finregress capm — Capital asset pricing model (CAPM)+

+This command is part of StataNow.

Description

`finregress capm` fits a time-series capital asset pricing model (CAPM), which relates assets' characteristics to risk factors of interest. A collection of dependent variables, usually asset returns, is regressed on a common set of independent variables, usually including a benchmark return (an overall market return) and possibly additional factors. Both the independent and dependent variables can be adjusted for a risk-free return. Coefficients for all assets are estimated simultaneously, allowing for joint tests across assets.

Quick start

Regress returns `r1`, `r2`, and `r3` on a market portfolio return `rmkt`

```
finregress capm r1 r2 r3 = rmkt
```

Same as above, but adjust dependent variables by subtracting risk-free rate `rf`

```
finregress capm r1 r2 r3 = rmkt, rfrate(rf)
```

Same as above, but adjust independent variable `rmkt` for the risk-free rate

```
finregress capm r1 r2 r3 = rmkt, rfrate(rf) adjust
```

Three-factor CAPM, adjusting `rmkt` but not variable `x1` or `x2`

```
finregress capm r1 r2 r3 = rmkt x1 x2, rfrate(rf) adjust(rmkt)
```

Same as above, but include an economic recession indicator `rec`

```
finregress capm r1 r2 r3 = rmkt x1 x2 i.rec, rfrate(rf) adjust(rmkt)
```

Menu

Statistics > Financial statistics > Capital asset pricing model (CAPM)

Syntax

Basic syntax

finregress capm *depvars* [= [*indepvars*]] [*if*] [*in*] [*weight*] [, *options*]

Syntax with a risk-free rate

finregress capm *depvars* [= [*indepvars*]] [*if*] [*in*] [*weight*] , rfrate(*varname*) [*options*]

options	Description
Model	
rfrate(*varname*)	specify risk-free rate variable and adjust all *depvars*
adjust	adjust all independent variables by risk-free rate variable
adjust(*varlist*)	adjust independent variables in *varlist* by risk-free rate variable
noconstant	omit constant term
SE/Robust	
vce(*vcetype*)	*vcetype* may be robust, unadjusted, cluster *clustvar*, or hac *hacspec*
Reporting	
level(#)	set confidence level; default is level(95)
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
coeflegend	display legend instead of statistics

You must tsset your data before using finregress capm; see [TS] **tsset**.

indepvars may contain factor variables; see **[U] 11.4.3 Factor variables**.

depvars, *indepvars*, *varname*, and *varlist* may contain time-series operators; see **[U] 11.4.4 Time-series varlists**.

collect is allowed; see **[U] 11.1.10 Prefix commands**.

aweights and fweights are allowed; see **[U] 11.1.6 weight**.

coeflegend does not appear in the dialog box.

See **[U] 20 Estimation and postestimation commands** for more capabilities of estimation commands.

Options

Model

rfrate(*varname*) specifies the risk-free rate used in the CAPM. If rfrate() is not specified, no adjustment is made, and the risk-free rate is implicitly set to 0 in all periods. When rfrate() is specified, it indicates that you wish to transform all dependent variables by subtracting the same risk-free rate, the variable in rfrate().

adjust and adjust(*varlist*) specify which independent variables are to be adjusted by subtracting the risk-free rate in rfrate(). adjust adjusts all independent variables. adjust(*varlist*) adjusts only the independent variables specified in *varlist*.

noconstant omits the constants in the CAPM.

SE/Robust

vce(*vcetype*) specifies the type of standard error reported, which includes types that are robust to some kinds of misspecification (robust) and that allow for intragroup correlation (cluster *clustvar*); see [R] ***vce_option***.

vce(unadjusted) requests conventional standard errors appropriate under homoskedasticity and no autocorrelation. See [R] **gmm** for details.

vce(hac *hacspec*) requests a heteroskedasticity- and autocorrelation-consistent (HAC) variance–covariance matrix. The full syntax of *hacspec* is one of the following:

vce(hac *kernel* [#]) requests a HAC variance–covariance matrix using the specified kernel (see below) with optional # lags. The bandwidth of a kernel is equal to # + 1. If # is not specified, a kernel with $N - 2$ lags is used, where N is the sample size.

vce(hac *kernel* opt [#]) requests a HAC variance–covariance matrix using the specified kernel (see below), and the lag order is selected using Newey and West's (1994) optimal lag-selection algorithm. # is an optional tuning parameter that affects the lag order selected; see the discussion in *Methods and formulas* in [R] **ivregress**.

kernel may be one of the following:

bartlett or nwest requests the Bartlett (Newey–West) kernel.

parzen or gallant requests the Parzen (Gallant 1987) kernel.

quadraticspectral or andrews requests the quadratic spectral (Andrews 1991) kernel.

