



17.0
MP-Parallel Edition

Statistics and Data Science

Copyright 1985-2021 StataCorp LLC
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC <https://www.stata.com>
979-696-4600 stata@stata.com

Stata license: Unlimited-user 64-core network perpetual
Serial number: 18461036
Licensed to: TEAM BTCR
TEAM BTCR

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5,000; see [help set maxvar](#).

```
1 . import excel "G:\Academic\Phd research\Novelty\Projects\Integrate\code\Analyse\robust\log-extra-control-variable\
(6 vars, 674 obs)
```

```
2 . desc
```

Contains data
Observations: **674**
Variables: **6**

Variable name	Storage type	Display format	Value label	Variable label
Doi	str28	%28s		Doi
Novelty	double	%10.0g		Novelty
Year	int	%10.0g		Year
label	byte	%10.0g		label
Author	int	%10.0g		Author
Aff	byte	%10.0g		Aff

Sorted by:

Note: Dataset has changed since last saved.

```
3 . logit label Novelty Year Author Aff
```

```
Iteration 0: log likelihood = -467.1812
Iteration 1: log likelihood = -453.5014
Iteration 2: log likelihood = -453.44979
Iteration 3: log likelihood = -453.44969
Iteration 4: log likelihood = -453.44969
```

Logistic regression

Number of obs = **674**
LR chi2(4) = **27.46**
Prob > chi2 = **0.0000**
Pseudo R2 = **0.0294**

Log likelihood = **-453.44969**

label	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
Novelty	-.1879272	.0511342	-3.68	0.000	-.2881483	-.087706
Year	.0028112	.0197305	0.14	0.887	-.0358599	.0414823
Author	.0592139	.0253209	2.34	0.019	.0095859	.1088419
Aff	-.123376	.037486	-3.29	0.001	-.1968471	-.0499048
_cons	-6.194325	39.70993	-0.16	0.876	-84.02437	71.63572

4 . logistic label Novelty Year Author Aff

Logistic regression

Number of obs = 674

LR chi2(4) = 27.46

Prob > chi2 = 0.0000

Pseudo R2 = 0.0294

Log likelihood = -453.44969

label	Odds ratio	Std. err.	z	P> z	[95% conf. interval]	
Novelty	.8286751	.0423736	-3.68	0.000	.7496504	.9160301
Year	1.002815	.0197861	0.14	0.887	.9647754	1.042355
Author	1.061002	.0268655	2.34	0.019	1.009632	1.114986
Aff	.8839313	.033135	-3.29	0.001	.8213162	.95132
_cons	.002041	.0810472	-0.16	0.876	3.23e-37	1.29e+31

Note: _cons estimates baseline odds.

5 . import excel "G:\Academic\Phd research\Novelty\Projects\Integrate\code\Analyse\robust\log-extra-control-variable\"
(6 vars, 678 obs)

6 . logit label Novelty Year Author Aff

Iteration 0: log likelihood = -469.95379

Iteration 1: log likelihood = -455.5436

Iteration 2: log likelihood = -453.31512

Iteration 3: log likelihood = -453.29699

Iteration 4: log likelihood = -453.29698

Logistic regression

Number of obs = 678

LR chi2(4) = 33.31

Prob > chi2 = 0.0000

Pseudo R2 = 0.0354

Log likelihood = -453.29698

label	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
Novelty	-.1127912	.0303542	-3.72	0.000	-.1722843	-.0532982
Year	.0184846	.0196708	0.94	0.347	-.0200695	.0570387
Author	.0712828	.0263403	2.71	0.007	.0196568	.1229088
Aff	-.1474805	.0395085	-3.73	0.000	-.2249157	-.0700452
_cons	-37.3296	39.59541	-0.94	0.346	-114.9352	40.27597

7 . logistic label Novelty Year Author Aff

Logistic regression

Number of obs = 678

LR chi2(4) = 33.31

Prob > chi2 = 0.0000

Pseudo R2 = 0.0354

Log likelihood = -453.29698

label	Odds ratio	Std. err.	z	P> z	[95% conf. interval]	
Novelty	.8933372	.0271165	-3.72	0.000	.8417399	.9480973
Year	1.018657	.0200378	0.94	0.347	.9801306	1.058697
Author	1.073885	.0282864	2.71	0.007	1.019851	1.130781
Aff	.8628793	.0340911	-3.73	0.000	.7985835	.9323516
_cons	6.14e-17	2.43e-15	-0.94	0.346	1.21e-50	3.10e+17

Note: _cons estimates baseline odds.

```
8 . import excel "G:\Academic\Phd research\Novelty\Projects\Integrate\code\Analyse\robust\log-extra-control-variable\
(6 vars, 3,800 obs)
```

```
9 . logit label Novelty Year Author Aff
```

```
Iteration 0: log likelihood = -1138.0352
Iteration 1: log likelihood = -1110.4516
Iteration 2: log likelihood = -1107.1949
Iteration 3: log likelihood = -1107.1776
Iteration 4: log likelihood = -1107.1776
```

