Reproducible and automated reporting using Stata

Kristin MacDonald

Director of Statistical Services
StataCorp LLC

2019 Nordic and Baltic Stata Users Group meeting
Stata’s commands for report generation allow you to create complete documents that include formatted text, summary statistics, regression results, and graphs in any of the following formats:

- Word\textsuperscript{®}
- Excel\textsuperscript{®}
- PDF
- HTML
Stata’s commands for creating reports come in two varieties:

1. **Dynamic document commands**—`dyntext` and `dyndoc`
   These commands create text files, HTML files, and Word documents that incorporate the full output from Stata commands. You can use the Markdown text-formatting language to customize the look of your report.

2. **`put*` commands**—`putdocx`, `putpdf`, and `putexcel`
   These commands create Word documents, PDFs, and Excel files that insert results from Stata commands into formatted text and tables in your document.

Whether you choose the dynamic document commands or the `put*` commands, you can create documents that are reproducible and dynamic.
Reproducible

- Reproducibility is at the core of Stata’s design.
- Use the **version 16** command, and any commands you run today will produce the same results many years from now.
- Use the **datasignature** command, you can verify that your data have not changed.
- Incorporate these tools for reproducibility into the do-files or text files that create your reports, and those reports will also be reproducible. Rerun your commands at any time and re-create your report.
Dynamic

- Update reports as data changes.
- Say you need to run the same report monthly. Rerun the commands that created the report with the updated dataset.
- All Stata results in the report are updated automatically.
A whirlwind tour:

- Word documents with `putdocx`
- PDFs with `putpdf`
- Excel files with `putexcel`
- HTML documents with `dyndoc`
- Word documents with `dyndoc`
- Converting files from one type to another
Content for our report:

- We have fictional data on the long-term care choices of 151 patients.
- Each patient selected a nursing home, in-home care, or an assisted-living facility.
- We fit a conditional logistic regression model to model the choices using `cmclogit`.
- We include two case-specific covariates: marital status and a five-category health status rating.
- We also include two alternative-specific covariates: monthly cost and distance from nearest relative.
. use ltcare, clear
. cmset id carealt

    caseid variable:  id
    alternatives variable:  carealt

. cmtab, choice(chosen)

Tabulation of chosen alternatives (chosen = 1)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursing home</td>
<td>58</td>
<td>38.41</td>
<td>38.41</td>
</tr>
<tr>
<td>In-home care</td>
<td>63</td>
<td>41.72</td>
<td>80.13</td>
</tr>
<tr>
<td>Assisted living</td>
<td>30</td>
<td>19.87</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>151</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>
. cmclogit chosen cost reldist, casevars(i.married i.hlthstat)

Iteration 0:  log likelihood = -119.36634
Iteration 1:  log likelihood = -87.257826
Iteration 2:  log likelihood = -80.030792
Iteration 3:  log likelihood = -79.791309
Iteration 4:  log likelihood = -79.790189
Iteration 5:  log likelihood = -79.790189

Conditional logit choice model
Case ID variable: id
Alternatives variable: carealt

Number of obs = 453
Number of cases = 151
Alts per case: min = 3
avg = 3.0
max = 3

Wald chi2(12) = 52.13
Prob > chi2 = 0.0000

Log likelihood = -79.790189

| chosen  | Coef.    | Std. Err. | z      | P>|z|  | [95% Conf. Interval] |
|---------|----------|-----------|--------|------|----------------------|
| carealt |          |           |        |      |                      |
| cost    | -.8927349| .1747136  | -5.11  | 0.000| -1.235167 - .5503025 |
| reldist | -.005768 | .0036172  | -1.59  | 0.111| -.0128575 .0013215  |
## Reproducible and automated reporting

<table>
<thead>
<tr>
<th>Nursing_home</th>
<th>(base alternative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In_home_care</td>
<td></td>
</tr>
<tr>
<td>married</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>3.682328</td>
</tr>
<tr>
<td>hlthstat</td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>0.7347672</td>
</tr>
<tr>
<td>Fair</td>
<td>4.894007</td>
</tr>
<tr>
<td>Average</td>
<td>6.52018</td>
</tr>
<tr>
<td>Good</td>
<td>4.800188</td>
</tr>
<tr>
<td>_cons</td>
<td>-3.44643</td>
</tr>
</tbody>
</table>

| Assisted_living |                    |
| married        |                    |
| Married        | 3.822477          |
| hlthstat       |                    |
| Poor           | -1.573831         |
| Fair           | 4.180285          |
| Average        | 6.179002          |
| Good           | 4.689159          |
| _cons          | -4.314829         |
We want to evaluate what would happen if cost of care in nursing homes increases by 15%.

