method

SEs from joint maximum likelihood estimation

$$\text{logLike} = \text{logLike}_{\text{Mediator}} + \text{logLike}_{\text{Outcome}}$$

	Estimate	95% CI		
		Bootstrap	DM (GMM)	DM (ML)
NIE $= \beta_2\alpha_1$	0.12	(0.035, 0.211)	(0.031, 0.209)	(0.031, 0.208)
CDE $= \beta_1$	-0.58	(-0.697, -0.451)	(-0.697, -0.457)	(-0.697, -0.457)
Total effect	-0.46	(-0.610, -0.304)	(-0.605, -0.309)	(-0.605, -0.308)
% TE mediated	-0.26	(-0.397, -0.198)	(-0.517, -0.008)	(-0.515, -0.009)

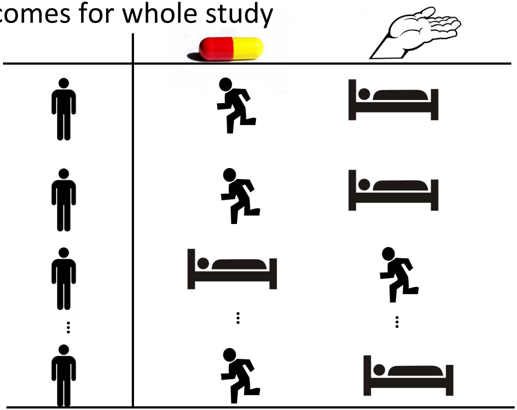
DM: delta-method

GMM SEs here are heteroskedasticity robust SEs

II. Structural mean models

Structural mean models

Potential outcomes for whole study



Average treatment effect = $E[Y(X=1)] - E[Y(X=0)]$

binary outcome: causal risk difference

Causal risk ratio = $E[Y(X=1)] \; / \; E[Y(X=0)]$

Causal odds ratio = $\text{odds}[Y(X=1)] \; / \; \text{odds}[Y(X=0)]$

Multiplicative SMM

- X exposure/treatment
- Y outcome
- Z instrument
- $Y(X = 0)$ exposure/treatment free potential outcome

Hernan & Robins 2006

$$\frac{E[Y|X,Z]}{E[Y(0)|X,Z]} = \exp(\psi X)$$

ψ : log causal risk ratio

Rearrange for $Y(0)$: $\quad Y(0) = Y \exp(-\psi X)$

Mutliplicative SMM: estimation with multiple instruments

Under the instrumental variable assumptions Robins 1994:

$$Y(0) \perp\!\!\!\perp Z$$
$$Y \exp(-\psi X) \perp\!\!\!\perp Z$$

Mutliplicative SMM: estimation with multiple instruments

Under the instrumental variable assumptions Robins 1994:

$$Y(0) \perp\!\!\!\perp Z$$
$$Y \exp(-\psi X) \perp\!\!\!\perp Z$$
$$\text{trick: } Y \exp(-\psi X) - Y(0) \perp\!\!\!\perp Z$$

Mutliplicative SMM: estimation with multiple instruments

Under the instrumental variable assumptions Robins 1994:

$$Y(0) \perp\!\!\!\perp Z$$
$$Y \exp(-\psi X) \perp\!\!\!\perp Z$$
$$\text{trick: } Y \exp(-\psi X) - Y(0) \perp\!\!\!\perp Z$$

Moment conditions

Z=0,1

$$E[(Y \exp(-\psi X) - Y(0))1] = 0$$
$$E[(Y \exp(-\psi X) - Y(0))Z_1] = 0$$

Mutliplicative SMM: estimation with multiple instruments

Under the instrumental variable assumptions Robins 1994:

$$Y(0) \perp\!\!\!\perp Z$$
$$Y \exp(-\psi X) \perp\!\!\!\perp Z$$
$$\text{trick: } Y \exp(-\psi X) - Y(0) \perp\!\!\!\perp Z$$