Logistic regression

Number of obs = 3,800
LR chi2(4) = 61.72
Prob > chi2 = 0.0000
Pseudo R2 = 0.0271

Log likelihood = -1107.1776

label	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
Novelty	-.1719622	.0420652	-4.09	0.000	-.2544085	-.0895159
Year	-.0548038	.0136278	-4.02	0.000	-.0815138	-.0280938
Author	.0532249	.0176399	3.02	0.003	.0186513	.0877984
Aff	-.1160053	.0249682	-4.65	0.000	-.1649422	-.0670685
_cons	112.1653	27.42948	4.09	0.000	58.40448	165.9261

```
10 . logistic label Novelty Year Author Aff
```

Logistic regression

Number of obs = 3,800
LR chi2(4) = 61.72
Prob > chi2 = 0.0000
Pseudo R2 = 0.0271

Log likelihood = -1107.1776

label	Odds ratio	Std. err.	z	P> z	[95% conf. interval]	
Novelty	.842011	.0354194	-4.09	0.000	.775375	.9143737
Year	.9466709	.0129011	-4.02	0.000	.92172	.9722972
Author	1.054667	.0186042	3.02	0.003	1.018826	1.091768
Aff	.8904705	.0222335	-4.65	0.000	.8479427	.9351312
_cons	5.16e+48	1.42e+50	4.09	0.000	2.32e+25	1.15e+72

Note: _cons estimates baseline odds.

```
11 . import excel "G:\Academic\Phd research\Novelty\Projects\Integrate\code\Analyse\robust\log-extra-control-variable\
(6 vars, 650 obs)
```

```
12 . logit label Novelty Year Author Aff
```

```
Iteration 0: log likelihood = -450.54567
Iteration 1: log likelihood = -433.92827
Iteration 2: log likelihood = -433.84804
Iteration 3: log likelihood = -433.848
```

Logistic regression

Number of obs = 650
LR chi2(4) = 33.40
Prob > chi2 = 0.0000
Pseudo R2 = 0.0371

Log likelihood = -433.848

label	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
Novelty	-.2456592	.0585806	-4.19	0.000	-.3604749	-.1308434
Year	.0021586	.0205778	0.10	0.916	-.0381731	.0424903
Author	.0646336	.0247834	2.61	0.009	.016059	.1132082
Aff	-.1384142	.0382844	-3.62	0.000	-.2134502	-.0633782
_cons	-4.977578	41.40872	-0.12	0.904	-86.13717	76.18202

13 . logistic label Novelty Year Author Aff

Logistic regression

Number of obs = 650
 LR chi2(4) = 33.40
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0371

Log likelihood = -433.848

label	Odds ratio	Std. err.	z	P> z	[95% conf. interval]	
Novelty	.7821888	.0458211	-4.19	0.000	.697345	.8773552
Year	1.002161	.0206222	0.10	0.916	.9625463	1.043406
Author	1.066768	.0264381	2.61	0.009	1.016189	1.119865
Aff	.8707379	.0333357	-3.62	0.000	.8077924	.9385884
_cons	.0068907	.2853364	-0.12	0.904	3.90e-38	1.22e+33

Note: _cons estimates baseline odds.

14 . import excel "G:\Academic\Phd research\Novelty\Projects\Integrate\code\Analyse\robust\log-extra-control-variable\log-extra-control-variable.xlsx" (6 vars, 3,918 obs)

15 . logit label Novelty Year Author Aff

Iteration 0: log likelihood = -1153.539
 Iteration 1: log likelihood = -1119.3414
 Iteration 2: log likelihood = -1116.1994
 Iteration 3: log likelihood = -1116.1898
 Iteration 4: log likelihood = -1116.1898

Logistic regression

Number of obs = 3,918
 LR chi2(4) = 74.70
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0324

Log likelihood = -1116.1898

label	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
Novelty	-.1111869	.0226019	-4.92	0.000	-.1554857	-.0668881
Year	-.0457766	.0141883	-3.23	0.001	-.0735851	-.017968
Author	.0641053	.0188581	3.40	0.001	.0271441	.1010664
Aff	-.1195483	.024848	-4.81	0.000	-.1682495	-.0708471
_cons	94.31011	28.55851	3.30	0.001	38.33646	150.2838

16 . logistic label Novelty Year Author Aff

Logistic regression

Number of obs = 3,918
 LR chi2(4) = 74.70
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0324

Log likelihood = -1116.1898

label	Odds ratio	Std. err.	z	P> z	[95% conf. interval]	
Novelty	.8947715	.0202235	-4.92	0.000	.8559993	.9352999
Year	.9552554	.0135535	-3.23	0.001	.9290571	.9821925
Author	1.066205	.0201066	3.40	0.001	1.027516	1.10635
Aff	.8873211	.0220482	-4.81	0.000	.845143	.9316043
_cons	9.09e+40	2.59e+42	3.30	0.001	4.46e+16	1.85e+65

Note: _cons estimates baseline odds.

17 .