We use **margins** to estimate changes in the expected proportions of individuals choosing each long-term care option.
. margins, alternative(1) at(cost=generate(cost)) at(cost=generate(cost*1.15))

Predictive margins
Model VCE : OIM
Expression : Pr(carealt|1 selected), predict()
Alternative : Nursing home

<table>
<thead>
<tr>
<th>1._at</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cost</td>
<td>= cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2._at</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cost</td>
<td>= cost*1.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Delta-method | Margin        | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|--------------|---------------|-----------|------|------|----------------------|
| _outcome#_at | Nursing home#1 | .384106   | .021637 | 17.75 | 0.000 | .3416982 .4265137 |
|              | Nursing home#2 | .3320989  | .0254219 | 13.06 | 0.000 | .2822728 .381925 |
|              | In-home care#1 | .4172185  | .030231 | 13.80 | 0.000 | .3579668 .4764702 |
|              | In-home care#2 | .4518933  | .0332705 | 13.58 | 0.000 | .3866843 .5171024 |
|              | Assisted living#1 | .1986755 | .0265481 | 7.48  | 0.000 | .1466422 .2507088 |
|              | Assisted living#2 | .2160078 | .0291166 | 7.42  | 0.000 | .1589404 .2730752 |
. marginsplot, xdim(_outcome) ///
> legend(order(1 "Current cost" 2 "15% increase") cols(1) ring(0) ///
> bplacement(neast)) ylabel(.1(.1).6) ///
> title("Effect of 15% increase in nursing-home cost") ///
> ytitle(Expected probability) xtitle(" ") noci graphregion(margin(r=10))

Variables that uniquely identify margins: _atopt _outcome
Multiple at() options specified:
    _atoption=1: cost=generate(cost)
    _atoption=2: cost=generate(cost*1.15)

. graph export cost.png, replace
(file cost.png written in PNG format)
Effect of 15% increase in nursing–home cost

Current cost

15% increase

Expected probability

Nursing home

In–home care

Assisted living
Reproducible and automated reporting

```
. margins, alternative(1) at(cost=generate(cost)) at(cost=generate(cost*1.15)) ///
> contrast(at(r) nowald)
```

```
Contrasts of predictive margins

Number of obs = 453

Model VCE : OIM
Expression : Pr(carealt|1 selected), predict()
Alternative : Nursing home

1._at : cost = cost
2._at : cost = cost*1.15
```

<table>
<thead>
<tr>
<th>_at@_outcome</th>
<th>Contrast</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2 vs 1) Nursing home</td>
<td>-.0520071</td>
<td>.0107312</td>
<td>-.0730398 -.0309743</td>
</tr>
<tr>
<td>(2 vs 1) In-home care</td>
<td>.0346748</td>
<td>.0075127</td>
<td>.0199501 .0493994</td>
</tr>
<tr>
<td>(2 vs 1) Assisted living</td>
<td>.0173323</td>
<td>.0067909</td>
<td>.0040223 .0306423</td>
</tr>
</tbody>
</table>
putdocx
Basic `putdocx` structure

```
putdocx begin

putdocx save myfile.docx
```
Basic `putdocx` structure

```bash
putdocx begin

putdocx paragraph
putdocx text = ("mytext")

putdocx save myfile.docx
```
Basic `putdocx` structure

```
putdocx begin

putdocx paragraph
putdocx text = ("mytext")

putdocx textblock begin
    mytext
    mytext
putdocx textblock end

putdocx save myfile.docx
```
Basic `putdocx` structure

```plaintext
putdocx begin

putdocx paragraph
putdocx text = ("mytext")

putdocx textblock begin
   mytext
putdocx textblock end

putdocx paragraph
putdocx image myimage.png

putdocx save myfile.docx
```
Basic `putdocx` structure

```plaintext
putdocx begin

putdocx paragraph
putdocx text = ("mytext")

putdocx textblock begin
    mytext
putdocx textblock end

putdocx paragraph
putdocx image myimage.png

putdocx table tablename = etable

putdocx save myfile.docx
```
• We can start with this basic structure and add to it to create a report with the results from our choice model.

• We begin our document and add a header with the Stata logo. We also put page numbers in the footer.

version 16
putdocx begin, header(head1) footer(foot1)

putdocx paragraph, toheader(head1) halign(right)
putdocx image stata16logo.png, height(.2in)

putdocx paragraph, tofooter(foot1)
putdocx pagenumber
We add a heading and an introduction to our report.

