

Intro 1 — Interpretation of choice models

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

Description

Choice models have a reputation for being difficult to interpret. The coefficients that we estimate when we fit a choice model rarely allow a researcher to directly test hypotheses of interest. From the sign of the coefficient, we can determine the direction of an effect, and in the cases of conditional logit and mixed logit models, we can compute odds ratios and relative-risk ratios. Beyond this, the coefficients are almost uninterpretable.

There is no need for concern, however. Stata has a unique command, `margins`, that can be used after any of the `cm` choice model commands to easily answer questions of real scientific interest. In this entry, we introduce you to `margins`, and we give you a preview of the types of inferences that you can make when you use `margins` with the results of your choice models. To stay focused, we use `cmlogit` to fit all models in this introduction. That is not a limitation. The inferences we make and the associated `margins` commands would be the same for any of the `cm` estimation commands.

Remarks and examples

stata.com

Remarks are presented under the following headings:

*[Interpretation of coefficients](#)**[Inferences from margins](#)**[Expected choice probabilities](#)**[Effects of a continuous covariate](#)**[Effects of a categorical covariate](#)**[Effects of an alternative-specific covariate](#)**[More inferences using margins](#)*

Interpretation of coefficients

We demonstrate with a transportation example. We have 210 individuals who choose a method of travel between two cities—by airplane, train, bus, or car. We also have information on the length of time each method will take, each individual's income, and the number of people traveling together.

```
. use https://www.stata-press.com/data/r16/travel
. generate time = travelttime+termtime
```

Above, we generated a `time` variable representing total travel time. We use `cmset` to specify that the `id` variable identifies the travelers and the `mode` variable records the possible methods of travel, called alternatives. Then, we fit a conditional logistic regression model where `choice` indicates the selected travel method. Of our covariates, `time` is the only one that varies across the four travel alternatives. `income` and `partysize` are listed in the `casevars()` option because they have only one value per individual (or case) rather than a value for each alternative.

```
. cmset id mode
      caseid variable: id
      alternatives variable: mode
```

```

. cmlogit choice time, casevars(income partysize)
Iteration 0:  log likelihood = -249.36629
Iteration 1:  log likelihood = -236.01608
Iteration 2:  log likelihood = -235.65162
Iteration 3:  log likelihood = -235.65065
Iteration 4:  log likelihood = -235.65065

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

Number of obs      =      840
Number of cases    =      210
Alts per case: min =        4
                  avg =       4.0
                  max =        4

Wald chi2(7)      =       71.14
Prob > chi2       =       0.0000

Log likelihood = -235.65065

```

choice	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
mode						
time	-.0041641	.0007588	-5.49	0.000	-.0056512	-.002677
air	(base alternative)					
train						
income	-.0613414	.0122637	-5.00	0.000	-.0853778	-.0373051
partysize	.4123606	.2406358	1.71	0.087	-.0592769	.883998
_cons	3.39349	.6579166	5.16	0.000	2.103997	4.682982
bus						
income	-.0363345	.0134318	-2.71	0.007	-.0626605	-.0100086
partysize	-.1370778	.3437092	-0.40	0.690	-.8107354	.5365798
_cons	2.919314	.7658496	3.81	0.000	1.418276	4.420351
car						
income	-.0096347	.0111377	-0.87	0.387	-.0314641	.0121947
partysize	.7350802	.2184636	3.36	0.001	.3068993	1.163261
_cons	.7471042	.6732971	1.11	0.267	-.5725338	2.066742

What can we determine from these results? The coefficient on `time` is negative, so the probability of choosing a method of travel decreases as the travel time increases. For the train alternative, the coefficient on `income` is negative. Because air travel is the base alternative, this negative coefficient tells us that as income increases, people are less likely to choose a train over an airplane. For the car alternative, the coefficient on `partysize` is positive. As party size increases, people are more likely to choose a car over an airplane.

