Estimating Structural Mean Models with Multiple Instruments using Generalised Method of Moments

Paul Clarke Tom Palmer Frank Windmeijer

University of Bristol

Bristol, RSS, 10/05/11

- Additive, Multiplicative and Logistic Structural Mean Models.
- Reformulate moment conditions in such a way that Generalised Method of Moments (GMM) estimation techniques can be applied.
- Extension to multivalued/multiple instruments is then straightforward, again using GMM to estimate the causal parameters, throughout making the no effect modification (NEM) assumption.
- One-step GMM projection of multiple instruments is same as proposal by Bowden and Vansteelandt (2010)
- Provide code for estimation of the SMMs by GMM in R and Stata.
- Extend multiple instruments LATE result of Imbens and Angrist (1994) to multiplicative Local Risk Ratio.

- We apply the GMM estimation procedures to estimate the causal effect of adiposity on hypertension as in Timpson et al. (2010), using genetic markers as instruments for adiposity. The data are from the Copenhagen General Population Study.

Y, X and Z are binary. Z is instrumental variable; X is exposure, and Y is outcome.

The multiplicative SMM is

$$\frac{E\left[Y|X,Z\right]}{E\left[Y\left(0\right)|X,Z\right]} = \exp\left(\theta_0 + \theta_1 Z\right) X,$$

where $Y\left(0\right)$ is the exposure- or treatment-free potential outcome.

Assuming NEM, $\theta_1 = 0$, and exploiting the conditional mean independence (CMI), assumption

$$E\left[Y\left(0\right)|Z=1\right] = E\left[Y\left(0\right)|Z=0\right] = E\left[Y\left(0\right)\right],$$

it follows that

$$\begin{aligned} E\left[\left\{Y\exp\left(-X\theta_0\right) - \alpha_0\right\}|Z=1\right] &= 0; \\ E\left[\left\{Y\exp\left(-X\theta_0\right) - \alpha_0\right\}|Z=0\right] &= 0, \end{aligned}$$

where $\alpha_0 = E\left[Y\left(0\right)\right]$.

For a multivalued instrument Z with values 0, 1, 2, From NEM and CMI we get the moment conditions

$$\begin{aligned} E\left[\left\{Y\exp\left(-X\theta_0\right)-\alpha_0\right\}|Z=2\right] &= 0; \\ E\left[\left\{Y\exp\left(-X\theta_0\right)-\alpha_0\right\}|Z=1\right] &= 0; \\ E\left[Y\exp\left(-X\theta_0\right)-\alpha_0\right] &= 0, \end{aligned}$$

with $\alpha_0 = E\left[Y(0)\right]$.

If we denote the indicator variables $Z_j = 1\{Z = j\}$ and let $S = (1, Z_1, Z_2)'$, then

$$E\left[\left\{\frac{Y}{\exp(X\theta_0)} - \alpha_0\right\}|S\right] = 0,$$

or

$$E\left[\frac{Y - \exp(\alpha_0^* + X\theta_0)}{\exp(X\theta_0)}|S\right] = 0,$$

where $\alpha_0^* = \ln(E[Y_0])$. But, by dividing by the constant α_0, it then also follows that

$$E\left[\frac{Y - \exp(\alpha_0^* + X\theta_0)}{\exp(\alpha_0^* + X\theta_0)}|S\right] = 0.$$

Let $g_i(\delta_0) = s_i\left(\frac{y_i}{\exp(x_i\theta_0)} - \alpha_0\right)$, $\delta_0 = (\alpha_0, \theta_0)'$, then $E[g_i(\delta_0)] = 0$. The GMM estimator $\widehat{\delta}$ is the solution to

$$\widehat{\delta} = \arg\min_{\delta}\left(\frac{1}{n}\sum_{i=1}^{n} g_i(\delta)\right)' W_n^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} g_i(\delta)\right).$$