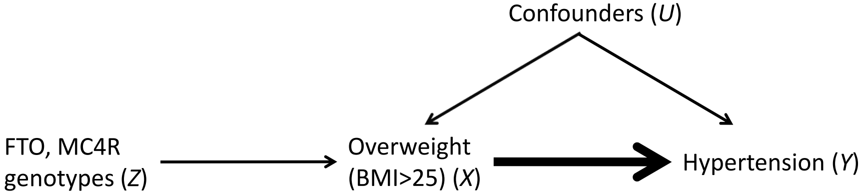
Moment conditions

Z=0,1,2,3

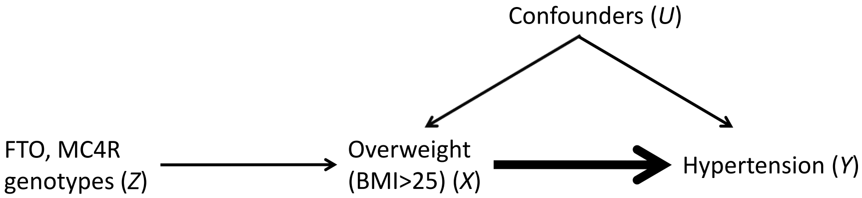
Over-identified

$$E[(Y \exp(-\psi X) - Y(0))1] = 0$$
$$E[(Y \exp(-\psi X) - Y(0))Z_1] = 0$$
$$E[(Y \exp(-\psi X) - Y(0))Z_2] = 0$$
$$E[(Y \exp(-\psi X) - Y(0))Z_3] = 0$$

Copenhagen example descriptive statistics 1



Copenhagen example descriptive statistics 1



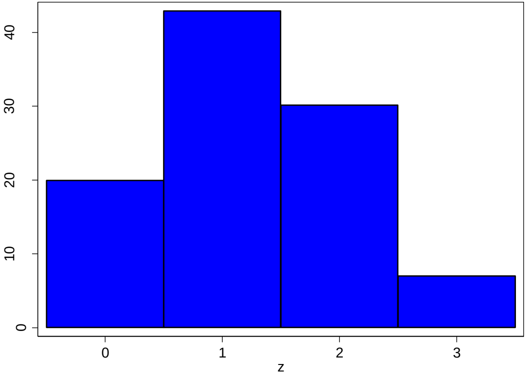
	No Hypertension	Hypertension	Total
Not Overweight	10,066 42%	13,909 58%	23,975
Overweight	6,906 22%	24,642 78%	31,548
Total	16,972 31%	38,551 69%	55,523 χ^2 P<0.001

Risk ratio for hypertension 1.35 (1.32, 1.37)

Copenhagen example descriptive statistics 2

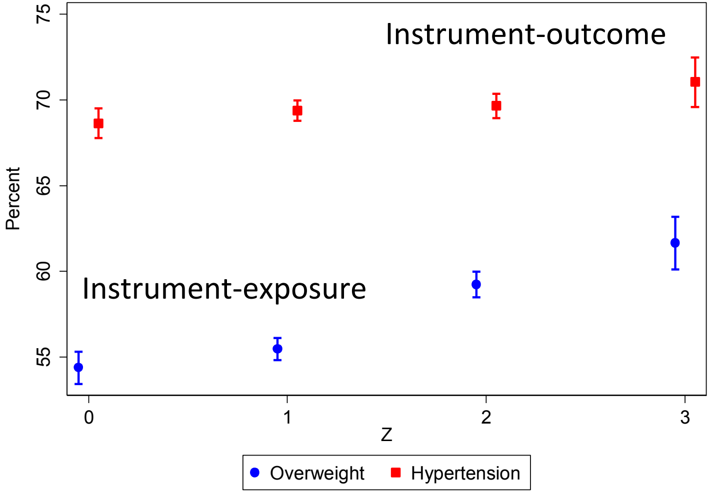
Distribution of instrument (Z)

FTO	*MC4R*	*Z*	Freq
0	0	0	0.20
0	1	1	0.15
1	0	1	0.27
1	1	2	0.21
2	0	2	0.09
2	1	3	0.07