Inferences from margins

The output from `cmlogit`, or from any other `cm` command, gives us a little information. However, it does not typically answer questions a researcher is truly interested in. In our example, for instance, we might want to know the following:

- What percentage of individuals do we expect to choose air travel?
- What is the effect of income? How does the probability of selecting car travel change as income increases from \$30,000 to \$40,000 or from \$40,000 to \$50,000? How does the probability of selecting train travel change?
- What if new security measures are implemented and wait times at the airport increase by 60 minutes? How does that affect the probability of choosing each method of travel?

We can answer each of these questions and many others using `margins`.

Expected choice probabilities

Let's start with the first question. We simply type `margins` without any options to obtain the average predicted probability of choosing each method of travel.

```
. margins
Predictive margins                Number of obs   =       840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
```

	Delta-method		z	P> z	[95% Conf. Interval]	
	Margin	Std. Err.				
<code>_outcome</code>						
air	.2761905	.0275268	10.03	0.000	.2222389	.330142
train	.3	.0284836	10.53	0.000	.2441731	.3558269
bus	.1428571	.0234186	6.10	0.000	.0969576	.1887567
car	.2809524	.028043	10.02	0.000	.2259891	.3359156

Based on this model and our random sample of travelers between the two cities, we expect 28% of individuals to travel by air. We also expect 30% to travel by train, 14% to travel by bus, and 28% to travel by car.

Effects of a continuous covariate

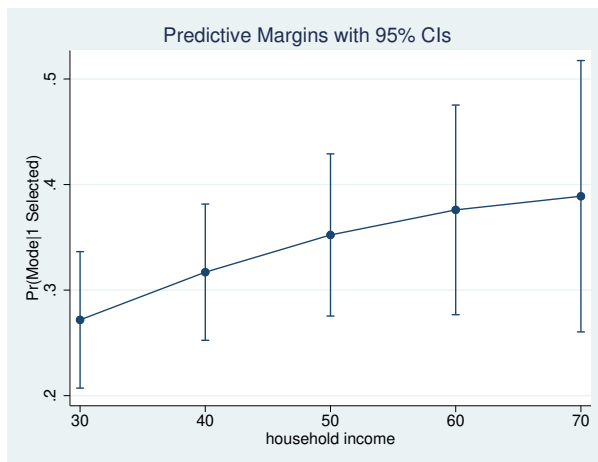
What does our model say would happen if values of a covariates change? We will first explore the effect of income. What would the expected probability of choosing car travel be if everyone made \$30,000? What if everyone made \$40,000? We can estimate expected probabilities of car travel for counterfactual income levels from \$30,000 to \$70,000 in \$10,000 increments by using the `at()` option with `margins`. We type

```
. margins, at(income=(30(10)70)) outcome(car)
Predictive margins                Number of obs   =       840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
Outcome      : car
1._at       : income           =       30
2._at       : income           =       40
3._at       : income           =       50
4._at       : income           =       60
5._at       : income           =       70
```

	Delta-method		z	P> z	[95% Conf. Interval]	
	Margin	Std. Err.				
<code>_at</code>						
1	.2717914	.0329811	8.24	0.000	.2071497	.3364331
2	.3169817	.0329227	9.63	0.000	.2524544	.3815091
3	.3522391	.0391994	8.99	0.000	.2754097	.4290684
4	.3760093	.050679	7.42	0.000	.2766802	.4753383
5	.3889296	.0655865	5.93	0.000	.2603825	.5174768

We can plot these probabilities to visualize the effect of income.

```
. marginsplot
Variables that uniquely identify margins: income
```



The expected probability of choosing car transportation increases as income increases. But are these differences statistically significant? We can test for differences in the expected probabilities for each \$10,000 increase in income. For this, we use the `contrast()` option and request reverse adjacent (`ar`) contrasts. We also simplify the output that reports a test of the differences by including the `nowald` and `effects` options.

```
. margins, at(income=(30(10)70)) outcome(car)
> contrast(atcontrast(ar) nowald effects)

Contrasts of predictive margins          Number of obs    =      840
Model VCE      : OIM
Expression     : Pr(mode|1 selected), predict()
Outcome       : car
1._at         : income           =      30
2._at         : income           =      40
3._at         : income           =      50
4._at         : income           =      60
5._at         : income           =      70
```