A one-step GMM estimator, $\widehat{\delta}_1$, is obtained by choosing an initial weight matrix, e.g. $W_n = \frac{1}{n}\sum_i s_i s_i'$. The efficient two-step GMM estimator is obtained as

$$\widehat{\delta}_2 = \arg\min_{\delta}\left(\frac{1}{n}\sum_{i=1}^{n} g_i(\delta)\right)' W_n^{-1}\left(\widehat{\delta}_1\right)\left(\frac{1}{n}\sum_{i=1}^{n} g_i(\delta)\right)$$

where

$$W_n\left(\widehat{\delta}_1\right) = \frac{1}{n}\sum_{i=1}^{n} g_i\left(\widehat{\delta}_1\right) g_i\left(\widehat{\delta}_1\right)'.$$

Under standard regularity conditions the limiting distributions of the one- and two-step GMM estimators are

$$\sqrt{n}\left(\widehat{\delta}_1 - \delta_0\right) \xrightarrow{d} N\left(0, \left(C_0' W C_0\right)^{-1} C_0 W \Omega_0 W C_0 \left(C_0' W C_0\right)^{-1}\right)$$
$$\sqrt{n}\left(\widehat{\delta}_2 - \delta_0\right) \xrightarrow{d} N\left(0, \left(C_0' \Omega_0 C_0\right)^{-1}\right)$$

where

$$C_0 = E\left[\frac{\partial g_i(\delta)}{\partial \delta'}|_{\delta_0}\right];$$
$$\Omega_0 = E\left[g_i(\delta_0)\, g_i(\delta_0)'\right];$$
$$W = \text{plim}\,(W_n)$$

e.g. $W = E\left[s_i s_i'\right]$ when $W_n = \frac{1}{n}\sum_i s_i s_i'$.

As the instrument is discrete, no further efficiency gains can be made and the 2-step GMM estimator is asymptotically efficient, see Chamberlain (1987).

As we now have more moment conditions than unknown parameters in the model we get that when the moment conditions are valid

$$J\left(\widehat{\delta}_2\right) = n\left(\frac{1}{n}\sum_{i=1}^{n} g_i\left(\widehat{\delta}_2\right)\right)' W_n^{-1}\left(\widehat{\delta}_1\right)\left(\frac{1}{n}\sum_{i=1}^{n} g_i\left(\widehat{\delta}_2\right)\right) \xrightarrow{d} \chi_q^2,$$

with q the degree of overidentification, in this example $q = 1$. $J\left(\widehat{\delta}_2\right)$ is a test for the validity of the model assumptions and is known as Hansen's J-test, Hansen (1982).

Considering the one-step GMM estimator with $W_n = \frac{1}{n}\sum_i s_i s_i'$.
The first-order condition, using $g_i(\delta) = s_i\left(\frac{y_i}{\exp(x_i\theta)} - \alpha\right)$:

$$\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial g_i(\delta)}{\partial\delta}\right)' W_n^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} g_i(\delta)\right) = 0$$

or

$$D'S\left(S'S\right)^{-1}S'v = 0$$

$$\begin{aligned} D &= \left\{d_i'\right\};\; S = \left\{s_i'\right\};\; v = \left\{v_i\right\} \\ d_i &= \left(\begin{array}{c} 1 \\ \frac{y_i}{\exp(x_i\theta)}x_i \end{array}\right); v_i = \frac{y_i}{\exp(x_i\theta)} - \alpha \end{aligned}$$

The instruments get combined into the projection $S\left(S'S\right)^{-1}S'D$, i.e. a constant 1 and the linear projection of $\frac{y_i}{\exp(x_i\theta)}x_i$ on s_i, the projection as proposed by Bowden and Vansteelandt (2010).

