Linking frames in Stata

Jeff Pitblado

StataCorp LLC

2023 German Stata Conference
Berlin, Germany

Outline

Frames

Frames

- ▶ Stata starts with a frame named `default`.
- ▶ `default` is the current frame.
- ▶ The current frame is where commands have access to data.

Example

```
. frames dir
  default  0 x 0

. frame
  (current frame is default)
```

Frames

Creating frames

► Use `frame create` to create a new frame.
► The current frame is not affected.

Example

```
. frame create play

. frames dir
  default  0 x 0
  play     0 x 0

. frame
  (current frame is default)
```

Frames

Change the current frame
Change the current frame with `frame change`.

Example

```
. frame
  (current frame is default)

. frame change play

. frame
  (current frame is play)
```

Frames

Frame prefix

The `frame` prefix syntax let's you run a command in a specified frame.

Example

```
. sysuse census
(1980 Census data by state)

. frame default : sysuse auto
(1978 automobile data)

. frames dir
  default  74 x 12; 1978 automobile data
  play     50 x 13; 1980 Census data by state
```

Frames

Frame block
The `frame` block syntax let's you run multiple commands in a specified frame.

Example
```
. frame default {
.         clear
.         sysuse cancer
(Patient survival in drug trial)
. }

. frames dir
  default  48 x 8; Patient survival in drug trial
  play     50 x 13; 1980 Census data by state
```

Frames

Be careful when using macros with the prefix syntax.

Example

```
. frame default : display c(N)
48

. frame default : display `c(N)´
50

. frame default {
.        display c(N)
48
. }

. frame default {
.        display `c(N)´
48
. }
```

Frames management

```
. frame
  (current frame is play)

. frame rename play census

. frame
  (current frame is census)

. frame copy default cancer

. frames dir
  cancer   48 x 8; Patient survival in drug trial
  census   50 x 13; 1980 Census data by state
  default  48 x 8; Patient survival in drug trial

. frame drop default

. frames dir
  cancer   48 x 8; Patient survival in drug trial
  census   50 x 13; 1980 Census data by state
```

Frames

Copy data into a new frame

Use `frame put` to copy selected variables and observations from the current frame to a new frame.

Example

```
. frame cancer : ///
>           frame put if died == 0, into(survivors)

. frames dir
  cancer      48 x 8; Patient survival in drug trial
  census      50 x 13; 1980 Census data by state
* survivors   17 x 8; Patient survival in drug trial

Note: Frames marked with * contain unsaved data.
```

Frames

Post results to a frame

Use `frame post` to add new observations to a frame.

Example

```
frames reset
frame create results ///
       str18 coefficient ///
       double estimate ///
       double se
sysuse auto
generate double logmpg = log(mpg)
regress logmpg turn trunk, eform(exp(b))
local cols : colname e(b)
foreach x of local cols {
       frame post results ///
               ("`x´") (_r_b[`x´]) (_r_se[`x´])
}
frame results : list
```

Frame post, continued

Example

```
. regress, eform(exp(b)) noheader nopvalue
```

logmpg	exp(b)	Std. err.	[95% conf. interval]	
turn	.9658032	.0053207	.9552522	.9764709
trunk	.9838846	.0055748	.9728313	.9950635
_cons	102.2701	18.77131	70.92596	147.4661

```
. frame results : list
```

	coeffi_t	estimate	se
1.	turn	.96580324	.00532068
2.	trunk	.9838846	.00557482
3.	_cons	102.27012	18.771306

Linking frames

Linking frames

Use frlink to link related datasets in different frames.

Example

We will work with data on people, linking them to family data and another group-level dataset.

Linking frames - person data

Person data

- ► Each observation represents a person
- ► `family_id` is the family identifier
- ► `group_id` is the group identifier
- ► `x` and `y` variables are measurements/properties of the person

Example

```
. frames reset

. use person

. frame rename default person

. describe, simple
family_id   x1          x3          y2
group_id    x2          y1
```

Linking frames - family data

Family data

► Each observation represents a family

► `family_id` is the family identifier

► `x` variables are measurements/properties of the family

Example

```
. frame create family

. frame family : use family

. frame family : describe, simple
family_id  x1          x2          x3
```

Linking frames - group data

Group data

- ▶ Each observation represents a group
- ▶ `group_id` is the group identifier
- ▶ `x` variables are measurements/properties of the group

Example

```
. frame create group

. frame group : use group

. frame group : describe, simple
group_id  x1          x2          x3
```

Linking frames

Linking frames

- ▶ Use `family_id` to link the person data to the family data.
- ▶ Use `group_id` to link the person data to the group data.
- ▶ `frlink` creates new variables that contain the linkage characteristics and point to observations in the other frames.

Example

```
. frlink m:1 family_id, frame(family)
(26 observations in frame person unmatched)

. frlink m:1 group_id, frame(group)
(all observations in frame person matched)

. describe, simple
family_id   x2          y2
group_id    x3          family
x1          y1          group
```

Linking frames

Save a set of frames to disk
▶ Use `frames save` to save multiple frames to a single file.
▶ Option `linked` will also add any linked frames to the list of specified frames to save.

Example
```
. frames save linked, frames(person) linked
file linked.dtas saved
```

Linking frames

Describe frames

Use `frames describe` to get a peek at the fames in memory or saved to disk.

Example

```
. frames describe using linked, simple
_____

Frame: person
family_id  x2        y2
group_id   x3        family
x1         y1        group
_____


_____

Frame: family
family_id  x1        x2        x3
_____


_____

Frame: group
group_id   x1        x2        x3
_____
```

Linking frames

Load a set of frames from disk

▶ Use `frames use` to load your saved frames into Stata.