Reporting

level(#); see [R] **Estimation options**.

display_options: noci, nopvalues, noomitted, vsquish, noemptycells, baselevels, allbaselevels, nofvlabel, fvwrap(#), fvwrapon(*style*), cformat(%*fmt*), pformat(%*fmt*), sformat(%*fmt*), and nolstretch; see [R] **Estimation options**.

The following option is available with finregress capm but is not shown in the dialog box:

coeflegend; see [R] **Estimation options**.

Remarks and examples

A CAPM relates asset characteristics, usually returns, to independent variables, often called risk factors or simply factors. These regressions can be run asset by asset or fit jointly on a collection of returns.

The classic CAPM relates the returns of each asset i at time t, r_{it}, to the overall market return, r_t^m. The rate of return one may obtain without incurring any risk, r_t^f, is also relevant because only returns in excess of this risk-free rate are considered gains. The model is

$$r_{it} - r_t^f = \alpha_i + \beta_i(r_t^m - r_t^f) + e_{it}$$

The coefficient β_i is called the market beta of asset i. Mechanically, β is a regression coefficient like any other. It has the standard interpretation: if the market return rises by one unit, then the stock's return is expected to rise by β units. Several values of β have special interpretation.

- $\beta < 0$ indicates that the stock moves against the market.
- $\beta = 0$ indicates that the stock's return is unrelated to the market return.
- $0 < \beta < 1$ indicates that the stock's return fluctuates in the same direction as the market but less than one for one.
- $\beta = 1$ indicates that the stock's return fluctuates one for one with the market return.
- $\beta > 1$ indicates that the stock's return varies more than one for one with the market.

The intercept α_i represents expected returns that are not captured by the independent variables in the model. The main assumption of the CAPM is that the intercept terms α_i's are zero. A test of this assumption is carried out by the Gibbons–Ross–Shanken procedure.

▷ Example 1: Single-factor CAPM

We use our fictional financial prices dataset and describe the first few variables.

```
. use https://www.stata-press.com/data/r19/finex
(Fictional stock price data)

. describe datestr-wgt

Variable      Storage   Display    Value
    name         type    format    label      Variable label
-------------------------------------------------------------------------------
datestr         str11   %11s                  String date
datem           int     %tm                   Monthly date
sp500           double  %10.0g                S&P 500
vol             float   %9.0g                 Volatility index
fedfunds        float   %9.0g                 Federal funds rate
acme            float   %9.0g                 Aciron Medical, Inc.
bat             float   %9.0g                 Boron Advanced Technologies
iron            float   %9.0g                 Industrial Operations Network
dune            float   %9.0g                 Digital Urban Network Enterprise
tyr             float   %9.0g                 Tyndale Resources Group
glo             float   %9.0g                 Green Logistics, Inc.
spa             float   %9.0g                 Space Rocket MFG
wgt             float   %9.0g                 Widget Gadgets
```

Because the dataset contains prices and `finregress` works on returns, we use `finreturns` to convert price series into monthly log returns. We multiply by 100 so that these returns are approximately percentage changes.

```
. quietly finreturns acme-tyr, log(lnr_) multiply(100)
```

Next, we use `finreturns` to calculate monthly log returns for the stock market (`sp500`), which will serve as our independent variable or overall market factor. The question of interest is how excess returns on individual assets vary with excess returns on this benchmark.

```
. quietly finreturns sp500, log(lnr_mkt) multiply(100)
```

To measure excess returns, we use the return in excess of a risk-free rate. The `fedfunds` variable contains the annualized federal funds interest rate and serves as our risk-free rate. Because the asset returns are all measured in monthly log changes, we approximate the monthly risk-free return by constructing the annual log risk-free return and dividing by 12.

```
. generate double rf = 100 * log(1 + fedfunds/100)/12
```

With this preprocessing complete, we are ready to fit a CAPM relating individual excess log stock returns to the excess log return on the market.