For the binary case considered here,

$$E\left[\frac{YX}{\exp(X\theta)}|S\right] = \frac{1}{\exp(\theta)}E\left[YX|S\right],$$

and this is therefore equivalent to using the linear projection of $y_i x_i$ on s_i as an instrument. It is also clear from this that the same one-step GMM estimate of θ is obtained by specifying the moment conditions as

$$\begin{aligned} g_i(\theta) &= \widetilde{s}_i\left(\frac{y_i}{\exp(x_i\theta)}\right); \\ \widetilde{s}_i &= \left(z_{i1} - \overline{z}_1, z_{i2} - \overline{z}_2\right)', \end{aligned}$$

with $\overline{z}_j$ the sample average of z_{ij}. This transformation is generally used in G-estimation, see e.g. Vansteelandt and Goetgebheur (2003).

Considering the logistic SMM, the moment conditions are derived from

$$\text{logit}\left\{E\left[Y|X,Z\right]\right\} - \text{logit}\left\{E\left[Y\left(0\right)|X,Z\right]\right\} = \xi_0 X.$$

For Z taking the values $0, 1, 2$, this results in,

$$\begin{aligned} E\left[\left\{\text{expit}\left(\text{logit}\left(E\left[Y|X,Z\right]\right) - X\xi_0\right) - \alpha_0\right\}|Z=2\right] &= 0; \\ E\left[\left\{\text{expit}\left(\text{logit}\left(E\left[Y|X,Z\right]\right) - X\xi_0\right) - \alpha_0\right\}|Z=1\right] &= 0; \\ E\left[\text{expit}\left(\text{logit}\left(E\left[Y|X,Z\right]\right) - X\xi_0\right) - \alpha_0\right] &= 0, \end{aligned}$$

where, again, $\alpha_0 = E\left[Y\left(0\right)\right]$.

Following Vansteelandt at al. (2010), let the saturated model be

$$\begin{aligned}\text{logit}\{E[Y|X,Z]\} &= \text{logit}\{P(Y=1|X,Z_1,Z_2)\} \\ &= \beta_0+\beta_1 X+\beta_2 Z_1+\beta_3 Z_2+\beta_4 XZ_1+\beta_5 XZ_2 \\ &= m(X,Z_1,Z_2;\beta)\end{aligned}$$

and let $\widehat{\beta}$ be an estimate of β. The logistic SMM estimate can then be obtained from minimising the GMM criterion

$$\left(\frac{1}{n}\sum_{i=1}^{n} g_i\left(\delta,\widehat{\beta}\right)\right)' W_n^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} g_i\left(\delta,\widehat{\beta}\right)\right)$$

where

$$\begin{aligned}g_i\left(\delta,\widehat{\beta}\right) &= s_i\left\{\text{expit}\left(m\left(x_i,z_{i1},z_{i2};\widehat{\beta}\right)-\xi x_i\right)-\alpha\right\} \\ &= s_i\left\{q_i\left(\xi;\widehat{\beta}\right)-\alpha\right\};\quad s_i=\begin{pmatrix}1 & z_{i1} & z_{i2}\end{pmatrix}'.\end{aligned}$$

Considering the one-step GMM estimator with $W_n = \frac{1}{n}\sum_i s_i s_i'$, the first-order condition leads to

$$D'S\left(S'S\right)^{-1}S'v = 0$$

where

$$\begin{aligned} D &= \left\{d_i'\right\};\ S = \left\{s_i'\right\} \\ d_i' &= \left(\begin{array}{cc} 1 & q_i\left(\xi;\widehat{\beta}\right)\left(1-q_i\left(\xi;\widehat{\beta}\right)\right)x_i \end{array}\right);\ v_i = \left\{q_i\left(\xi;\widehat{\beta}\right)-\alpha\right\} \end{aligned}$$

so the instruments get combined in the projection $S\left(S'S\right)^{-1}S'D$, i.e. a constant and the linear projection of $q_i\left(\xi,\widehat{\beta}\right)\left(1-q_i\left(\xi,\widehat{\beta}\right)\right)x_i$ on s_i, again the projection as suggested by Bowden and Vansteelandt (2010).