---

putdocx paragraph, style(Heading1)
putdocx text ("Introduction")

putdocx textblock begin
We are interested in studying the determinants of long-term care choices. We are also concerned that the cost of nursing home care is likely to increase by approximately 15%, and we want to know the expected impact of this change on the probability of individuals selecting each type of long-term care available.
putdocx textblock end
We add a description of our data.

```
use ltcare, clear
cmset id carealt
local ncase = r(n_cases)
total chosen, over(carealt)

putdocx textblock begin
In our sample of <<dd_docx_display: ‘ncase’>> patients,
<<dd_docx_display: _b[c.chosen@1bn.carealt]>>
selected a nursing home,
<<dd_docx_display: _b[c.chosen@2.carealt]>>
selected in-home care, and
<dd_docx_display: _b[c.chosen@3.carealt]>> selected an
assisted-living facility.
putdocx textblock end
```
We describe our model and add a table of the results.

We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status.

The results are presented in Table 1:

```
putdocx paragraph, style(Heading1)
putdocx text ("Model")

putdocx textblock begin
We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status.

The results are presented in Table 1:
putdocx textblock end

cmclogit chosen cost reldist, casevars(i.married i.hlthstat)
cmclogit, cformat(%5.3f) pformat(%4.2f) sformat(%5.3f)
putdocx table cmcl = etable, title("Table 1")
```
We present the results of our **margins** command graphically.

```
putdocx paragraph, style(Heading1)
putdocx text ("Effect of increasing cost of care in nursing homes")

margins, alternative(1) at(cost=generate(cost)) ///
at(cost=generate(cost*1.15)) ///

marginsplot, xdim(_outcome) ///
legend(order(1 "Current cost" 2 "15% increase") ///
cols(1) ring(0) bplacement(neast)) ylabel(.1(.1).6) ///
title("Effect of 15% increase in nursing-home cost") ///
ytitle(Expected probability) xtitle(" ") noci ///
graphregion(margin(r=10)))

graph export cost.png, replace

putdocx paragraph, halign(center)
putdocx image cost.png, height(3in)
```
We summarize the expected changes in probabilities when for increased nursing-home prices.

Finally, we save our document.

```
margins, alternative(1) at(cost=generate(cost)) ///
    at(cost=generate(cost*1.15)) ///
contrast(at(r) nowald)
```

```
putdocx textblock begin
If nursing home costs increase by 15%, we anticipate a percentage point decrease in the number of individuals selecting long-term care in nursing homes. In turn, we expect the number of individuals selecting in-home care to increase by and the number of individuals selecting assisted-living facilities to increase by percentage points.
putdocx textblock end
```

```
putdocx save myrpt, replace
```
Introduction

We are interested in studying the determinants of long-term care choices. We are also concerned that the cost of nursing home care is likely to increase by approximately 15%, and we want to know the expected impact of this change on the probability of individuals selecting each type of long-term care available.

In our sample of 153 patients, 58 selected a nursing home, 65 selected in-home care, and 30 selected an assisted-living facility.

Model

We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status. The results are presented in Table 1:

| Table 1 | chool | Cost | Std. Err. | z | P>|z| | [95% Conf. Interval] |
|---------|-------|------|-----------|---|------|-------------------|
| overall | cost  | -0.893 | 0.175 | -5.110 | 0.00 | -1.235 | -0.550 |
|         | value | -0.006 | 0.004 | -1.595 | 0.11 | -0.015 | 0.001 |

Nursing Home

<table>
<thead>
<tr>
<th>In-home care</th>
<th>Married</th>
<th>3.682</th>
<th>0.795</th>
<th>4.634</th>
<th>0.00</th>
<th>2.125</th>
<th>5.240</th>
</tr>
</thead>
</table>

| Health | Poor | 0.735 | 1.346 | 0.546 | 0.59 | -1.904 | 3.373 |
|        | Fair | 4.894 | 1.462 | 3.347 | 0.00 | 2.028 | 7.760 |
|        | Average | 6.520 | 1.752 | 3.722 | 0.00 | 3.087 | 9.954 |
|        | Good | 4.800 | 1.788 | 2.696 | 0.01 | 1.310 | 8.290 |
|        | care | -3.446 | 3.133 | -2.624 | 0.01 | -6.020 | -0.872 |

| Assistant living | Married | 3.822 | 0.877 | 4.300 | 0.00 | 2.104 | 5.543 |
| Health | Poor | 1.574 | 1.603 | -0.982 | 0.33 | -4.716 | 3.869 |
|        | Fair | 4.180 | 1.424 | 2.936 | 0.00 | 1.389 | 6.971 |
|        | Average | 6.179 | 1.708 | 3.618 | 0.00 | 2.382 | 9.526 |
|        | Good | 4.689 | 1.736 | 2.717 | 0.01 | 1.306 | 8.072 |
|        | care | -4.215 | 1.298 | -3.224 | 0.00 | -6.859 | -1.771 |
Effect of increasing cost of care in nursing homes

