

# Package ‘MIIVsem’

May 19, 2017

**Encoding** UTF-8

**Title** Model Implied Instrumental Variable (MIIV) Estimation of  
Structural Equation Models

**Version** 0.5.2

**Description** Functions for estimating structural equation models using  
instrumental variables.

**URL** <https://github.com/zackfisher/MIIVsem>

**BugReports** <https://github.com/zackfisher/MIIVsem/issues>

**Depends** R (>= 2.10)

**Imports** stats, utils, methods, lavaan, numDeriv, Matrix, car, boot

**License** GPL-2

**LazyData** true

**ByteCompile** true

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**Author** Zachary Fisher [aut, cre],  
Keneth Bollen [aut],  
Kathleen Gates [aut],  
Mikko Rönkkö [aut]

**Maintainer** Zachary Fisher <[fish.zachary@gmail.com](mailto:fish.zachary@gmail.com)>

**Repository** CRAN

**Date/Publication** 2017-05-19 21:40:55 UTC

## R topics documented:

MIIVsem-package . . . . .	2
bollen1989a . . . . .	3
bollen1989b . . . . .	4
bollen1989c . . . . .	5
bollen1996 . . . . .	6

estimatesTable . . . . .	7
felson1979 . . . . .	8
miive . . . . .	9
miivs . . . . .	16
print.miive . . . . .	20
print.miivs . . . . .	20
summary.miive . . . . .	21
summary.miivs . . . . .	21

<b>Index</b>	<b>23</b>
--------------	-----------

---

MIIVsem-package	<i>Model Implied Instrumental Variable Estimation of Structural Equation Models</i>
-----------------	---

---

## Description

**MIIVsem** is a package for estimating structural equation models using model-implied instrumental variables (MIIVs).

## Author(s)

**Maintainer:** Zachary Fisher <fish.zachary@gmail.com>

Authors:

- Keneth Bollen
- Kathleen Gates
- Mikko Rönkkö

## See Also

Useful links:

- <https://github.com/zackfisher/MIIVsem>
- Report bugs at <https://github.com/zackfisher/MIIVsem/issues>

---

bollen1989a

*Industrialization-Democracy Data*

---

### Description

A dataset from Bollen (1989) containing measures of political democracy and industrialization for 75 developing countries in 1960 and 1965. The variables are as follows:

### Usage

bollen1989a

### Format

A data frame with 75 rows and 9 variables

### Details

- y1. freedom of the press, 1960
- y2. freedom of political opposition, 1960
- y3. fairness of elections, 1960
- y4. effectiveness of elected legislature, 1960
- y5. freedom of the press, 1965
- y6. freedom of political opposition, 1965
- y7. fairness of elections, 1965
- y8. effectiveness of elected legislature, 1965
- x1. natural log of GNP per capita, 1960
- x2. natural log of energy consumption per capita, 1960
- x3. arcsin of square root of percentage of labor force in industry, 1960

### References

Bollen, K. A. (1989). Structural equation models. New York: Wiley-Interscience.

### Examples

```
## Not run:
model <- '
  Eta1 =~ y1 + y2 + y3 + y4
  Eta2 =~ y5 + y6 + y7 + y8
  Xi1  =~ x1 + x2 + x3
  Eta1 ~ Xi1
  Eta2 ~ Xi1
  Eta2 ~ Eta1
```

```
y1  ~~ y5
y2  ~~ y4
y2  ~~ y6
y3  ~~ y7
y4  ~~ y8
y6  ~~ y8 devtools::build_win()
```

```
## End(Not run)
```

---

bollen1989b

*Union sentiment data*

---

### Description

A dataset from McDonald and Clelland (1984) reanalyzed by Bollen (1989) containing data on union sentiment of southern nonunion textile workers.

### Usage

```
bollen1989b
```

### Format

A data frame with 173 rows and 5 variables

### Details

- deferenc. deference (submissiveness) to managers
- laboract. support for labor activism
- unionsen. sentiment towards unions
- yrsmill. log of years spent in textile mill
- age. centered age

### References

Bollen, K. A. 1989. Structural Equations with Latent Variables. New York: Wiley

McDonald, A. J., & Clelland, D. A. (1984). Textile Workers and Union Sentiment. *Social Forces*, 63(2), 502–521.