Conditioning on the estimate $\widehat{\beta}$ will lead to invalid inference for the standard GMM estimation results, as they ignore the first stage estimation. Gouriéroux, Monfort and Renault (1996) show that correct inference for this so-called Two-Stage GMM estimator (2SGMM) is obtained from a first-order expansion around the true values β_0 and δ_0, resulting in

$$\sqrt{n}\left(\widehat{\delta}_{1,\widehat{\beta}} - \delta_0\right) \xrightarrow{d} N\left(0, \left(C_0' W C_0\right)^{-1} C_0 W \Omega_0^* W C_0 \left(C_0' W C_0\right)^{-1}\right),$$

where Ω_0^* is the variance of the limiting normally distributed

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_i\left(\delta_0, \beta_0\right) + E\frac{\partial g_i\left(\delta_0, \beta_0\right)}{\partial \beta'} \sqrt{n}\left(\widehat{\beta} - \beta_0\right).$$

Alternatively, one could specify the joint moments as

$$h_i(\zeta) = \begin{pmatrix} r_i \left(y_i - \mathrm{expit}\left\{m\left(X, Z_1, Z_2; \beta\right)\right\}\right) \\ s_i \left\{\mathrm{expit}\left(m\left(x_i, z_{i1}, z_{i2}; \beta\right) - \xi x_i\right) - \alpha\right\} \end{pmatrix}$$

where $r_i = \left(1, x_i, z_{1i}, z_{2i}, x_i z_{1i}, x_i z_{2i}\right)'$, and estimate $\zeta = \left(\beta', \alpha, \xi\right)'$ jointly by minimising the GMM criterion

$$\left(\frac{1}{n}\sum_{i=1}^{n} h_i(\zeta)\right)' W_n^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} h_i(\zeta)\right).$$

Gouriéroux, Monfort and Renault (1996) show that the asymptotic distributions of the 2SGMM and the joint GMM estimates for δ are the same. An important advantage of using the joint moments is that standard GMM software can be used to estimate the parameters ζ. For example, the *gmm* command in Stata or the *gmm()* function in R can straightforwardly be employed.

Stata syntax:

```
logit y x z1 z2 xz1 xz2
matrix from = e(b)
predict xblog, xb
gmm (invlogit(xblog - x*{psi}) - {ey0}), instruments(z1 z2)
twostep
matrix from = (from,e(b));
* SEs incorrect here
gmm (y - invlogit({logit:x z1 z2 xz1 xz2} + {logitconst})) ///
    (invlogit({logit:} + {logitconst} - x*{psi}) - {ey0}), ///
    instruments(1:x z1 z2 xz1 xz2) instruments(2:z1 z2) ///
    winitial(unadjusted, independent) from(from)
```

Some Monte Carlo Results

We generate data from a logistic SMM model, satisfying the NEM and CMI restrictions. The data is generated from

$$E[Y|X, Z_1, Z_2] = \text{expit}(\beta_0 + (\beta_1 + \xi_0)X + \beta_2 Z_1 + \beta_3 Z_2 + \beta_4 X Z_1 + \beta_5 X Z_2).$$

We set the treatment effect $\xi_0 = 0.6$. We further set $P(Z = 1) = 0.3$; $P(Z = 2) = 0.2$; $P(X = 1|Z = z) = p_{10} + 0.15 \times z$; $E[Y(0)] = 0.19$; $E[Y] = 0.25$; $\beta_1 = 0.15$; $\beta_4 = -0.6$ and $\beta_5 = 0.6$. The other parameters are such that CMI and NEM hold: $\beta_0 = -1.518$; $\beta_2 = 0.3183$; $\beta_3 = -0.5202$; and $p_{10} = 0.4404$.