Data come in many formats. If our dependent and independent variables were already adjusted for risk-free rate `rf`, we would fit the CAPM by typing

```
. finregress capm lnr_acme-lnr_tyr = lnr_mkt
```

In our case, however, the variables have not been adjusted, so we type

```
. finregress capm lnr_acme-lnr_tyr = lnr_mkt, rfrate(rf) adjust
Capital asset pricing model
Sample: 1955m2 thru 2019m12                           Number of obs = 779
```

	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
lnr_acme						
lnr_mkt	.1316993	.0102147	12.89	0.000	.111679	.1517197
_cons	.0322367	.0378174	0.85	0.394	-.041884	.1063575
lnr_bat						
lnr_mkt	1.530731	.0114377	133.83	0.000	1.508314	1.553148
_cons	.0760026	.0416309	1.83	0.068	-.0055925	.1575977
lnr_iron						
lnr_mkt	1.871524	.012273	152.49	0.000	1.847469	1.895579
_cons	.0556327	.0454103	1.23	0.221	-.0333698	.1446352
lnr_dune						
lnr_mkt	1.878887	.0123726	151.86	0.000	1.854637	1.903137
_cons	.0895071	.046235	1.94	0.053	-.0011119	.1801261
lnr_tyr						
lnr_mkt	1.087576	.0117232	92.77	0.000	1.064599	1.110553
_cons	.0282078	.0443328	0.64	0.525	-.0586828	.1150984

```
Notes: Dependent variables adjusted for risk-free rate rf.
       Independent variable lnr_mkt adjusted for risk-free rate rf.
```

The `rfrate()` option specifies which variable our asset returns are in "excess" of. When `rfrate()` is specified, dependent variables (in this case `lnr_acme-lnr_tyr`) are transformed prior to estimation, generating excess returns r^e_{it} as

$$r^e_{it} = r_{it} - r^f_t$$

In addition, by specifying `adjust`, we also transformed `lnr_mkt` prior to estimation by subtraction of the same risk-free rate variable. Thus, the regression actually run, for asset i, is

$$(r_{it} - r^f_t) = \alpha_i + \beta_i(r^m_t - r^f_t) + e_{it}$$

Turning to the output, we see that each block in the output table corresponds to one asset return. Two coefficients were estimated for each return series: a coefficient on the market and an intercept. All the chosen assets covary positively with the market, with response coefficients ranging from 0.13 to 1.88. The units on both sides of the equation are monthly excess log returns, so the interpretation is that a one-unit change in excess log return in the market is associated with an expected β_i change in excess log return in each asset, measured monthly. The intercept can be interpreted as the excess log return in that asset when the excess log return on the market is 0. For example, the ACME stock (equation lnr_acme) has an intercept of about 0.03, indicating that even when excess log market return is 0, the ACME stock still generates an excess monthly log return of about 0.03 or approximately 3 basis points.

◁

There is no requirement that the CAPM uses a single independent variable or factor. Indeed, multifactor models have become popular; see Fama and French (1992, 1996) and Campbell, Lo, and MacKinlay (1997, chap. 6).

▷ Example 2: Multifactor CAPM

Continuing with the previous example, we add the vol factor as an independent variable to the model for our five assets. This factor is a return on a fictional volatility index. We use the volatility index as a control, and we do not want to adjust it for the risk-free rate.

```
. finregress capm lnr_acme-lnr_tyr = lnr_mkt vol, rfrate(rf) adjust(lnr_mkt)
Capital asset pricing model
Sample: 1955m2 thru 2019m12                          Number of obs = 779
```

	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
lnr_acme						
lnr_mkt	.1483932	.0119328	12.44	0.000	.1250054	.171781
vol	-.0966061	.0356362	-2.71	0.007	-.1664517	-.0267604
_cons	-.0086329	.0419574	-0.21	0.837	-.0908679	.0736021
lnr_bat						
lnr_mkt	1.52412	.0131144	116.22	0.000	1.498417	1.549824
vol	.0382555	.0400207	0.96	0.339	-.0401837	.1166947
_cons	.0921868	.0450832	2.04	0.041	.0038254	.1805482
lnr_iron						
lnr_mkt	1.851205	.0145758	127.01	0.000	1.822636	1.879773
vol	.1175864	.0443886	2.65	0.008	.0305864	.2045865
_cons	.1053782	.0479793	2.20	0.028	.0113406	.1994159
lnr_dune						
lnr_mkt	1.872238	.014166	132.16	0.000	1.844473	1.900003
vol	.0384778	.0399057	0.96	0.335	-.039736	.1166915
_cons	.1057853	.049148	2.15	0.031	.0094571	.2021135
lnr_tyr						
lnr_mkt	1.083229	.0131535	82.35	0.000	1.057448	1.109009
vol	.0251582	.0392261	0.64	0.521	-.0517235	.1020398
_cons	.0388511	.0483987	0.80	0.422	-.0560086	.1337109

```
Notes: Dependent variables adjusted for risk-free rate rf.
       Independent variable lnr_mkt adjusted for risk-free rate rf.
```

The volatility factor appears with coefficients ranging from −0.10 on the lnr_acme return to 0.12 on the lnr_iron return. The results from these regressions provide information on the level and distribution of the coefficients across assets. We can characterize which assets covary strongly with particular independent variables. One key prediction of the CAPM that is not tested here is whether the coefficients themselves vary systematically with average returns. To investigate this question, see [FIN] **finregress fmb**.