## Examples

```
## Not run:
model <- '
  unionsen ~ deferenc + laboract + yrsmill
  deferenc ~ age
  laboract ~ age + deferenc
  yrsmill ~ age
'

## End(Not run)
```

---

bollen1989c

*Subjective class data*

---

## Description

The following data is from Bollen (1989) using data from Kluegel et al. (1977). These data include measures of actual income (inc) and occupational prestige (occ), measures of respondents' subjective assessments of income (subinc), occupational prestige (subocc), and overall SES status (subgen).

## Usage

```
bollen1989c
```

## Format

A data frame with 432 rows and 5 variables

## Details

- occ. actual occupational prestige
- inc. actual income
- subocc. respondents' subjective assessments of prestige
- subinc. respondents' subjective assessments of income
- subgen. respondents' subjective assessments of SES status

## References

- Bollen, K. A. 1989. *Structural Equations with Latent Variables*. New York: Wiley
- Kluegel, J. R., Singleton, R., & Starnes, C. E. (1977). Subjective Class Identification: A Multiple Indicator Approach. *American Sociological Review*, 42(4), 599–611.

## Examples

```
## Not run:
model <- '
subinc ~ inc + subocc
subocc ~ occ + subinc
subgen ~ subinc + subocc
subinc ~~ subocc + subgen
subocc ~~ subgen
inc    ~~ occ
'

## End(Not run)
```

---

bollen1996

*Perceived accessibility data*


---

## Description

Data come from a survey that was conducted in rural clusters of Tanzania in 1993. The goal was to collect information on the perceived accessibility of a specific family planning facility that serviced each cluster. Six informants were chosen: 3 female and 3 male. New informants were chosen for each cluster. Each informant was independently asked to rate the accessibility of the facility, and how easy it was to get to the facility. More specifically the women informants were asked to rate how women of childbearing age perceived the accessibility and easiness and men were asked to rate how accessible and easy men perceived access to the clinic to be. Higher values indicate greater accessibility and ease of travel. The female informants' ratings are 1 to 3 and the male informants' ratings are 4 to 6.

## Usage

```
bollen1996
```

## Format

A data frame with 220 rows and 12 variables

## Details

- access1.
- access2.
- access3.
- access4.
- access5.
- access6.

- easy1.
- easy2.
- easy3.
- easy4.
- easy5.
- easy6.

## References

Bollen, K. A., Speizer, I. S., & Mroz, T. A. (1996). Family Planning Facilities in Rural Tanzania: His and Her Perceptions of Time and Distance.

## Examples

```
## Not run:
model <- '
  femaleAccess =~ access1 + access2 + access3
  maleAccess   =~ access4 + access5 + access6
  femaleEasy   =~ easy1   + easy2   + easy3
  maleEasy     =~ easy4   + easy5   + easy6
'

## End(Not run)
```

---

<code>estimatesTable</code>	<i>return a dataframe of parameter estimates for a fitted model.</i>
-----------------------------	--

---

## Description

return a dataframe of parameter estimates for a fitted model.

## Usage

```
estimatesTable(x, v = NULL, sarg = FALSE)
```

## Arguments

<code>x</code>	An object of class <code>miive</code>
<code>v</code>	A list containing variance and covariance parameter information.
<code>sarg</code>	Logical. Should Sargan test results be included.

---

felson1979

*Attractiveness and academic ability*

---

### Description

This data comes from a study by Felson and Borhnstedt (1979) of perceived attractiveness and academic ability in teenagers, sixth through ninth grade. The six variables are perception of academic ability (academic), perception of physical attractiveness (attract), grade point average (gpa), height, weight, and a strangers' rating of attractiveness (rating).

### Usage

```
felson1979
```

### Format

A data frame with 209 rows and 7 variables

### Details

- acad.
- athl.
- attract.
- gpa.
- height.
- weight.
- rating.

### References

Felson, R.B. & Bohrnstedt, G.W. (1979). "Are the good beautiful or the beautiful good?" The relationship between children's perceptions of ability and perceptions of physical attractiveness. *Social Psychology Quarterly*, 42, 386–392.

### Examples

```
## Not run:
model <- '
acad ~ gpa + attract
attract ~ height + weight + rating + acad
'

## End(Not run)
```



miive

*Model-implied instrumental variable (MIIV) estimation***Description**

Estimate structural equation models using model-implied instrumental variables (MIIVs).