Table 2. Estimation results for Logistic SMM

Instruments		Z	Z_1, Z_2	Z_1, Z_2
Moments		2SGMM/joint	2SGMM	joint GMM
One-Step	α	0.1912	0.1905	0.1907
		(.0168)	(.0153)	(.0153)
		[.0167]	[.0152]	[.0152]
	ξ	0.5970	0.6033	0.6001
		(.1905)	(.1729)	(.1731)
		[.1899]	[.1722]	[.1721]
Two-Step	α		0.1904	0.1911
			(.0153)	(.0154)
			[.0152]	[.0152]
	ξ		0.6038	0.5957
			(.1729)	(.1735)
			[.1722]	[.1722]
Hansen J			0.9882	0.9827
rej-freq 5%			0.0503	0.0495

Notes: Sample size 10,000. Means of 10,000 MC replications;

We find the Logistic SMM estimators behave well, also for instruments with 6 or even 11 values, although we find that the 2SGMM estimator has a small upward bias for the designs we considered. For example, for an instrument with values $0, 1, 2, ..., 10$, we get means (sd) of the two-step GMM estimates of 0.6323 (0.1073) for 2SGMM and 0.5999 (0.1066) for the joint moments GMM estimator.

Going back to the design with Z taking the values $0, 1, 2$, we change the parameter of Z_2 to $\beta_3 + \tau$ with $\tau = 0.25$. The estimators are severely biased. The GMM estimates using the joint moments has a mean of 1.2805, with a standard deviation of 0.1511. The mean (variance) of Hansen's J-test is equal to 1.26 (3.09) with a rejection frequency at the 5% level of only 8.5%.

In contrast, if we change the parameter of Z_1 to $\beta_2 + \tau$ with $\tau = 0.1$, the estimator has a much smaller bias, with a mean of 0.5527 and standard deviation of 0.1660, but the J-test has much more power in this case, rejecting 49.4% of the times at the 5% level.

Multiple Instruments and Local Treatment Effects

For the single instrument case when the NEM assumption does not hold, but when a monotonicity condition

$$P\left[X\left(1\right)-X\left(0\right)\geq 0\right]=1$$

does hold, then the local risk ratio, defined as

$$LRR=\frac{E\left[Y\left(1\right)|X\left(1\right)>X\left(0\right)\right]}{E\left[Y\left(0\right)|X\left(1\right)>X\left(0\right)\right]}$$

is identified and estimated by the multiplicative SMM. Likewise, the linear SMM estimates in that case the so-called Local Average Treatment Effect (LATE) or Complier Average Causal Effect (CACE), see Imbens and Angrist (1994), defined as

$$LATE=E\left[Y\left(1\right)-Y\left(0\right)|X\left(1\right)>X\left(0\right)\right].$$

Let the values for Z, $\{0, 1, 2, ..., K\}$ be ordered such that $E[YX|Z = k] > E[YX|Z = k - 1]$.We show that, for the one-step GMM estimator,

$$e_z^{-\theta} = \sum_{k=1}^{K} \mu_k e_{k,k-1}^{-\theta}$$

which is a weighted average. Also,

$$e_z^{\theta} = \sum_{k=1}^{K} \tau_k e_{k,k-1}^{\theta}$$

with

$$\begin{aligned} \tau_k &= \left(E\left[Y(X-1)|Z=k\right] - E\left[Y(X-1)|Z=k-1\right]\right) \\ &\cdot \frac{\sum_{l=k}^{K} \pi_l \left(E\left[YX|Z=l\right] - E\left[YX\right]\right)}{\sum_{l=0}^{K} \pi_l E\left[Y(X-1)|Z=l\right]\left(E\left[YX|Z=l\right] - E\left[YX\right]\right)}. \end{aligned}$$

This constitutes a weighted average if $E[YX|Z = k] > E[YX|Z = k - 1]$ and $E[Y(X - 1)|Z = k] > E[Y(X - 1)|Z = k - 1]$.