If nursing home costs increase by 15%, we anticipate a 5.2 percentage point decrease in the number of individuals selecting long-term care in nursing homes. In turn, we expect the number of individuals selecting in-home care to increase by 3.5% and the number of individuals selecting assisted living facilities to increase by 1.7 percentage points.
putpdf
Basic `putpdf` structure

```stata
putpdf begin

putpdf save myfile.pdf
```
Basic `putpdf` structure

```plaintext
putpdf begin

putpdf paragraph
putpdf text = ("mytext")

putpdf save myfile.pdf
```

Basic `putpdf` structure

```plaintext
putpdf begin

putpdf paragraph
putpdf text = ("mytext")

putpdf paragraph
putpdf image myimage.png

putpdf save myfile.pdf
```
Basic `putpdf` structure

```plaintext
putpdf begin

putpdf paragraph
putpdf text = ("mytext")

putpdf paragraph
putpdf image myimage.png

putpdf table tablenam = etable

putpdf save myfile.pdf
```
version 16
putpdf clear
putpdf begin

putpdf paragraph, font(, 16)
putpdf text ("Introduction")

putpdf paragraph
#delimit ;
putpdf text ("We are interested in studying the determinants
of long-term care choices. We are also concerned that the cost
of nursing home care is likely to increase by approximately 15%,
and we want to know the expected impact of this change on the
probability of individuals selecting each type of long-term care
available.");
#delimit cr

use ltcare, clear
cmset id carealt
local ncase = r(n cases)
total chosen, over(carealt)

putpdf paragraph
putpdf text ("In our sample of ")
putpdf text ("ncase")
putpdf text (" patients, ")
putpdf text ("b[c.chosen@1bn.carealt]")
putpdf text (" selected a nursing home, ")
putpdf text ("b[c.chosen@2.carealt]")
putpdf text (" selected in-home care, and ")
putpdf text ("b[c.chosen@3.carealt]")
putpdf text (" selected an assisted-living facility.")
Reproducible and automated reporting

K. L. MacDonald (StataCorp) 30 August 2019
Reproducible and automated reporting

```stata
76    putpdf paragraph
77    putpdf text ("If nursing home costs increase by 15%, we anticipate a ")
78    putpdf text (":\$3.1f 100*r(b)\{1,1}\)
79    putpdf text (" percentage point decrease in the number of individuals ")
80    putpdf text ("selecting long-term care in nursing homes. In turn, we ")
81    putpdf text ("expect the number of individuals selecting 
82    putpdf text ("in-home ")
83    putpdf text ("care to increase by ")
84    putpdf text (":\$3.1f 100*r(b)\{1,2}\)
85    putpdf text (" and the number of individuals selecting ")
86    putpdf text ("assisted-living facilities to increase by ")
87    putpdf text (":\$3.1f 100*r(b)\{1,3}\)
88    putpdf text (" percentage points."))
```

```stata
putpdf save myrpt, replace
```
Introduction

We are interested in studying the determinants of long-term care choices. We are also concerned that the cost of nursing home care is likely to increase by approximately 15%, and we want to know the expected impact of this change on the probability of individuals selecting each type of long-term care available.

In our sample of 151 patients, 58 selected a nursing home, 63 selected in-home care, and 30 selected an assisted-living facility.

Model

We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status.