**Usage**

```
miive(model = model, data = NULL, instruments = NULL, sample.cov = NULL,
      sample.mean = NULL, sample.nobs = NULL, sample.cov.rescale = TRUE,
      estimator = "2SLS", se = "standard", bootstrap = 1000L,
      boot.ci = "norm", missing = "listwise", est.only = FALSE,
      var.cov = FALSE, miiv.check = TRUE, ordered = NULL,
      sarg.adjust = "none")
```

**Arguments**

model	A model specified using lavaan model syntax or a <code>miivs</code> object returned by the <code>miivs</code> function. See Details for more information about permissible operators and example model syntax.
data	A data frame, list or environment or an object coercible by <code>as.data.frame</code> to data frame. The most common application is to supply a data.frame.
instruments	This allows user to specify the instruments for each equation. See Details and the <code>miivs.out</code> argument of <code>summary.miivs</code> for more information on the correct input format. External (auxiliary) instruments can be supplied, however, the <code>miiv.check</code> argument must be set to FALSE. In the typical application, the program will choose the MIIVs for each equation based on the model specification. To view the model implied instruments after estimation see the <code>eq.info</code> argument of <code>summary.miive</code> .
sample.cov	Numeric matrix. A sample variance-covariance matrix. The rownames and colnames attributes must contain all the observed variable names indicated in the model syntax.
sample.mean	A sample mean vector. If <code>sample.cov</code> is provided and the <code>sample.mean</code> argument is NULL, intercepts for all endogenous variables will not be estimated.
sample.nobs	Number of observations in the full data frame.
sample.cov.rescale	If TRUE, the sample covariance matrix provided by the user is internally rescaled by multiplying it with a factor (N-1)/N.
estimator	Options "2SLS" or "GMM" for estimating the model parameters. Default is "2SLS". Currently, only 2SLS is supported.
se	If "standard", asymptotic standard errors are computed. If "bootstrap" (or "boot"), bootstrap standard errors are computed.
bootstrap	Number of bootstrap draws, if bootstrapping is used. The default is 1000

<code>boot.ci</code>	Method for calculating bootstrap confidence intervals. Options are normal approximation ("norm"), basic bootstrap interval ("basic"), percentile interval ("perc"), and adjusted bootstrap percentile ("bca"). The default is normal approximation. See <a href="#">boot.ci</a> for more information.
<code>missing</code>	Default is "listwise" however, a maximum likelihood related missing data method called "twostage" is also available. See Details below on missing for more information.
<code>est.only</code>	If TRUE, only the coefficients are returned.
<code>var.cov</code>	If TRUE, variance and covariance parameters are estimated.
<code>miiv.check</code>	Default is TRUE. <code>miiv.check</code> provides a check to determine whether user-supplied instruments are implied by the model specification (e.g. valid MIVs). When auxiliary or external instruments are provided <code>miiv.check</code> should be set to FALSE.
<code>ordered</code>	A vector of variable names to be treated as ordered factors in generating the polychoric correlation matrix and subsequent PIV estimates. See details on ordered below for more information.
<code>sarg.adjust</code>	Adjustment methods used to adjust the p-values associated with the Sargan test due to multiple comparisons. Defaults is none. For options see <a href="#">p.adjust</a> .

## Details

- `model`

The following model syntax operators are currently supported: `=~`, `~`, `~~` and `*`. See below for details on default behavior, descriptions of how to specify the scaling indicator in latent variable models, and how to impose equality constraints on the parameter estimates.

### Example using Syntax Operators

In the model below, '`L1 =~ Z1 + Z2 + Z3`' indicates the latent variable L1 is measured by 3 indicators, Z1, Z2, and Z3. Likewise, L2 is measured by 3 indicators, Z4, Z5, and Z6. The statement '`L1 ~ L2`' specifies latent variable L1 is regressed on latent variable L2. '`Z1 ~~ Z2`' indicates the error of Z2 is allowed to covary with the error of Z3. The label LA3 appended to Z3 and Z6 in the measurement model constrains the factor loadings for Z3 and Z6 to equality. For additional details on constraints see Equality Constraints and Parameter Restrictions.

```
model <- '
  L1 =~ Z1 + Z2 + LA3*Z3
  L2 =~ Z4 + Z5 + LA3*Z6
  L1 ~ L2
  Z2 ~~ Z3
'
```

### Scaling Indicators

Following the **lavaan** model syntax, latent variables are defined using the `=~` operator. For first order factors, the scaling indicator chosen is the first observed variable on the RHS of an equation. For the model below Z1 would be chosen as the scaling indicator for L1 and Z4 would be chosen as the scaling indicator for L2.

```
model <- '
  L1 =~ Z1 + Z2 + Z3
  L2 =~ Z4 + Z5 + Z6
'
```

