

intro 8 — Robust and clustered standard errors

[Description](#) [Options](#) [Remarks and examples](#) [Also see](#)

Description

`sem` and `gsem` provide two options to modify how standard-error calculations are made: `vce(robust)` and `vce(cluster clustvar)`. These standard errors are less efficient than the default standard errors, but they are valid under less restrictive assumptions.

These options are allowed only when default estimation method `method(ml)` is used or when option `method(mlmv)` is used. `ml` stands for maximum likelihood, and `mlmv` stands for maximum likelihood with missing values; see [\[SEM\] intro 4](#), [\[SEM\] sem](#), and [\[SEM\] gsem](#).

Also see [\[SEM\] intro 9](#), entitled *Standard errors, the full story*.

Options

`vce(vctype)` specifies how the VCE, and thus the standard errors, is calculated. VCE stands for variance–covariance matrix of the estimators. The standard errors that `sem` and `gsem` report are the square roots of the diagonal elements of the VCE.

`vce(oim)` is the default. `oim` stands for observed information matrix (OIM). The information matrix is the matrix of second derivatives, usually of the log-likelihood function. The OIM estimator of the VCE is based on asymptotic maximum-likelihood theory. The VCE obtained in this way is valid if the errors are independent and identically distributed normal, although the estimated VCE is known to be reasonably robust to violations of the normality assumption, at least as long as the distribution is symmetric and normal-like.

`vce(robust)` specifies an alternative calculation for the VCE, called robust because the VCE calculated in this way is valid under relaxed assumptions. The method is formally known as the Huber/White/sandwich estimator. The VCE obtained in this way is valid if the errors are independently distributed. It is not required that the errors follow a normal distribution, nor is it required that they be identically distributed from one observation to the next. Thus the `vce(robust)` VCE is robust to heteroskedasticity of the errors.

`vce(cluster clustvar)` is a generalization of the `vce(robust)` calculation that relaxes the assumption of independence of the errors and replaces it with the assumption of independence between clusters. Thus the errors are allowed to be correlated within clusters.

Remarks and examples

[stata.com](#)

The `vce()` option is allowed by `sem` and `gsem`. In the rest of this entry, we will use `sem` in illustrations, but everything we say applies equally to `gsem`.

The `vce(robust)` option,

```
. sem ..., ... vce(robust)
```

and the `vce(cluster clustvar)` option,

```
. sem ..., ... vce(cluster clustvar)
```

relax assumptions that are sometimes unreasonable for a given dataset and thus produce more accurate standard errors in those cases. Those assumptions are homoskedasticity of the variances of the errors—`vce(robust)`—and independence of the observations—`vce(cluster clustvar)`. `vce(cluster clustvar)` relaxes both assumptions.

Homoskedasticity means that the variances of the errors are the same from observation to observation. Homoskedasticity can be unreasonable if, for instance, the error corresponds to a dependent variable of income or socioeconomic status. It would not be unreasonable to instead assume that, in the data, the variance of income or socioeconomic status increases as the mean increases. In such cases, rather than typing

```
. sem (y<-...) (...) (...<-x1) (...<-x2)
```

you would type

```
. sem (y<-...) (...) (...<-x1) (...<-x2), vce(robust)
```

Independence implies that the observations are uncorrelated. If you have observations on people, some of whom live in the same neighborhoods, it would not be unreasonable to assume instead that the error of one person is correlated with those of others who live in the same neighborhood because neighborhoods tend to be homogeneous. In such cases, if you knew the neighborhood, rather than typing

```
. sem (y<-...) (...) (...<-x1) (...<-x2)
```

you would type

```
. sem (y<-...) (...) (...<-x1) (...<-x2), vce(cluster neighborhood)
```

Understand that if the assumptions of independent and identically distributed normal errors are met, the `vce(robust)` and `vce(cluster clustvar)` standard errors are less efficient than the standard `vce(oim)` standard errors. Less efficient means that for a given sample size, the standard errors jump around more from sample to sample than would the `vce(oim)` standard errors. `vce(oim)` standard errors are unambiguously best when the standard assumptions of homoskedasticity and independence are met.

Also see

[SEM] [intro 7](#) — Postestimation tests and predictions

[SEM] [intro 9](#) — Standard errors, the full story

[SEM] [sem option method\(\)](#) — Specifying method and calculation of VCE

[SEM] [gsem estimation options](#) — Options affecting estimation