◁

Stored results

`finregress capm` stores the following in `e()`:

Scalars	
`e(N)`	number of observations
`e(k_dv)`	number of dependent variables
`e(k_indepvars)`	number of independent variables, not including the constant
`e(tmin)`	minimum time
`e(tmax)`	maximum time
Macros	
`e(cmd)`	`finregress`
`e(subcmd)`	`capm`
`e(cmdline)`	command as typed
`e(depvar)`	names of dependent variables
`e(indepvars_unadj)`	names of unadjusted independent variables
`e(indepvars_adj)`	names of adjusted independent variables
`e(wtype)`	weight type
`e(wexp)`	weight expression
`e(title)`	title in estimation output
`e(vcetype)`	title used to label Std. err.
`e(vce)`	*vcetype* specified in `vce()`
`e(rfrate)`	risk-free rate variable, if specified
`e(tmaxs)`	formatted maximum time
`e(tmins)`	formatted minimum time
`e(tvar)`	time variable
`e(properties)`	`b V`
`e(estat_cmd)`	program used to implement `estat`
`e(predict)`	program used to implement `predict`
`e(asbalanced)`	factor variables `fvset` as `asbalanced`
`e(asobserved)`	factor variables `fvset` as `asobserved`
Matrices	
`e(b)`	coefficient vector
`e(V)`	variance–covariance matrix of the estimator
`e(alpha)`	vector of intercepts
Functions	
`e(sample)`	marks estimation sample

In addition to the above, the following is stored in `r()`:

Matrices	
`r(table)`	matrix containing the coefficients with their standard errors, test statistics, *p*-values, and confidence intervals

Note that results stored in `r()` are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.

Methods and formulas

Let r_{it} be the return for asset i and time t, where $i = 1, 2, \ldots, K$ and $t = 1, 2, \ldots, T$. The CAPM relates r_{it} to the overall market return, r_t^m, while adjusting for the risk-free rate, r_t^f.

We write

$$r_{it} - r_t^f = \alpha_i + \beta_i(r_t^m - r_t^f) + e_{it}$$

for the single-factor CAPM and

$$r_{it} - r_t^f = \alpha_i + \beta_{1i}(x_{1t} - r_t^f) + \beta_{2i}(x_{2t} - r_t^f) + \cdots + e_{it}$$

for the multifactor CAPM. α_i's and the β's are estimated parameters, and e_{it} is the error. Notice the i-index on the coefficients; each asset return series is allowed to have its own intercept and beta. In many applications, one of the independent variables is the excess return on a benchmark or market index; the coefficient on this variable is referred to as an asset's market beta. The intercept, which represents average excess returns that are not captured by the independent variables, is called Jensen's alpha.

The CAPM can also be thought of as a factor structure on the behavior of returns. Let $\mathbf{r}_t$ be an $K \times 1$ column vector of (adjusted) returns. Let $\mathbf{f}_t$ be an $M \times 1$ column vector of (adjusted) factors. Note that these factors are observed independent variables in this case. Then the CAPM is

$$\mathbf{r}_t = \boldsymbol{\alpha} + \boldsymbol{\Lambda}\mathbf{f}_t + \mathbf{e}_t$$

where $\boldsymbol{\alpha}$ is an $K \times 1$ vector of intercept terms, $\boldsymbol{\Lambda}$ is an $K \times M$ matrix of coefficients, and $\mathbf{e}_t$ is an $K \times 1$ vector of error terms.

The collection of time-series regressions is estimated jointly by the generalized method of moments; see [R] **gmm** for details.

References

Andrews, D. W. K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59: 817–858. https://doi.org/10.2307/2938229.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press. https://doi.org/10.2307/j.ctt7skm5.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47: 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x.

———. 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51: 55–84. https://doi.org/10.1111/j.1540-6261.1996.tb05202.x.

Gallant, A. R. 1987. *Nonlinear Statistical Models*. New York: Wiley. https://doi.org/10.1002/9780470316719.

Newey, W. K., and K. D. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61: 631–653. https://doi.org/10.2307/2297912.

Also see

[FIN] **finregress capm postestimation** — Postestimation tools for finregress capm+

[FIN] **finportfolio** — Financial portfolio selection+

[FIN] **finregress fmb** — Fama–MacBeth regression+

[FIN] **finreturns** — Generate financial returns+

[U] 20 Estimation and postestimation commands

For suggested citations, see the FAQ on citing Stata documentation.