### Equality Constraints and Parameter Restrictions

Within- and across-equation equality constraints on the factor loading and regression coefficients can be imposed directly in the model syntax. To specify equality constraints between different parameters equivalent labels should be prepended to the variable name using the `*` operator. For example, we could constrain the factor loadings for the two non-scaling indicators of L1 to equality using the following model syntax.

```
model <- '
  L1 =~ Z1 + LA2*Z2 + LA2*Z3
  L2 =~ Z4 + Z5 + Z6
'
```

Researchers also can constrain the factor loading and regression coefficients to specific numeric values in a similar fashion. Below we constrain the regression coefficient of L1 on L2 to 1.

```
model <- '
  L1 =~ Z1 + Z2 + Z3
  L2 =~ Z4 + Z5 + Z6
  L3 =~ Z7 + Z8 + Z9
  L1 ~ 1*L2 + L3
'
```

### Higher-order Factor Models

For example, in the model below, the scaling indicator for the higher-order factor H1 is taken to be Z1, the scaling indicator that would have been assigned to the first lower-order factor L1. The intercepts for lower-order latent variables are set to zero, by default

```
model <- '
  H1 =~ L1 + L2 + L3
  L1 =~ Z1 + Z2 + Z3
  L2 =~ Z4 + Z5 + Z6
  L3 =~ Z7 + Z8 + Z9
'
```

### Model Defaults

In addition to those relationships specified in the model syntax **MIIVsem** will automatically include the intercepts of any observed or latent endogenous variable. The intercepts for any scaling indicators and lower-order latent variables are set to zero by default. Covariances among exogenous latent and observed variables are included when `var.cov = TRUE`. Where appropriate the covariances of the errors of latent and observed dependent variables are automatically included in the model specification. These defaults correspond to those used by **lavaan** and `auto = TRUE`, except that endogenous latent variable intercepts are estimated by default, and the intercepts of scaling indicators are fixed to zero.

### Invalid Specifications

Certain model specifications are not currently supported. For example, the scaling indicator of a latent variable is not permitted to cross-load on another latent variable. In the model below Z1, the scaling indicator for L1, cross-loads on the latent variable L2. Executing a search on the model below will result in the warning: *miivs: scaling indicators with a factor complexity greater than 1 are not currently supported.*

```
model <- '
  L1 =~ Z1 + Z2 + Z3
  L2 =~ Z4 + Z5 + Z6 + Z1
'
```

In addition, **MIIVsem** does not currently support relations where the scaling indicator of a latent variable is also the dependent variable in a regression equation. The model below would not be valid under the current algorithm.

```
model <- '
  L1 =~ Z1 + Z2 + Z3
  Z1 ~ Z4
  Z4 ~ Z5 + Z6
'
```

- **instruments**

To utilize this option you must first define a list of instruments using the syntax displayed below. Here, the dependent variable for each equation is listed on the LHS of the `~` operator. In the case of latent variable equations, the dependent variable is the scaling indicator associated with that variable. The instruments are then given on the RHS, separated by `+` signs. The instrument syntax is then enclosed in single quotes. For example,

```
customIVs <- '
  y1 ~ z1 + z2 + z3
  y2 ~ z4 + z5
'
```

After this list is defined, set the `instruments` argument equal to the name of the list of instruments (e.g. `customIVs`). Note, that `instruments` are specified for an equation, and not for a specific endogenous variable. If only a subset of dependent variables are listed in the `instruments` argument, only those equations listed will be estimated. If external or auxiliary instruments (instruments not otherwise included in the model) are included the `miiv.check` argument should be set to `FALSE`.

- **sample.cov**

The user may provide a sample covariance matrix in lieu of raw data. The rownames and colnames must contain the observed variable names indicated in the model syntax. If `sample.cov` is not `NULL` the user must also supply a vector of sample means (`sample.mean`), and the number of sample observations (`sample.nobs`) from which the means and covariances were calculated. If no vector of sample means is provided intercepts will not be estimated. **MIIVsem** does not support bootstrap standard errors or polychoric instrumental variable estimation when the sample moments, rather than raw data, are used as input.

- **sample.mean**

A vector of length corresponding to the row and column dimensions of the `sample.cov` matrix. The names of `sample.mean` must match those in the `sample.cov`. If the user supplies a covariance matrix but no vector of sample means intercepts will not be estimated.