Copenhagen example descriptive statistics 3

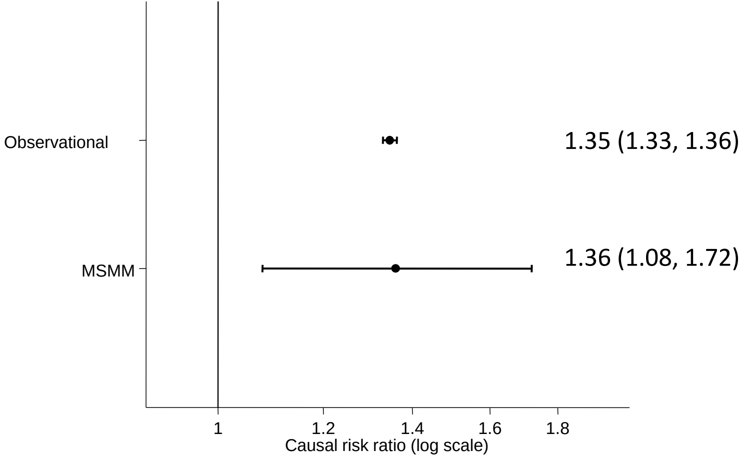
Exposure (over-weight) & outcome (hypertension) by instrument



P=0.007

P<0.001
R^2=0.002

Copenhagen example Multiplicative SMM estimates



MSMM: Hansen over-identification test $P = 0.31$
$E[Y(0)] = 0.58$ (0.50, 0.65)

How does GMM deal with multiple instruments?

GMM estimator solution to:

$$\frac{\partial m'(\psi)}{\partial \psi} W^{-1} m(\psi) = 0$$

- MSMM: instruments combined into linear projection of $YX \exp(-X\psi)$ on $Z = (1, Z_1, Z_2)'$ Bowden & Vansteelandt 2010

(double) Logistic SMM

$\text{logit}(p) = \log(p/(1-p)), \text{expit}(x) = e^x/(1+e^x)$

Goetghebeur, 2010

$$\text{logit}(E[Y|X,Z]) - \text{logit}(E[Y(0)|X,Z]) = \psi X$$

ψ : log causal odds ratio

Rearrange for $Y(0)$: $Y(0) = \text{expit}(\text{logit}(Y) - \psi X)$

(double) Logistic SMM

$\text{logit}(p) = \log(p/(1-p)), \text{expit}(x) = e^x/(1+e^x)$

Goetghebeur, 2010

$$\text{logit}(E[Y|X,Z]) - \text{logit}(E[Y(0)|X,Z]) = \psi X$$

$$\psi : \text{ log causal odds ratio}$$

$$\text{Rearrange for } Y(0): \quad Y(0) = \text{expit}(\text{logit}(Y) - \psi X)$$

- Can't be estimated in a single step Robins (1999)
- First stage association model Vansteelandt (2003):
  (i) logistic regression of Y on X & Z & interactions
  (ii) predict Y, estimate LSMM using predicted Y

(double) Logistic SMM moment conditions

Association model moment conditions

Logistic regression using GMM

$$E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))1] = 0$$
$$E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))X] = 0$$
$$E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))Z] = 0$$
$$E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))XZ] = 0$$

(double) Logistic SMM moment conditions

Association model moment conditions

Logistic regression using GMM

$$E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))1] = 0$$
$$E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))X] = 0$$
$$E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))Z] = 0$$
$$E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))XZ] = 0$$

Causal model moment conditions

$$E[(\text{expit}(\text{logit}(\widehat{p}) - \psi X) - Y(0))1] = 0$$
$$E[(\text{expit}(\text{logit}(\widehat{p}) - \psi X) - Y(0))Z] = 0$$