The results are presented in Table 1:

|       | Cost     | Std. Err | z      | P>|z| | 95% Conf. Interval |
|-------|----------|----------|--------|------|-------------------|
| frail | 0.893    | 0.175    | -5.110 | 0.00 | -1.235 to -0.550  |
| frail | 0.006    | 0.004    | -1.505 | 0.11 | -0.013 to 0.001   |
| Nursing_home | (base alternative) |
| In_home care | Married |
| Married | 3.682    | 0.795    | 4.634  | 0.00 | 2.125 to 5.240    |
| Poor   | 0.735    | 1.346    | 0.546  | 0.59 | -1.904 to 3.373   |
| Fair   | 4.694    | 1.452    | 3.347  | 0.00 | 2.028 to 7.700    |
| Average| 6.520    | 1.752    | 3.722  | 0.00 | 3.087 to 9.954    |
| Good   | 4.800    | 1.780    | 2.696  | 0.01 | 1.310 to 8.290    |
| Assisted living | Married |
| Married | 3.622    | 0.877    | 4.360  | 0.00 | 2.104 to 5.541    |
| Poor   | -1.574   | 1.603    | -0.982 | 0.33 | -4.716 to 1.569   |
| Fair   | 4.180    | 1.424    | 2.936  | 0.00 | 1.389 to 6.971    |
| Average| 6.179    | 1.708    | 3.618  | 0.00 | 2.832 to 9.526    |
| Good   | 4.889    | 1.726    | 2.717  | 0.01 | 1.306 to 8.072    |
| _cons  | -4.315   | 1.298    | -3.324 | 0.00 | -6.859 to -1.771  |
Effect of increasing cost of care in nursing homes

If nursing home costs increase by 15%, we anticipate a 5.2 percentage point decrease in the number of individuals selecting long-term care in nursing homes. In turn, we expect the number of individuals selecting in-home care to increase by 3.5 and the number of individuals selecting assisted-living facilities to increase by 1.7 percentage points.
putexcel
Basic `putexcel` structure

```stata
putexcel set myfile.xlsx
```
Basic `putexcel` structure

```stata
putexcel set myfile.xlsx
putexcel A1 = "mytext"
```
Basic **putexcel** structure

```plaintext
putexcel set myfile.xlsx
putexcel A1 = "mytext"
putexcel A2 = image(myimage.png)
```
Basic `putexcel` structure

```stata
putexcel set myfile.xlsx
putexcel A1 = "mytext"
putexcel A2 = image(myimage.png)
putexcel B10 = etable
```
Reproducible and automated reporting

K. L. MacDonald (StataCorp)  30 August 2019  45 / 74

```stata
version 16
putexcel clear
putexcel set myrpt, replace
putexcel A1 = "Long-term care choices", font(, 16)
use ltcare, clear
cmset id carealt
putexcel B3 = "Sample:", font(, 14)
putexcel B4 = "Nursing home"
putexcel B5 = "In-home care"
putexcel B6 = "Assisted living"
putexcel B7 = "Total"
putexcel B7:C7, border(bottom) fpattern(solid, lightgray)
putexcel B4:B7, border(left)
putexcel C4:C7, border(right)
putexcel B4:C4, border(top)
putexcel C7 = `r(n_cases)'
total chosen, over(carealt)
matrix b = e(b)'
putexcel C4 = matrix(b)
putexcel B10 = "Conditional logistic regression model", font(, 14)
cmclogit chosen cost reldist, casevars(i.married i.hlthstat)
cmclogit, cformat($5.3f) pformat($5.3f) sformat($5.3f)
putexcel B11 = etable
putexcel B11:B37, border(left)
putexcel H11:H37, border(right)
```
Reproducible and automated reporting