- `sample.cov.rescale`  
Default is TRUE. If the sample covariance matrix provided by the user should be internally rescaled by multiplying it with a factor  $(N-1)/N$ .
- `estimator`  
The default estimator is 2SLS. For equations with continuous variables only and no restrictions the estimates are identical to those described in Bollen (1996, 2001). If restrictions are present a restricted MIIV-2SLS estimator is implemented using methods similar to those described by Greene (2003) but adapted for moment based estimation. 2SLS coefficients and overidentification tests are constructed using the sample moments for increased computational efficiency.  
If an equation contains ordered categorical variables, declared in the `ordered` argument, the PIV estimator described by Bollen and Maydeu-Olivares (2007) is implemented. The PIV estimator does not currently support exogenous observed predictors of endogenous categorical variables. See details of the `ordered` argument for more information about the PIV estimator.
- `se` When `se` is set to "boot" or "bootstrap" standard errors are computed using a nonparametric bootstrap assuming an independent random sample. If `var.cov = TRUE` nonconvergence may occur and any datasets with improper solutions will be recorded as such and discarded. Bootstrapping is implemented using the `boot` by resampling the observations in data and refitting the model with the resampled data. The number of bootstrap replications is set using the `bootstrap` argument, the default is 1000. Here, the standard errors are based on the standard deviation of successful bootstrap replications. Note, the Sargan test statistic is calculated from the original sample and is not a bootstrap-based estimate. When `se` is set to "standard" standard errors for the MIIV-2SLS coefficients are calculated using analytic expressions. For equations with categorical endogenous variables, the asymptotic distribution of the coefficients is obtained via a first order expansion where the matrix of partial derivatives is evaluated at the sample polychoric correlations. For some details on these standard errors see Bollen & Maydeu-Olivares (2007, p. 315). If `var.cov = TRUE` only point estimates for the variance and covariance estimates are calculated. To obtain standard errors for the variance and covariance parameters we recommend setting `se = "bootstrap"`. Analytic standard errors for the variance covariance parameters accounting for the first stage estimation have been derived and will be available in future releases.
- `missing` There are two ways to handle missing data in **MIIVsem**. First, missing data may be handled by listwise deletion (`missing = "listwise"`), In this case any row of data containing missing observation is excluded from the analysis and the sample moments are adjusted accordingly. Estimation then proceeds normally. The second option for handling missing data is through a two-stage procedures `missing = "twostage"` where consistent estimates of the saturated populations means and covariance are obtained in the first stage. These quantities are often referred to as the "EM means" and "EM covariance matrix." In the second stage the saturated estimates are used to calculate the MIIV-2SLS structural coefficients. Bootstrap standard errors are recommended but will be computationally burdensome due to the cost of calculating the EM-based moments at each bootstrap replication.
- `ordered` For equations containing ordered categorical variables MIIV-2SLS coefficients are estimated using the approach outlined in Bollen & Maydeu-Olivares (2007). The asymptotic distribution of the these coefficients is obtained via a first order expansion where the matrix of

partial derivatives is evaluated at the sample polychoric correlations. For some details on these standard errors see Bollen & Maydeu-Olivares (2007, p. 315). If `var.cov = TRUE` only point estimates for the variance and covariance estimates are calculated using the DWLS estimator in **lavaan**. To obtain standard errors for the variance and covariance parameters we recommend the bootstrap approach. Analytic standard errors for the variance covariance parameters in the presence of endogenous categorical variables will be available in future releases. Currently **MIIVsem** does not support exogenous variables in equations with categorical endogenous variables.

### Sargan's Test of Overidentification

An essential ingredient in the MIIV-2SLS approach is the application of overidentification tests when a given model specification leads to an excess of instruments. Empirically, overidentification tests are used to evaluate the assumption of orthogonality between the instruments and equation residuals. Rejection of the null hypothesis implies a deficit in the logic leading to the instrument selection. In the context of MIIV-2SLS this is the model specification itself. By default, **MIIVsem** provides Sargan's overidentification test (Sargan, 1958) for each overidentified equation in the system. When cross-equation restrictions or missing data are present the properties of the test are not known. When the system contains many equations the `sarg.adjust` option provides methods to adjust the p-values associated with the Sargan test due to multiple comparisons. Defaults is none. For other options see [p.adjust](#).