	Delta-method				
	Contrast	Std. Err.	z	P> z	[95% Conf. Interval]
2._at	.0451903	.016664	2.71	0.007	.0125296 .0778511
3._at	.0352574	.017903	1.97	0.049	.0001681 .0703466
4._at	.0237702	.0190387	1.25	0.212	-.013545 .0610854
5._at	.0129204	.0200549	0.64	0.519	-.0263866 .0522273

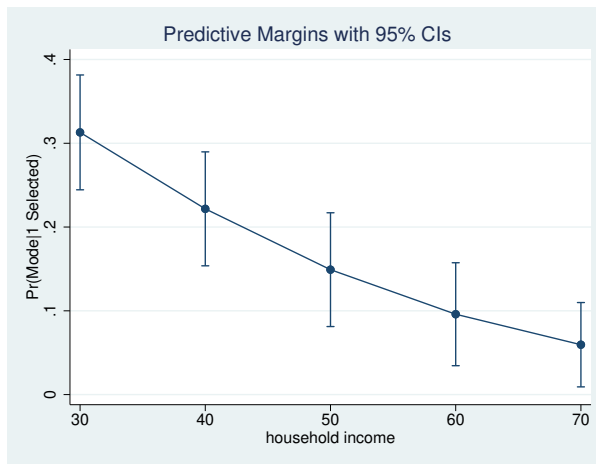
From the first line in this table, we see that the effect of having \$40,000 instead of \$30,000 is a 0.045 increase in the expected probability of selecting car travel. Having \$50,000 instead of \$40,000 increases the expected probability of car travel by 0.035. Both of these effects are significant at a 5% significance level, but increases in income from \$50,000 to \$60,000 and from \$60,000 to \$70,000 do not lead to significantly different probabilities of car travel.

We can also evaluate the effect of income on the probability of taking a train. We include the `outcome(train)` option instead of `outcome(car)` in our `margins` command.

```
. margins, at(income=(30(10)70)) outcome(train)
Predictive margins                                Number of obs      =          840
Model VCE    : OIM
Expression   : Pr(mode|1 selected), predict()
Outcome      : train
1._at       : income          =          30
2._at       : income          =          40
3._at       : income          =          50
4._at       : income          =          60
5._at       : income          =          70
```

	Delta-method			z	P> z	[95% Conf. Interval]	
	Margin	Std. Err.					
_at							
1	.3129898	.034968	8.95	0.000	.2444538	.3815257	
2	.2217291	.0346969	6.39	0.000	.1537244	.2897338	
3	.1491281	.0346442	4.30	0.000	.0812268	.2170294	
4	.0959391	.0313489	3.06	0.002	.0344965	.1573818	
5	.0595511	.0256786	2.32	0.020	.009222	.1098802	

```
. marginsplot
Variables that uniquely identify margins: income
```



As income levels increase, the expected probability of choosing train transportation decreases.

We can again test for differences in the probabilities using reverse adjacent contrasts.

```
. margins, at(income=(30(10)70)) outcome(train)
> contrast(atcontrast(ar) nowald effects)
```

```
Contrasts of predictive margins          Number of obs    =          840
Model VCE      : OIM
Expression     : Pr(mode|1 selected), predict()
Outcome        : train
1._at         : income           =           30
2._at         : income           =           40
3._at         : income           =           50
4._at         : income           =           60
5._at         : income           =           70
```

	Delta-method				[95% Conf. Interval]	
	Contrast	Std. Err.	z	P> z		
_at						
(2 vs 1)	-.0912606	.0174009	-5.24	0.000	-.1253659	-.0571554
(3 vs 2)	-.072601	.0104846	-6.92	0.000	-.0931505	-.0520516
(4 vs 3)	-.053189	.0066782	-7.96	0.000	-.0662779	-.0401
(5 vs 4)	-.036388	.0064958	-5.60	0.000	-.0491195	-.0236566

This time, we find a significant decrease in the expected probability of choosing train transportation for each \$10,000 increase in income.