As an example, consider an instrument that takes the values $Z = \{0, 1, 2, 3\}$, with Y and X generated form a bivariate normal distribution as

$$\begin{aligned} X &= 1\{c_0 + c_1 Z_1 + c_2 Z_2 + c_3 Z_3 - V > 0\} \\ Y &= 1\{b_0 + b_1 X - U > 0\} \\ \begin{pmatrix} U \\ V \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right) \end{aligned}$$

with, as before, $Z_j = 1\{Z = j\}$. We set $\pi_l = P(Z = l) = 0.25$ for all l; the c_l parameters are such that $P(X = 1|Z = l) = 0.1 + 0.1 \times l$; $b_0 = \Phi^{-1}(0.4)$; $b_1 = 0.5$ and $\rho = 0.8$.

The population local risk ratios are then equal to

$$
\begin{aligned}
LRR_{1,0} &= e^{\theta}_{1,0} = \frac{E\left[Y(1)\,|X(1) > X(0)\right]}{E\left[Y(0)\,|X(1) > X(0)\right]} = 1.1585; \\
LRR_{2,1} &= e^{\theta}_{2,1} = \frac{E\left[Y(1)\,|X(2) > X(1)\right]}{E\left[Y(0)\,|X(2) > X(1)\right]} = 1.3227; \\
LRR_{3,2} &= e^{\theta}_{3,2} = \frac{E\left[Y(1)\,|X(3) > X(2)\right]}{E\left[Y(0)\,|X(3) > X(2)\right]} = 1.5303,
\end{aligned}
$$

and the population values of the τ_k are given by
$\tau_1 = 0.3725;\ \tau_2 = 0.3991;\ \tau_3 = 0.2285.$

The one-step GMM estimator will thus be an estimate for the weighted average $\tau_1 LRR_{1,0} + \tau_2 LRR_{2,1} + \tau_3 LRR_{3,2} = 1.3090$. Table 3 presents some estimation results confirming this, for a sample of size 40,000 and for 10,000 Monte Carlo replications.

Table 3. Risk ratio estimation results

	$e^{\theta}_{1,0}$	$e^{\theta}_{2,1}$	$e^{\theta}_{3,2}$	e^{θ}	τ_1	τ_2	τ_3
mean	1.164	1.330	1.542	1.311	0.373	0.399	0.228
st. dev.	0.094	0.121	0.160	0.038	0.027	0.032	0.022

Notes: Estimation results from 10,000 MC replications. Sample size 40,000.

Further, using the two-step GMM results, Hansen's J-test rejects the null 47% of the time at the 5% level, therefore clearly having power to reject this violation of the NEM assumption.

We apply the estimation procedures described above to estimate the causal effect of adiposity on hypertension as in Timpson et al. (2010), using genetic markers as instruments for adiposity. The data are from the Copenhagen General Population Study and the full details of the variable definitions and selection criteria are described in Timpson et al. (2010).

The outcome variable is whether an individual has hypertension, defined as a systolic blood pressure of $>$140 mmHg, diastolic blood pressure of $>$ 90 mmHg, or the taking of antihypertensive drugs.

The intermediate adiposity phenotype is being overweight, defined as having a BMI$>$25.

The two Single Nucleotide Polymorphisms (SNPs) that were used as instruments are the *FTO* and *MC4R* loci, see Frayling et al. (2007) and Loos et al. (2008).

The combinations for the four values of instrument combinations are given in Table 4.

Table 4. Combinations of instruments

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

Table 5 gives the frequency distributions for the hypertension (Y) and overweight (X) variables.