### References

- Bollen, K. A. (1996). An Alternative 2SLS Estimator for Latent Variable Models. *Psychometrika*, 61, 109-121.
- Bollen, K. A. (2001). Two-stage Least Squares and Latent Variable Models: Simultaneous Estimation and Robustness to Misspecifications. In R. Cudeck, S. Du Toit, and D. Sorbom (Eds.), *Structural Equation Modeling: Present and Future, A Festschrift in Honor of Karl Joreskog* (pp. 119-138). Lincoln, IL: Scientific Software.
- Bollen, K. A., & Maydeu-Olivares, A. (2007). A Polychoric Instrumental Variable (PIV) Estimator for Structural Equation Models with Categorical Variables. *Psychometrika*, 72(3), 309.
- Freedman, D. (1984). On Bootstrapping Two-Stage Least-Squares Estimates in Stationary Linear Models. *The Annals of Statistics*, 12(3), 827-842.
- Greene, W. H. (2000). *Econometric analysis*. Upper Saddle River, N.J: Prentice Hall.
- Hayashi, F. (2000). *Econometrics*. Princeton, NJ: Princeton University Press
- Sargan, J. D. (1958). The Estimation of Economic Relationships using Instrumental Variables. *Econometrica*, 26(3), 393-415.
- Savalei, V. (2010). Expected versus Observed Information in SEM with Incomplete Normal and Nonnormal Data. *Psychological Methods*, 15(4), 352-367.
- Savalei, V., & Falk, C. F. (2014). Robust Two-Stage Approach Outperforms Robust Full Information Maximum Likelihood With Incomplete Nonnormal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2), 280-302.

### See Also

[MIIVsem](#)[miivs](#)

**Examples**

```
# Example 1

bollen1989a_model <- '

  Eta1 =~ y1 + y2 + y3 + y4
  Eta2 =~ y5 + y6 + y7 + y8
  Xi1  =~ x1 + x2 + x3

  Eta1 ~ Xi1
  Eta2 ~ Xi1
  Eta2 ~ Eta1

  y1  ~~ y5
  y2  ~~ y4
  y2  ~~ y6
  y3  ~~ y7
  y4  ~~ y8
  y6  ~~ y8

'
```

```
# Example 2

bollen1989a_model <- '

  Eta1 =~ y1 + y2 + y3 + y4
  Eta2 =~ y5 + y6 + y7 + y8
  Xi1  =~ x1 + x2 + x3

  Eta1 ~ Xi1
  Eta2 ~ Xi1
  Eta2 ~ Eta1

  y1  ~~ y5
  y2  ~~ y4
  y2  ~~ y6
  y3  ~~ y7
  y4  ~~ y8
  y6  ~~ y8

'
```

```
my_instruments <- '

  y1 ~ x2 + x3
  y5 ~ y2 + y3 + y4 + x2
  y2 ~ y3 + y7 + y8 + x2
  y3 ~ y2 + y4 + y6 + y8
  y4 ~ y3 + y6
  y6 ~ y3 + y4 + y7 + x2
  y7 ~ y2 + y4 + y6 + y8
  y8 ~ y2 + y3 + y7 + x2

'
```

```

      x2 ~ y1 + y5 + y2 + y3 + y4 + y6
      x3 ~ y1 + y5 + y2 + y3 + y4 + y6
    ,

miive(model = bollen1989a_model,
      data = bollen1989a,
      instruments = my_instruments)
# Example 3

bollen1989a_model_r <- '

Eta1 =~ y1 + 12*y2 + 13*y3 + 14*y4
Eta2 =~ y5 + 12*y6 + 13*y7 + 14*y8
Xi1  =~ x1 + x2 + 0.5*x3

Eta1 ~ Xi1
Eta2 ~ Xi1
Eta2 ~ Eta1

y1  ~~ y5
y2  ~~ y4
y2  ~~ y6
y3  ~~ y7
y4  ~~ y8
y6  ~~ y8

'

miive(model = bollen1989a_model_r, data = bollen1989a)

```

---

miivs

---

*Model-implied instrumental variable (MIIV) search*


---

## Description

A key step in the MIIV-2SLS approach is to transform the SEM by replacing the latent variables with their scaling indicators minus their errors. Upon substitution the SEM is transformed from a model with latent variables to one containing observed variables with composite errors. The `miivs` function automatically makes this transformation. The `miivs` function will also identify equation-specific model-implied instrumental variables in simultaneous equation models without latent variables.

## Usage

```
miivs(model)
```

## Arguments

`model` A model specified using lavaan model syntax. See the `model` argument within the `lavaanify` function for more information. See the documentation below for a description of how to specify the scaling indicator in latent variable models and impose equality constraints on the parameter estimates.



## Details

- model

A model specified using the model syntax employed by **lavaan**. The following model syntax operators are currently supported:  $\approx$ ,  $\sim$ ,  $\sim\sim$  and  $*$ . See below for details on default behaviors, how to specify the scaling indicator in latent variable models, and how to impose equality constraints on the parameter estimates.