We do not need to limit our analysis to just car or train travel. How does the probability of each method of travel change with income?

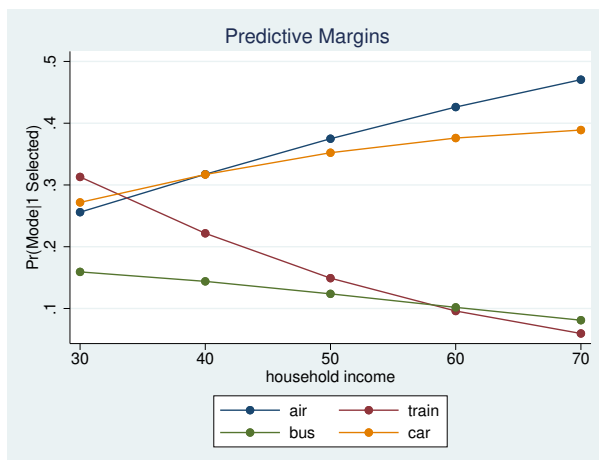
We can type

```
. margins, at(income=(30(10)70))
```

This produces lots of output, so we do not show it here. We instead show you the graph of all the results. We omit the confidence intervals so that it is easy to see the probabilities for all four methods of travel.

```
. marginsplot, noci
```

Variables that uniquely identify margins: income _outcome



As income goes up, the expected probabilities of selecting bus and train transportation decrease, and the expected probabilities of choosing air and car transportation increase.

This graph allows us to visually compare the travel methods at each income level. We can formally test for differences in the expected probabilities of the travel methods. For instance, at the \$30,000 income level, are all four methods of travel equally likely to be selected? We can use the `contrast(outcomejoint)` option to request this test.

```
. margins, at(income=30) contrast(outcomejoint)
Contrasts of predictive margins          Number of obs    =          840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
at            : income          =          30
```

	df	chi2	P>chi2
_outcome	3	12.80	0.0051

We find that at least one of these expected probabilities is significantly different from the others.

We might want to ask a more specific question. Is there a difference in the expected probabilities of selecting train and bus travel when income is \$50,000? We use the `outcome()` option to specify these two methods of travel and the `contrast(outcomecontrast(r))` option to request that `margins` estimate the difference between the two probabilities.

```
. margins, at(income=50) outcome(bus train)
> contrast(outcomecontrast(r) nowald effects)
Contrasts of predictive margins          Number of obs    =          840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
at            : income          =          50
```

	Delta-method				
	Contrast	Std. Err.	z	P> z	[95% Conf. Interval]
_outcome (bus vs train)	-.0254125	.0504996	-0.50	0.615	-.1243899 .0735648

We do not find a significant difference in the probabilities of selecting bus and train travel at this income level.

Effects of a categorical covariate

The questions we answered above were about a continuous variable, `income`. What if we want to evaluate the effect of a categorical variable instead? We include the variable as a factor variable in our model and again use `margins`. To demonstrate, we create a variable representing income quartiles and change our model to use it rather than the continuous income covariate.

```
. xtile income_cat = income, nquantiles(4)
. label define quartiles 1 "Quartile 1" 2 "Quartile 2"
>                 3 "Quartile 3" 4 "Quartile 4"
. label values income_cat quartiles
. cmlogit choice time, casevars(i.income_cat partysize)
(output omitted)
```

We estimate the expected probability of selecting train travel for each income quartile by typing

```
. margins income_cat, outcome(train)
```