Table 5. Frequency distributions for hypertension overweight

	All		$Z = 0$		$Z = 1$		$Z = 2$		$Z = 3$	
	X									
Y	0	1	0	1	0	1	0	1	0	1
0	0.18	0.12	0.19	0.12	0.19	0.12	0.17	0.13	0.16	13
1	0.25	0.44	0.27	0.42	0.26	0.43	0.23	0.46	0.23	0.48

Table 6. Estimation results

SMM				
Linear	OLS	2SLS	GMM2	J-test
ψ	0.2009 (0.0039)	0.2091 (0.0819)	0.2094 (0.0819)	0.2965
Multiplicative	Gamma	GMM1	GMM2	J-test
θ	0.2974 (0.0063)	0.3090 (0.1192)	0.3104 (0.1192)	0.3071
Logistic	Logistic regression	GMM1	GMM2	J-test
ξ	0.9487 (0.0189)	1.0409 (0.4220)	1.0528 (0.4217)	0.2924

Notes: Sample size 55,523.

Standard errors in brackets; p-values are reported for the J-test.

Although the J-test results do not indicate that the NEM assumptions are not valid, we present in Table 7 the local risk ratio estimation results as described in Section 4. The most precisely estimated risk ratio is $LRR_{2,1} = e^{\theta}_{2,1}$ which gets therefore the largest weight, $\tau_2 = 0.81$.

Table 7. Risk ratio estimation results

	$e^{\theta}_{1,0}$	$e^{\theta}_{2,1}$	$e^{\theta}_{3,2}$	e^{θ}
Coeff	2.2065	1.1086	2.6935	1.3621
95% CI	0.548-8.884	0.791-1.553	0.588-12.336	1.078-1.720
Sample Size	34,896	40,552	20,627	55,523

$$\tau_1 = 0.1037;\ \tau_2 = 0.8082;\ \tau_3 = 0.0881$$

Following Vansteelandt and Goetghebeur (2003), we can use the same GMM format to estimate the logistic SMM with a continuous exposure X.

With a continuous exposure, parametric assumptions have to be made in order to identify causal parameters. Following Vansteelandt and Goetghebeur (2003) and Vansteelandt et al. (2010), we impose that the exposure effect is linear in the exposure on the odds ratio scale and independent of the instrumental variable:

$$\frac{odds\left(Y=1|X,Z\right)}{odds\left\{Y\left(0\right)=1|X,Z\right\}}=\exp\left(\xi_0 X\right).$$

Further, we specify the association model as

$$\begin{aligned}\text{logit}\{E[Y|X,Z]\} &= \text{logit}\{P(Y=1|X,Z_1,Z_2,Z_3)\} \\ &= \beta_0+\beta_1 X+\beta_2 Z_1+\beta_3 Z_2+\beta_4 Z_3 \\ &\quad +\beta_5 XZ_1+\beta_6 XZ_2+\beta_7 XZ_3 \\ &= m(X,Z_1,Z_2,Z_3;\beta)\end{aligned}$$

and estimate the parameters using the joint moment conditions.

For the continuous exposure we use $\left(BMI - \overline{BMI}\right)$, $10\left(\ln BMI - \overline{\ln BMI}\right)$ and $10\left(\ln RELBMI\right)$, where $\ln BMI$ is the natural logarithm of BMI, and $\ln RELBMI$ are the residuals of the regression of $\ln BMI$ on sex, age, age squared, ln(height) and an age-sex interaction, as used in Timpson et al. (2010) to represent relative BMI.

We subtract the mean from BMI and $\ln BMI$ to ensure that zero exposure is part of the data range. We further multiply the $\ln BMI$ and $\ln RELBMI$ by a factor 10 so that the estimated odds ratio is for an increase in exposure of approximately 10%.

Table 8. Estimation results for logistic SMM, continuous exposure

Exposure	*BMI*	ln *BMI*	ln *RELBMI*
ξ	0.1122	0.3035	0.2879
	(0.0384)	(0.1069)	(0.1016)
J-test	0.4714	0.4828	0.5004

Notes: Sample size 55,523. Two-step GMM estimates, using joint moments Standard errors in brackets; p-values are reported for the *J*-test.