### Example using Syntax Operators

In the model below, 'L1  $\approx$  Z1 + Z2 + Z3' indicates the latent variable L1 is measured by 3 indicators, Z1, Z2, and Z3. Likewise, L2 is measured by 3 indicators, Z4, Z5, and Z6. The statement 'L1  $\sim$  L2' specifies latent variable L1 is regressed on latent variable L2. 'Z1  $\sim\sim$  Z2' indicates the error of Z2 is allowed to covary with the error of Z3. The label LA3 appended to Z3 and Z6 in the measurement model equations constrains the factor loadings for Z3 and Z6 to equality. For additional details on constraints see Equality Constraints and Parameter Restrictions.

```
model <- '
  L1  $\approx$  Z1 + Z2 + LA3*Z3
  L2  $\approx$  Z4 + Z5 + LA3*Z6
  L1  $\sim$  L2
  Z2  $\sim\sim$  Z3
'
```

### Scaling Indicators

Following the **lavaan** model syntax, latent variables are defined using the  $\approx$  operator. For first order factors, the scaling indicator chosen is the first observed variable on the RHS of an equation. For the model below Z1 would be chosen as the scaling indicator for L1 and Z4 would be chosen as the scaling indicator for L2.

```
model <- '
  L1  $\approx$  Z1 + Z2 + Z3
  L2  $\approx$  Z4 + Z5 + Z6
'
```

### Equality Constraints and Parameter Restrictions

Within- and across-equation equality constraints on the factor loading and regression coefficients can be imposed directly in the model syntax. To specify equality constraints between different parameters equivalent labels should be prepended to the variable name using the  $*$  operator. For example, we could constrain the factor loadings for two non-scaling indicators of latent factor L1 to equality using the following model syntax.

```
model <- '
  L1  $\approx$  Z1 + LA2*Z2 + LA2*Z3
  L2  $\approx$  Z4 + Z5 + Z6
'
```

Researchers can also constrain the factor loadings and regression coefficients to specific numeric values in a similar fashion. Below we constrain the regression coefficient of L1 on L2 to 1.

```

model <- '
  L1 =~ Z1 + Z2 + Z3
  L2 =~ Z4 + Z5 + Z6
  L3 =~ Z7 + Z8 + Z9
  L1 ~ 1*L2 + L3
'
```

### Higher-order Factor Models

For example, in the model below, the scaling indicator for the higher-order factor H1 is taken to be Z1, the scaling indicator that would have been assigned to the first lower-order factor L1. The intercepts for lower-order latent variables are set to zero, by default

```

model <- '
  H1 =~ L1 + L2 + L3
  L1 =~ Z1 + Z2 + Z3
  L2 =~ Z4 + Z5 + Z6
  L3 =~ Z7 + Z8 + Z9
'
```

### Model Defaults

In addition to those relationships specified in the model syntax **MIIVsem** will automatically include the intercepts of any observed or latent endogenous variable. The intercepts for any scaling indicators and lower-order latent variables are set to zero. Covariances among exogenous latent and observed variables are included by default when `var.cov = TRUE`. Where appropriate the covariances of the errors of latent and observed dependent variables are also automatically included in the model specification. These defaults correspond to those used by **lavaan** and `auto = TRUE`, except that endogenous latent variable intercepts are estimated by default, and the intercepts of scaling indicators are fixed to zero.

### Invalid Specifications

Certain model specifications are not currently supported. For example, the scaling indicator of a latent variable is not permitted to cross-load on another latent variable. In the model below Z1, the scaling indicator for L1, cross-loads on the latent variable L2. Executing a search on the model below will result in the warning: *miivs: scaling indicators with a factor complexity greater than 1 are not currently supported.*

```

model <- '
  L1 =~ Z1 + Z2 + Z3
  L2 =~ Z4 + Z5 + Z6 + Z1
'
```

In addition, **MIIVsem** does not currently support relations where the scaling indicator of a latent variable is also the dependent variable in a regression equation. For example, the model below would not be valid under the current algorithm.

```

model <- '
  L1 =~ Z1 + Z2 + Z3
  Z1 ~ Z4
  Z4 ~ Z5 + Z6
'
```

The `miivs` function displays a table containing the following information for each equation in the system:

- LHS The "dependent" variable.
- RHS The right hand side variables of the transformed equation.
- MIIVs The model-implied instrumental variables for each equation.

### Value

A list of model equations.

### References

Bollen, K. A. (1996). An Alternative 2SLS Estimator for Latent Variable Models. *Psychometrika*, 61, 109-121.