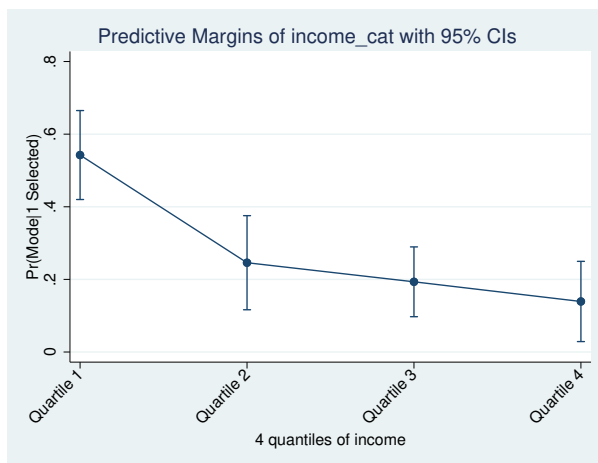
```
Predictive margins                    Number of obs    =          840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
Outcome       : train
```

	Delta-method				
	Margin	Std. Err.	z	P> z	[95% Conf. Interval]
income_cat					
Quartile 1	.5424554	.0624929	8.68	0.000	.4199715 .6649392
Quartile 2	.2459475	.0661098	3.72	0.000	.1163746 .3755203
Quartile 3	.1933788	.0490343	3.94	0.000	.0972733 .2894843
Quartile 4	.1391895	.0563599	2.47	0.014	.0287262 .2496529

We can again plot the results using marginsplot.

```
. marginsplot
```

```
Variables that uniquely identify margins: income_cat
```



We use the `ar.` operator to test for differences in the expected probabilities across adjacent income quartiles. This time, we attach the contrast operator to the name of the variable.

```
. margins ar.income_cat, outcome(train) contrast(nowald effects)
Contrasts of predictive margins          Number of obs    =          840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
Outcome       : train
```

	Delta-method				
	Contrast	Std. Err.	z	P> z	[95% Conf. Interval]
income_cat (Quartile 2 vs Quartile 1) (Quartile 3 vs Quartile 2) (Quartile 4 vs Quartile 3)	-.2965079	.0905087	-3.28	0.001	-.4739018 -.1191141
	-.0525687	.082642	-0.64	0.525	-.214544 .1094067
	-.0541893	.0744351	-0.73	0.467	-.2000794 .0917009

We find that the expected probability of choosing train travel is significantly different when moving from the first to the second income quartile.

Effects of an alternative-specific covariate

So far, we have explored the effects of the case-specific income variable. We can also ask questions about an alternative-specific variable such as travel time.

Perhaps new security measures are added. What would we expect if wait times at the airport increase by 60 minutes for all flights? We can use `margins'` `at()` option to evaluate this scenario.

```
. margins, at(time=generate(time+60)) alternative(air)
Predictive margins          Number of obs    =          840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
Alternative   : air
at            : time          = time+60
```

	Delta-method				
	Margin	Std. Err.	z	P> z	[95% Conf. Interval]
_outcome					
air	.238248	.0261139	9.12	0.000	.1870657 .2894303
train	.3138431	.0293829	10.68	0.000	.2562537 .3714325
bus	.1513295	.0243434	6.22	0.000	.1036173 .1990417
car	.2965794	.0289688	10.24	0.000	.2398016 .3533572

As we would anticipate, the expected probability of selecting air travel decreases when travel time increases. The probability of choosing air travel is now 0.24. If we look back at our first `margins` command, we can see that with the original travel times, the expected probability of choosing air travel was 0.28. Rather than looking at results from multiple commands, we can estimate probabilities with the original travel times and with the increased travel times all with a single `margins` command.

```

. margins, at(time=generate(time)) at(time=generate(time+60))
> alternative(air)
Predictive margins                                Number of obs   =       840
Model VCE    : OIM
Expression   : Pr(mode|1 selected), predict()
Alternative  : air
1._at       : time                               = time
2._at       : time                               = time+60