K. L. MacDonald (StataCorp) 30 August 2019 46 / 74

```stata
putexcel J3 = "Effect of increasing cost of care in nursing homes", font(14)
margins, alternative(1) at(cost=generate(cost)) ///
at(cost=generate(cost*1.15)) ///
marginsplot, xdim(_outcome) ///
legend(order(1 "Current cost" 2 "15% increase") ///
cols(1) ring(0) placement(neast) ylab(.1(.1).6) ///
title("Effect of 15% increase in nursing-home cost") ///
ytitle(Expected probability) xtitle(" ") noci ///
graphregion(margin(r=10))
graph export cost.png, replace height(300)
putexcel J5 = image(cost.png)
margins, alternative(1) at(cost=generate(cost)) ///
at(cost=generate(cost*1.15)) ///
contrast(at(r) nowald)
putexcel J22 = "Expected change in proportion of individuals selecting each"
putexcel J23 = "long-term care option when nursing-home costs increase 15%"
putexcel J25 = "Nursing home"
putexcel J26 = "In-home care"
putexcel J27 = "Assisted living"
putexcel K25 = `:display r(b)[1,1]`, nformat(percent_d2)
putexcel K26 = `:display r(b)[1,2]`, nformat(percent_d2)
putexcel K27 = `:display r(b)[1,3]`, nformat(percent_d2)
```

```
### Long-term care choices

<table>
<thead>
<tr>
<th>Sample</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursing home</td>
<td>58</td>
</tr>
<tr>
<td>In-home care</td>
<td>63</td>
</tr>
<tr>
<td>Assisted living</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td>151</td>
</tr>
</tbody>
</table>

### Conditional logistic regression model

| chosen          | Coef. | Std. Err. | z     | P>|z| [95% Interval] |
|-----------------|-------|-----------|-------|-----------------|
| carealt         |       |           |       |                 |
| cost            | 0.893 | 0.175     | -5.11 | 0.0000          |
| reldist         | -0.006| 0.004     | -1.595| 0.1110          |
| Nursing_home    |       |           |       |                 |
| In_home         |       |           |       |                 |
| married         |       |           |       |                 |
| Married         | 3.682 | 0.795     | 4.634 | 0.0000          |
| hitstat         |       |           |       |                 |
| Poor            | 0.735 | 1.346     | 0.546 | 0.5850          |
| Fair            | 4.894 | 1.462     | 3.347 | 0.0010          |
| Average         | 6.52  | 1.752     | 3.722 | 0.0000          |
| Good            | 4.8   | 1.785     | 2.656 | 0.0070          |
| Assisted        |       |           |       |                 |
| living          |       |           |       |                 |
| married         |       |           |       |                 |
| Married         | 3.822 | 0.877     | 4.36  | 0.0000          |
| hitstat         |       |           |       |                 |
| Poor            | -1.574| 1.603     | -0.982| 0.3260          |
| Fair            | 4.18  | 1.424     | 2.936 | 0.0030          |
| Average         | 6.179 | 1.708     | 3.618 | 0.0000          |
| Good            | 4.689 | 1.726     | 2.717 | 0.0070          |
| cons            |       |           |       |                 |
|                | -4.315| 1.298     | -3.324| 0.0010          |

Expected change in proportion of individuals selecting each long-term care option when nursing-home costs increase 15%:

- Nursing home: -5.20%
- In-home care: 3.47%
- Assisted living: 1.73%
dyndoc
Basic `dyndoc` structure

Markdown-formatted text

Title
====
Basic **dyndoc** structure

Markdown-formatted text

Title
=====

##Subtitle
Basic **dyndoc** structure

Markdown-formatted text

Title
=====

##Subtitle

text with **bolding**
Basic **dyndoc** structure

Markdown-formatted text

Title
======

##Subtitle

text with **bolding**

~~~

codeblock

~~~
Basic dyndoc structure

Dynamic tags

<<dd_do>>
  Stata commands
<<</dd_do>>
Basic **dyndoc** structure

**Dynamic tags**

```
<<dd_do>>
    Stata commands
<</dd_do>>

<<dd_display: expression>>
```
Basic dyndoc structure

Dynamic tags

<<dd_do>>
  Stata commands
<</dd_do>>

<<dd_display: expression>>

<<dd_graph:saving(mygraph.png) graphname(mygraph)>>
Reproducible and automated reporting

K. L. MacDonald (StataCorp)
```stata
```~
```stata
```~
```stata
```~
```stata
## Effect of increasing cost of care in nursing homes

```stata
```~
```stata
```~
```stata
```~
```stata
```~
```stata
```~
```stata
If nursing home costs increase by 15%, we anticipate a percentage point decrease in the number of individuals selecting long-term care in nursing homes. In turn, we expect the number of individuals selecting in-home care to increase by and the number of individuals selecting assisted-living facilities to increase by percentage points.
Convert the `myhtml.md` Markdown document to HTML.

```
. dyndoc myhtml.md
```
Long-term care choices

Introduction

We are interested in studying the determinants of long-term care choices. We are also concerned that the cost of nursing home care is likely to increase by approximately 15%, and we want to know the expected impact of this change on the probability of individuals selecting each type of long-term care available.

Our sample consists of 151 patients who made the following long-term care choices.

```
. ctab, choose(chosen)
```

Tabulation of chosen alternatives (chosen = 1)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursing home</td>
<td>58</td>
<td>38.41</td>
<td>38.41</td>
</tr>
<tr>
<td>In-home care</td>
<td>63</td>
<td>41.72</td>
<td>80.13</td>
</tr>
<tr>
<td>Assisted living</td>
<td>30</td>
<td>19.87</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>151</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Conditional logistic regression model

We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status.