Bentler, P. M., and Weeks, D. G. (1980). Linear Structural Equations with Latent Variables. *Psychometrika*, 45, 289–308.

### See Also

[miive](#)

### Examples

```
bollen1989a_model <- '
  Eta1 =~ y1 + y2 + y3 + y4
  Eta2 =~ y5 + y6 + y7 + y8
  Xi1  =~ x1 + x2 + x3

  Eta1 ~ Xi1
  Eta2 ~ Xi1
  Eta2 ~ Eta1

  y1   ~~ y5
  y2   ~~ y4
  y2   ~~ y6
  y3   ~~ y7
  y4   ~~ y8
  y6   ~~ y8
',

miivs(bollen1989a_model)
```

---

print.miive	<i>Print method for a MIIV estimation object</i>
-------------	--

---

**Description**

Print method for a MIIV estimation object

**Usage**

```
## S3 method for class 'miive'  
print(x, ...)
```

**Arguments**

x	a miive object
...	Optional arguments to print, not used by user.

---

print.miivs	<i>Print method for a MIIV search object</i>
-------------	--

---

**Description**

Print method for a MIIV search object

**Usage**

```
## S3 method for class 'miivs'  
print(x, ...)
```

**Arguments**

x	a miivs object
...	Optional arguments to print, not used by user.

---

summary.miive	<i>Summary information for a MIIV estimation object</i>
---------------	---

---

**Description**

Summary information for a MIIV estimation object

**Usage**

```
## S3 method for class 'miive'
summary(object, eq.info = FALSE, restrict.tests = FALSE,
        rsquare = FALSE, ...)
```

**Arguments**

object	An object of class miive
eq.info	A logical indicating whether equation-specific information should be printed. Useful in models with large numbers of variables.
restrict.tests	A logical indicating whether two test statistics for a large-sample wald test of linear hypotheses imposed on the MIIV-2SLS coefficient matrix should be provided. The first statistic is an approximate F and the second is Chi-square. Assumptions and additional details for each test are given by Greene (2000, p. 346-347) and Henningsen and Hamman (2007).
rsquare	A logical indicating whether R-square values for endogeneous variables are included in the output. Only available when var.cov is TRUE.
...	Optional arguments to summary, not used by user.

**References**

Greene, W. H. (2000). *Econometric analysis*. Upper Saddle River, N.J: Prentice Hall.

Henningsen, A., and Hamann, J.D. (2007). *systemfit: A Package for Estimating Systems of Simultaneous Equations in R*. *Journal of Statistical Software* 23(4), 1-40.

---

summary.miivs	<i>Summary information for a MIIV search object</i>
---------------	---

---

**Description**

Summary information for a MIIV search object

**Usage**

```
## S3 method for class 'miivs'
summary(object, miivs.out = FALSE, eq.info = FALSE, ...)
```

**Arguments**

object	An object of class <code>miivs</code>
<code>miivs.out</code>	A logical indicating whether the model-implied instrumental variables found for <code>model</code> should be printed to the console. This is a temporary convenience function to provide an editable, baseline format, for the <code>instruments</code> argument of <code>miive</code> .
<code>eq.info</code>	A logical indicating whether equation-specific information should be printed. Useful in models with a large number of variables.
<code>...</code>	Optional arguments to <code>summary</code> , not used by user.

# Index

- \*Topic **2sls**
    - miive, 9
  - \*Topic **MIIV-2SLS**
    - miive, 9
  - \*Topic **MIIV**
    - miive, 9
  - \*Topic **PIV**
    - miive, 9
  - \*Topic **SEM**
    - miive, 9
  - \*Topic **datasets**
    - bollen1989a, 3
    - bollen1989b, 4
    - bollen1989c, 5
    - bollen1996, 6
    - felson1979, 8
  - \*Topic **instrument**
    - miive, 9
  - \*Topic **least-squares**
    - miive, 9
  - \*Topic **tsls**
    - miive, 9
  - \*Topic **two-stage**
    - miive, 9
- bollen1989a, 3  
bollen1989b, 4  
bollen1989c, 5  
bollen1996, 6  
boot.ci, 10
- estimatesTable, 7
- felson1979, 8
- lavaanify, 16
- miive, 9, 19, 22  
miivs, 9, 16  
MIIVsem, 14  
MIIVsem (MIIVsem-package), 2
- MIIVsem-package, 2
- p.adjust, 10, 14  
print.miive, 20  
print.miivs, 20
- summary.miive, 9, 21  
summary.miivs, 9, 21