```

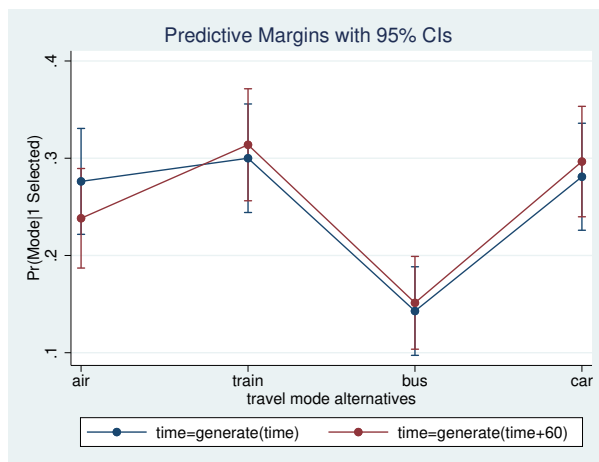
	Delta-method				[95% Conf. Interval]	
	Margin	Std. Err.	z	P> z		
_outcome#_at						
air#1	.2761905	.0277787	9.94	0.000	.2217453	.3306357
air#2	.238248	.0261139	9.12	0.000	.1870657	.2894303
train#1	.3	.0284715	10.54	0.000	.2441968	.3558032
train#2	.3138431	.0293829	10.68	0.000	.2562537	.3714325
bus#1	.1428571	.0232375	6.15	0.000	.0973125	.1884018
bus#2	.1513295	.0243434	6.22	0.000	.1036173	.1990417
car#1	.2809524	.0280657	10.01	0.000	.2259446	.3359602
car#2	.2965794	.0289688	10.24	0.000	.2398016	.3533572

Now it is easy to plot both scenarios together using `marginsplot`. This time we include the `xdimension()` option to place the four travel choices along the x axis.

```

. marginsplot, xdimension(_outcome)
Variables that uniquely identify margins: _atopt _outcome
Multiple at() options specified:
  _atoption=1: time=generate(time)
  _atoption=2: time=generate(time+60)

```



We see that the probability of air travel decreases, while the probability of choosing each of the other methods of travel increases a little.

We can again use contrasts to test for a difference. We use the `atcontrast(r)` to request comparisons to a reference level (the original travel times).

```
. margins, at(time=generate(time)) at(time=generate(time+60))
> alternative(air) contrast(atcontrast(r) nowald effects)

Contrasts of predictive margins          Number of obs      =          840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
Alternative   : air
1._at        : time              = time
2._at        : time              = time+60
```

	Delta-method				
	Contrast	Std. Err.	z	P> z	[95% Conf. Interval]
_at@_outcome					
(2 vs 1) air	-.0379425	.0056615	-6.70	0.000	-.0490387 -.0268462
(2 vs 1)					
train	.0138431	.0027216	5.09	0.000	.0085089 .0191773
(2 vs 1) bus	.0084724	.001936	4.38	0.000	.004678 .0122668
(2 vs 1) car	.015627	.0026002	6.01	0.000	.0105308 .0207232

Each of these differences is statistically significant. Of course, the researcher would need to decide whether an expected 3.8 percentage point decrease in air travel is meaningful.

We might take this even one step further. What if the speed of train transportation improves at the same time that speed of air travel declines? What if train travel is now 60 minutes faster while air travel is 60 minutes slower? We cannot specify all of these changes directly in the `at()` option, but we can create a new variable that represents this scenario.

```
. generate newtime = time
. replace newtime = time+60 if mode==1
(210 real changes made)
. replace newtime = time-60 if mode==2
(210 real changes made)
```