Problem: SEs incorrect - need association model uncertainty

LSMM joint estimation

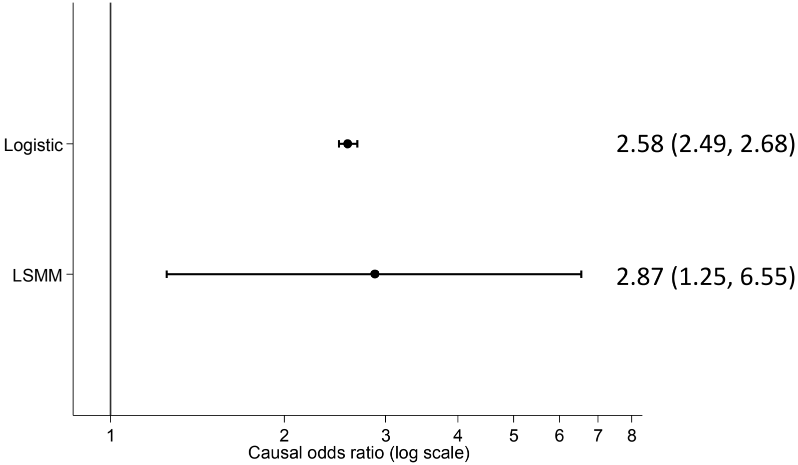
Joint estimation = correct SEs Gourieroux (1996)

Vansteelandt & Goetghebeur (2003)

$$
\begin{aligned}
E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))1] &= 0 \\
E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))X] &= 0 \\
E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))Z] &= 0 \\
E[(Y - \text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ))XZ] &= 0 \\
E[(\text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ - \psi X) - Y(0))1] &= 0 \\
E[(\text{expit}(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ - \psi X) - Y(0))Z] &= 0
\end{aligned}
$$

In example causal model SEs increase $\times 10$ from non-joint estimation

Copenhagen example LSMM estimates



LSMM: Hansen over-identification test $P = 0.29$
$\mathrm{E}[Y(0)] = 0.57$ (0.45, 0.68)

Issues estimating SMMs

- Weak identification
  - many values of causal parameter give independence condition close to zero
- GMM convergence at local/global minima
- Hence check estimated $E[Y(0)]$ approx baseline risk
- Sensitive to initial values: in another dataset
  - initial $CRR = 1$ gave $CRR > 1$
  - initial $CRR < 1$ gave $CRR < 1$
- Fit with centred Z (with/without constant E[Y(0)])
- Estimation MSMM/LSMM models with continuous X more problematic than binary X – centring X important for sensible estimates of E[Y(0)]

Summary

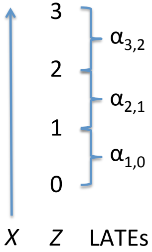
- Mediation models
  - GMM exact identification
  - Delta-method SEs alternative to bootstrapping
- Structural mean models
  - fit over-identified models with multiple instruments
  - check if estimated $E[Y(0)]$ is sensible – approx baseline risk
- Straightforward to implement in Stata & R

References

- Baron RM, Kenny DA (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. Journal of Personality and Social Psychology, 51(6), 1173–1182.
- Hansen. Large sample properties of generalized method of moments estimators Econometrica, 1982, 50, 1029-1054.
- Hicks R, Tingley D (2011) Causal mediation analysis. The Stata Journal, 11(4), 605–619.
- Imai K, Keele L, Yamamoto T (2010) Identification, Inference, and Sensitivity Analysis for Causal Mediation Effects, Statistical Sciences, 25(1) pp. 51-71.
- Emsley RA, Liu H (2013) paramed: Stata module to perform causal mediation analysis using parametric regression models. `http://ideas.repec.org/c/boc/bocode/s457581.html`
- Robins JM (1994) Correcting for non-compliance in randomized trials using structural nested mean models. CSTM. 23(8) 2379–2412

Local risk ratios for Multiplicative SMM

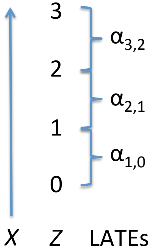
- Identification: NEM by Z ... what if it doesn't hold?
- Alternative assumption of monotonicity: $X(Z_k) \geq X(Z_{k-1})$
- Local Average Treatment Effect (LATE) Imbens 1994
  - effect among those whose exposures are changed (upwardly) by changing (counterfactually) the IV from Z_{k-1} to Z_k



$$\alpha_{All} = \lambda_1 \alpha_{1,0} + \lambda_2 \alpha_{2,1} + \lambda_3 \alpha_{3,2}$$

