

Chapter 12 Stata v10.1 Analysis Examples Syntax and Output

General Notes on Stata 10.1

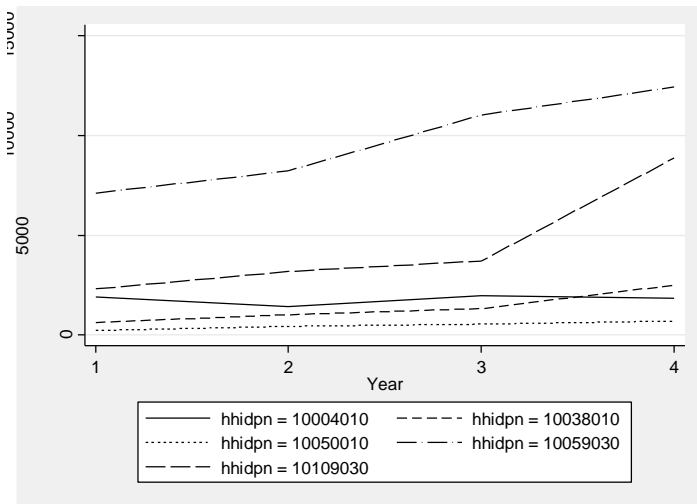
Given that this tool is used throughout the ASDA textbook this chapter includes only the syntax and output for the analysis examples provided in Chapter 12. Stata 10.1 is an excellent tool for survey data analysis as well as graphing and related data management tasks. It offers a very comprehensive set of svy commands as well as weighted graphics and convenient syntax and data management abilities. For these reasons, we use Stata as the primary software for the ASDA text.

The examples and syntax presented here assume that all data management including variable construction, labels for variable values and other preparation steps are complete. See the Stata documentation for assistance with these issues.

All analysis examples presented can be done in Stata 10.1 and are included in this chapter's output.

Please check the Stata documentation and also the ASDA web site for updates to Stata as new versions are released. For example, we have already included an example of how to use Stata 11.0 with the new "factor" variable features/syntax and compared this to the older "xi" type of syntax for including categorical variables in data analysis.

```
. xtline totassets000 if hhidpn <=10200000 & hhidpn != 10003030, ///
overlay ytitle(Total Assets (Thousands of Dollars)) ttitle(Year)
```



```
. *list a few data records to show data structure
. list hhidpn stratum secu baseweight weight year totalassets in 5/20
```

	hhidpn	stratum	secu	baseweight	weight	year	totalassets
5.	10299010	1	1	5633	5338	3	83000
6.	10299010	1	1	5633	4574	1	125000
7.	10299010	1	1	5633	8461	4	122000
8.	10299010	1	1	5633	5021	2	116000
9.	10394010	1	1	4166	4696	2	727000
10.	10394010	1	1	4166	5269	4	641000
11.	10394010	1	1	4166	4966	3	631000
12.	10394010	1	1	4166	4599	1	645000
13.	10395010	1	1	4162	4744	2	5797050
14.	10395010	1	1	4162	4564	1	5256000
15.	10395010	1	1	4162	5478	4	4698000
16.	10395010	1	1	4162	5316	3	2170000
17.	10397010	1	1	4166	4599	1	695000
18.	10397010	1	1	4166	5269	4	1221000
19.	10397010	1	1	4166	4696	2	507000
20.	10397010	1	1	4166	4966	3	1000000

```
* generate variables needed for analysis
gen yrssince00=0 if year==1
replace yrssince00=2 if year==2
replace yrssince00=4 if year==3
replace yrssince00=6 if year==4
tab yrssince00
```

```
* rescale total assets for easier understanding
gen totassets000 = totalassets / 1000
keep if edcat==4
```

```
* panel data plot for a few HH
xtset hhidpn year
*rescale weights using method 1 of rabe-hesketh and skrondal
gen sqw=llweight^2
egen sumsqw = sum (sqw), by (hhidpn)
egen sumw = sum(llweight), by (hhidpn)
gen llweight_r = llweight* sumw/sumsqw
```

```
* create a new variable for the gllamm command (unique id for stratum and secu)
gen newsecu = stratum *100 + secu
```

```
* generate another set of level 1 and 2 weights and run gllamm
gen pwt2 = baseweight
gen pwt1 = llweight_r
```

```
. xi: gllamm totassets000 yrssince00 , i(hhidpn) pweight(pwt) adapt cluster(newsecu)
```

```
Running adaptive quadrature
```

```
Iteration 0: log likelihood = -1.104e+08
Iteration 1: log likelihood = -1.099e+08
Iteration 2: log likelihood = -1.092e+08
Iteration 3: log likelihood = -1.091e+08
Iteration 4: log likelihood = -1.091e+08
Iteration 5: log likelihood = -1.091e+08
```

```
Adaptive quadrature has converged, running Newton-Raphson
```

```
Iteration 0: log likelihood = -1.091e+08
Iteration 1: log likelihood = -1.091e+08 (backed up)
Iteration 2: log likelihood = -1.091e+08
Iteration 3: log likelihood = -1.091e+08
```

```
number of level 1 units = 4350
number of level 2 units = 1112
```

```
Condition Number = 6410.0376
```

```
gllamm model
```

```
log likelihood = -1.091e+08
```

```
Robust standard errors for clustered data: cluster(newsecu)
```

totassets000	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
yrssince00	81.73993	13.04773	6.26	0.000	56.16684	107.313
_cons	741.527	78.63583	9.43	0.000	587.4036	895.6503

```
Variance at level 1
```

```
4402531 (1464316.4)
```

```
Variances and covariances of random effects
```

```
***level 2 (hhidpn)
```

```
var(1): 3510574.4 (1524032.8)
```

```
* run analogous model in svy regress
. gen wgt1_2 = pwt1*pwt2
(2516 missing values generated)
```

```
. svyset newsecu [pweight=wgt1_2]
```

```
    pweight: wgt1_2
          VCE: linearized
Single unit: missing
Strata 1: <one>
  SU 1: newsecu
  FPC 1: <zero>
```

```
. svy: regress totassets000 yrssince00
(running regress on estimation sample)
```

Survey: Linear regression

Number of strata	=	1	Number of obs	=	4447
Number of PSUs	=	104	Population size	=	11793472
			Design df	=	103
			F(1, 103)	=	39.74
			Prob > F	=	0.0000
			R-squared	=	0.0043

	Coef.	Linearized Std. Err.	t	P> t	[95% Conf. Interval]	
totassets000						
yrssince00	83.84876	13.30053	6.30	0.000	57.4703	110.2272
_cons	744.0301	79.56743	9.35	0.000	586.2269	901.8333