```
. clogit chosen cost reldist, casevars(1.married 1.hlthstat) ///
> cformat($6.3f) pformat($6.3f) sformat($6.3f)
```

Iteration 0:  log likelihood =  -119.36634
Iteration 1:  log likelihood =  -107.59925
Iteration 2:  log likelihood =  -97.93902
Iteration 3:  log likelihood =  -97.79138
Iteration 4:  log likelihood =  -97.79018
Iteration 5:  log likelihood =  -97.79018

Conditional logit choice model

Number of obs = 453
Number of cases = 151

Alternatives variable: casevars
Alts per case: min = 3
            avg =  3.0
            max =  3

Log likelihood = -97.79018
Wald chi2(12)  = 52.13
Prob > chi2    = 0.0000

chosen | Coef. Std. Err.   z   P>|z|   [95% Conf. Interval]
-------|-------------------|---------|------|------------------------|
chosen  | cost              | -0.203  | 0.175| 1.180      | 0.000     | -0.540   | 0.134     |
        | reldist           | -0.006  | 0.004| 0.150      | 0.037     | -0.008   | 0.000     |
Reproducible and automated reporting

K. L. MacDonald (StataCorp) 30 August 2019

Effect of increasing cost of care in nursing homes

If nursing home costs increase by 15%, we anticipate a 5.2 percentage point decrease in the number of individuals selecting long-term care in nursing homes. In turn, we expect the number of individuals selecting in-home care to increase by 3.5 and the number of individuals selecting assisted-living facilities to increase by 1.7 percentage points.
Convert the **myhtml.md** Markdown document to a Word document.

```
    . dyndoc myhtml.md, docx
```

Or convert the HTML document directly to a Word document.

```
    . html2docx myhtml.html
```
Long-term care choices

Introduction

We are interested in studying the determinants of long-term care choices. We are also concerned that the cost of nursing home care is likely to increase by approximately 15%, and we want to know the expected impact of this change on the probability of individuals selecting each type of long-term care available.

Our sample consists of 151 patients who made the following long-term care choices.

. tabulate, choice(chosen)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursing home</td>
<td>58</td>
<td>38.41</td>
<td>38.41</td>
</tr>
<tr>
<td>In-home care</td>
<td>63</td>
<td>41.72</td>
<td>80.13</td>
</tr>
<tr>
<td>Assisted living</td>
<td>30</td>
<td>19.87</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>151</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Conditional logistic regression model

We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status.

. clogit choice cost marital, casevar(id, married) ///
   & gformat(%5.3f) groupvar(id) gformat(%5.3f) Iteration 0:
   & log likelihood = -119.36634 Iteration 1:
   & log likelihood = -87.257826 Iteration 2:
   & log likelihood = -80.030792 Iteration 3:
   & log likelihood = -79.793359 Iteration 4:
   & log likelihood = -79.790189 Iteration 5:
   & log likelihood = -79.790189

   Conditional logit choice model
   Number of obs = 453
   Case ID variable: id
   Number of cases = 151
   Alternatives variable: marital
   Alts per case: min = 3
   avg = 3.0
Incorporate HTML code and style sheets to change the look of the document.
Reproducible and automated reporting

K. L. MacDonald (StataCorp)  30 August 2019  64 / 74

---

Long-term care choices

## Introduction

We are interested in studying the determinants of long-term care choices. We are also concerned that the cost of nursing home care is likely to increase by approximately 15%, and we want to know the expected impact of this change on the probability of individuals selecting each type of long-term care available.

```stata
<<dd_do:quietly>>
use ltcare, clear
cmset id carealt
<</dd_do>>
```

Our sample consists of 151 patients who made the following long-term care choices.

```stata
<</dd_do>>
```

## Conditional logistic regression model

We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status.

```stata
<</dd_do>>
```
Reproducible and automated reporting

K. L. MacDonald (StataCorp) 30 August 2019 65 / 74
We save the Markdown file that includes the header file as the `myhtml2.md`. Then we can type

```
.dyndoc myhtml2.md
```

to convert it to an HTML file.
Long-term care choices

Introduction

We are interested in studying the determinants of long-term care choices. We are also concerned that the cost of nursing home care is likely to increase by approximately 15%, and we want to know the expected impact of this change on the probability of individuals selecting each type of long-term care available.

Our sample consists of 151 patients who made the following long-term care choices.

```
. cmtab, choice(chosen)
```

Tabulation of chosen alternatives (chosen = 1)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursing home</td>
<td>58</td>
<td>38.41</td>
<td>38.41</td>
</tr>
<tr>
<td>In-home care</td>
<td>63</td>
<td>41.72</td>
<td>80.13</td>
</tr>
<tr>
<td>Assisted living</td>
<td>30</td>
<td>19.87</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>151</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Conditional logistic regression model

We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status.