Local risk ratios for Multiplicative SMM

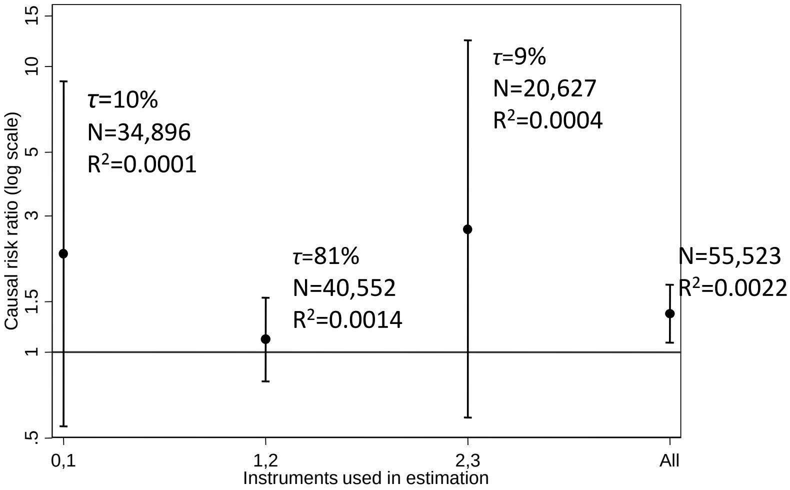
- Identification: NEM by Z ... what if it doesn't hold?
- Alternative assumption of monotonicity: $X(Z_k) \geq X(Z_{k-1})$
- Local Average Treatment Effect (LATE) Imbens 1994
  - effect among those whose exposures are changed (upwardly) by changing (counterfactually) the IV from Z_{k-1} to Z_k



$$\alpha_{All} = \lambda_1 \alpha_{1,0} + \lambda_2 \alpha_{2,1} + \lambda_3 \alpha_{3,2}$$

Similar result holds for MSMM: $e^{\psi}_{\text{All}} = \sum_{k=1}^{K} \tau_k e^{\psi}_{k,k-1}$

Copenhagen example local risk ratios



Check: $(0.10 \times 2.21) + (0.81 \times 1.11) + (0.09 \times 2.69) = 1.36$

R code I – gmm & msm packages

```
library(msm); library(gmm); library(foreign)
data <- read.dta("contmediatorcontoutcome.dta")

ex1Moments <- function(theta, x){
 M <- data[,"M"]; T <- data[,"T"]
 x <- data[,"x"]; Y <- data[,"Y"]
 # moments
 m1 <- (M - theta[1] - theta[2]*T - theta[3]*x)
 m2 <- (M - theta[1] - theta[2]*T - theta[3]*x)*T
 m3 <- (M - theta[1] - theta[2]*T - theta[3]*x)*x
 m4 <- (Y - theta[4] - theta[5]*T - theta[6]*M - theta[7]*x)
 m5 <- (Y - theta[4] - theta[5]*T - theta[6]*M - theta[7]*x)*T
 m6 <- (Y - theta[4] - theta[5]*T - theta[6]*M - theta[7]*x)*M
 m7 <- (Y - theta[4] - theta[5]*T - theta[6]*M - theta[7]*x)*x
 return(cbind(m1,m2,m3,m4,m5,m6,m7))
}

ex1 <- gmm(ex1Moments, x=data, t0=c(0,0,0,0,0,0,0), vcov="iid")
```

R code II – gmm & msm packages

```
print(summary(ex1)); print(cbind(coef(ex1), confint(ex1)))
estmean <- coef(ex1)
estvar <- ex1$vcov

# ACME
acme <- estmean[6]*estmean[2]
acmese <- deltamethod( ~ (x6*x2), estmean, estvar)
acmeci <- c(acme - 1.96*acmese, acme + 1.96*acmese)
print(acme); print(acmeci)
```