▶ All linkage information is present.

Example

```
. frames reset

. frames use linked
  family  200 x 4
  group   10 x 4
  person  653 x 9

. frame
  (current frame is default)

. frame change person
```

Variables in linked frames

Copy variables from linked frames

- ▶ Use `frget` to copy variables from a linked frame.
- ▶ Unmatched observations yield missing values.
- ▶ Consumes memory for each observation in the current frame.

Example

```
. frame family : describe, simple
family_id  x1          x2          x3

. frget x?, from(family) prefix(f)
(26 missing values generated)
(26 missing values generated)
(26 missing values generated)
(3 variables copied from linked frame)
```

Variables in linked frames

Copy variables from linked frames

Call `frget` separately for each linked frame.

Example

```
. frame group : describe, simple
group_id  x1          x2          x3

. frget x?, from(group) prefix(g)
(3 variables copied from linked frame)
```

Variables in linked frames

- ▶ `frget` copies values and associated metadata into the current frame.
- ▶ Metadata includes display format, value label, variable label, and characteristics (including notes).

Example

```
. describe x* fx* gx*
```

Variable name	Storage type	Display format	Value label	Variable label
x1	float	%9.0g		Coffee expenses last month
x2	float	%9.0g		Rating on that new movie
x3	byte	%9.0g	valid	Loves to eat Tex-Mex food
fx1	int	%9.0g		Last years vacation expenses
fx2	byte	%10.0g	rating	Overall opinion of internet service
fx3	float	%9.0g		Some family-wise measurement
gx1	float	%9.0g		Annual dues
gx2	byte	%9.0g	noyes	Supports local construction project
gx3	float	%9.0g		Some group-wise measurement

Variables in linked frames

New variables from `frget`

▶ Use them as you would any other variable in Stata.

▶ Changing their values or metadata in the current frame will not be reflected in the linked frames.

Example

```
. summarize *x?, separator(3)
```

Variable	Obs	Mean	Std. dev.	Min	Max
x1	653	12.52833	3.58515	6.25	18.75
x2	653	5.049005	1.434827	2.5	7.5
x3	653	.4793262	.4999554	0	1
fx1	627	2970.11	1184.664	1007	4992
fx2	627	1.958533	.7882137	1	3
fx3	627	1.473844	.3132613	1	2
gx1	653	23466.28	10269.87	11804	43436
gx2	653	.5068913	.5003358	0	1
gx3	653	1.497703	.298736	1	1.9

```
. drop fx? gx?
```

Alias variables

Alias variables

► An alias is a reference to another variable, usually in a different frame.

► Alias variables behave like regular variables, but you cannot modify their values (observations).

► Use `fralias add` to create aliases for variables in a linked frame.

Example

```
. fralias add x?, from(family) prefix(f)
(3 variables aliased from linked frame)

. fralias add x?, from(group) prefix(g)
(3 variables aliased from linked frame)
```

Alias variables

- ▶ Alias variables get their own copy of the metadata, but their observations remain with the linked frame.
- ▶ Their storage type is a reference to the target variable's type in the linked frame.

Example

```
. describe *x?
```

Variable name	Storage type	Display format	Value label	Variable label
x1	float	%9.0g		Coffee expenses last month
x2	float	%9.0g		Rating on that new movie
x3	byte	%9.0g	valid	Loves to eat Tex-Mex food
fx1	*int*	%9.0g		Last years vacation expenses
fx2	*byte*	%10.0g	rating	Overall opinion of internet service
fx3	*float*	%9.0g		Some family-wise measurement
gx1	*float*	%9.0g		Annual dues
gx2	*byte*	%9.0g	noyes	Supports local construction project
gx3	*float*	%9.0g		Some group-wise measurement

Alias variables

Describing aliases

Use `fralias describe` to show information about alias variables in the current frame.

Example

```
. fralias describe
```

Alias	Type	Target	Link	Frame
fx1	int	x1	family	family
fx2	byte	x2	family	family
fx3	float	x3	family	family
gx1	float	x1	group	group
gx2	byte	x2	group	group
gx3	float	x3	group	group

Alias variables

- ▶ Use them as you would any other variable in Stata.
- ▶ Changing their metadata in the current frame will not be reflected in the linked frames.
- ▶ You cannot change their values in the current frame.

Example

```
. summarize *x?, separator(3)
```

Variable	Obs	Mean	Std. dev.	Min	Max
x1	653	12.52833	3.58515	6.25	18.75
x2	653	5.049005	1.434827	2.5	7.5
x3	653	.4793262	.4999554	0	1
fx1	627	2970.11	1184.664	1007	4992
fx2	627	1.958533	.7882137	1	3
fx3	627	1.473844	.3132613	1	2
gx1	653	23466.28	10269.87	11804	43436
gx2	653	.5068913	.5003358	0	1
gx3	653	1.497703	.298736	1	1.9

Alias variables

You can save frames with alias variables.

Example

```
. frames save linked, frames(person) linked replace
file linked.dtas saved

. frames describe using linked, simple
```
───
```
Frame: person
family_id  x2        y2        fx1       gx1
group_id   x3        family    fx2       gx2
x1         y1        group     fx3       gx3
```
───

───
```
Frame: family
family_id  x1        x2        x3
```
───

───
```
Frame: group
group_id   x1        x2        x3
```
───

Summary

- ► Frames allow you to work with multiple datasets in memory.
- ► Linking frames allows you to work with data collected at different grouping levels.
- ► You can save a set of frames to a single file, and have Stata automatically include linked frames in the file.
- ► You can copy data into frames, and make copies of variables from linked frames.
- ► New alias variables provide a memory efficient way to use data from linked frames in the current frame.