```
. cnclogit chosen cost reldist, casevars(i.married i.hlthstat) ///
   > cformat(%5.2f) pformat(%5.2f) sformat(%5.3f)
```

Iteration 0:  log likelihood = -119.36634
Iteration 1:  log likelihood = -119.257826
Iteration 2:  log likelihood = -119.077428
Iteration 3:  log likelihood = -118.940429
Iteration 4:  log likelihood = -118.847341
Iteration 5:  log likelihood = -118.790189
Effect of increasing cost of care in nursing homes

If nursing home costs increase by 15%, we anticipate a 5.2 percentage point decrease in the number of individuals selecting long-term care in nursing homes. In turn, we expect the number of individuals selecting in-home care to increase by 3.5 and the number of individuals selecting assisted-living facilities to increase by 1.7 percentage points.
The same style is applied to a the Word document that is created when we type

```
dyndoc myhtml2.md, docx
```
Long-term care choices

Introduction

We are interested in studying the determinants of long-term care choices. We are also concerned that the cost of nursing home care is likely to increase by approximately 20%, and we want to know the expected impact of this change on the probability of individuals selecting each type of long-term care available.

Our sample consists of 151 patients who made the following long-term care choices.

```
. tabstat choice(chosen)
Tabulation of chosen alternatives (chosen = 1)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursing home</td>
<td>58</td>
<td>38.41</td>
<td>38.41</td>
</tr>
<tr>
<td>In-home care</td>
<td>63</td>
<td>41.72</td>
<td>80.13</td>
</tr>
<tr>
<td>Assisted living</td>
<td>30</td>
<td>19.87</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>151</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>
```

Conditional logistic regression model

We fit a conditional logistic regression model with alternative-specific covariates monthly cost and distance from nearest care-giving relative. We also include case-specific covariates marital status and health status.

```
. logit choice cost riskdist marriage 1.male 1.birthnat //
> > pformat(%2.2f) pformat(%4.2f) pformat(%5.3f)

Iteration 0:  log likelihood = -119.66634
Iteration 1:  log likelihood = -87.257826
Iteration 2:  log likelihood = -85.030782
Iteration 3:  log likelihood = -79.791009
Iteration 4:  log likelihood = -79.790189
Iteration 5:  log likelihood = -79.790189

Conditional logit choice model
Number of obs = 153
Case ID variable: id
Number of cases = 151
Alternatives variable: careact
Alts per case: min = 3
                          avg = 3.0
A few edits make it easy to produce dynamic reports. For instance, change

```
<<dd_do:quietly>>
use ltcare, clear
cmset id carealt
<</dd_do>>
```

to

```
<<dd_do:quietly>>
use ‘1’, clear
cmset id carealt
<</dd_do>>
```

Save the file with this update as `mythmld.md`, then type

```
   . dyndoc mythmld.md ltcare_q3.dta
```

to create the same report using the `ltcare_q3.dta`
Add more flexibility by passing additional arguments. For instance, we can specify the percentage increase in nursing-home cost that we want to analyze as a second argument. We change each instance of

15%

within text to in our Markdown document to

<<dd_display:‘2’>>
We also change the `margins` and `marginsplot` to respect the percentage specified in the second argument.

```
<<dd_do: quietly>>
margins, alternative(1) at(cost=generate(cost)) ///
   at(cost=generate(cost*1.'2'))

marginsplot, xdim(_outcome) ///
   legend(order(1 "Current cost" 2 "'2’% increase") ///
   cols(1) ring(0) bplacement(neast)) ylabel(.1(.1).6) ///
   title("Effect of ‘2’% increase in nursing-home cost") ///
   ytitle(Expected probability) xtitle("") noci ///
   graphregion(margin(r=10))
<</dd_do>>
```

After we save the modified file as `myhtmlmd.md`, we can type

```
. dyndoc myhtmlmd.md ltcare_q3.dta 20, saving(q3ptc20.html) replace
```

to evaluate the effect of a 20% increase in the cost nursing-home care and create the HTML file `q3pct20.html` that reports the results.
Whirlwind tour of reporting features?
Tornado?
Cyclone?

Try `putdocx`, `putpdf`, `putexcel`, or `dyndoc` the next time you need to automate a report in Word, HTML, Excel, or PDF format.