Now we can specify our `newtime` variable in the `at()` option. We also include the `alternative(simultaneous)` option to specify that the changes to air and train travel be made simultaneously.

```

. margins, at(time=generate(time)) at(time=generate(newtime))
> alternative(simultaneous)

Predictive margins                                Number of obs   =       840
Model VCE    : OIM
Expression   : Pr(mode|1 selected), predict()
1._at       : time                               = time
2._at       : time                               = newtime

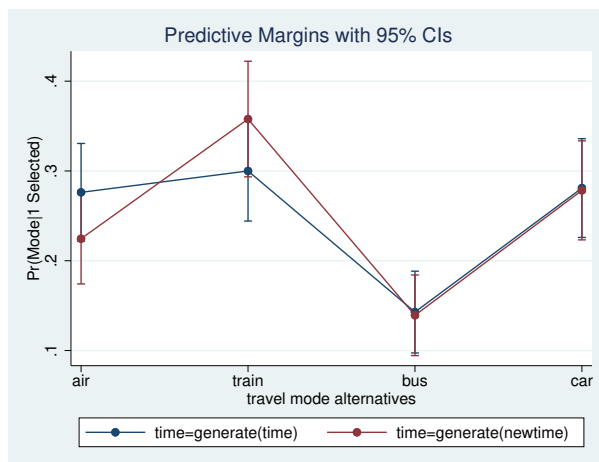
```

_outcome#_at	Delta-method			z	P> z	[95% Conf. Interval]	
	Margin	Std. Err.					
air#1	.2761905	.0277787	9.94	0.000	.2217453	.3306357	
air#2	.224539	.0256826	8.74	0.000	.1742021	.274876	
train#1	.3	.0284715	10.54	0.000	.2441968	.3558032	
train#2	.3578287	.0328487	10.89	0.000	.2934464	.4222111	
bus#1	.1428571	.0232375	6.15	0.000	.0973125	.1884018	
bus#2	.1392549	.0228972	6.08	0.000	.0943773	.1841325	
car#1	.2809524	.0280657	10.01	0.000	.2259446	.3359602	
car#2	.2783773	.0281348	9.89	0.000	.2232342	.3335205	

```

. marginsplot, xdimension(_outcome)
Variables that uniquely identify margins: _atopt _outcome
Multiple at() options specified:
  _atoption=1: time=generate(time)
  _atoption=2: time=generate(newtime)

```



Now it appears that the expected probability of selecting air travel decreases, the expected probability of selecting train travel increases, and the expected probabilities of selecting the other methods of transportation do not change much. Let's test for differences.

```

. margins, at(time=generate(time)) at(time=generate(newtime))
> alternative(simultaneous) contrast(atcontrast(r) nowald effects)
Contrasts of predictive margins          Number of obs      =          840
Model VCE      : OIM
Expression    : Pr(mode|1 selected), predict()
1._at         : time              = time
2._at         : time              = newtime

```

	Delta-method				[95% Conf. Interval]	
	Contrast	Std. Err.	z	P> z		
_at@_outcome						
(2 vs 1) air	-.0516514	.0080654	-6.40	0.000	-.0674594	-.0358435
(2 vs 1) train	.0578287	.0109404	5.29	0.000	.0363859	.0792716
(2 vs 1) bus	-.0036023	.0022959	-1.57	0.117	-.0081021	.0008976
(2 vs 1) car	-.002575	.0036587	-0.70	0.482	-.009746	.0045959

Now the expected probability of selecting air travel is 5.2 percentage points lower than it was with the original travel times, and the expected probability of selecting train travel is 5.7 percentage points higher.

More inferences using margins

Here we have demonstrated how you can use the results of your choice model to answer some interesting questions. But this is just a small sample of the types of inference that you can do using `margins` after a choice model estimator. For more examples, see [\[CM\] margins](#), [\[CM\] Intro 5](#), [\[CM\] Intro 6](#), [\[CM\] cmlogit](#), [\[CM\] cmmixlogit](#), and [\[CM\] cmxtmixlogit](#). Regardless of the `cm` command you use to fit your model, you may be interested in all of these examples because the same `margins` commands can be used after any of the models to estimate the same types of margins and test the same types of hypotheses.

Also see

[\[CM\] margins](#) — Adjusted predictions, predictive margins, and marginal effects