Rescaling results of nonlinear probability models to compare regression coefficients or variance components across hierarchically nested models

Dirk Enzmann & Ulrich Kohler

University of Hamburg, dirk.enzmann@uni-hamburg.de

Wissenschaftszentrum Berlin, kohler@wzb.eu

German Stata Users Group Meeting
Berlin, June 1, 2012

Title improved

Rescaling of fixed and random effects in hierarchically nested multilevel models

Dirk Enzmann & Ulrich Kohler

University of Hamburg, dirk.enzmann@uni-hamburg.de

Wissenschaftszentrum Berlin, kohler@wzb.eu

German Stata Users Group Meeting
Berlin, June 1, 2012

Outline

- The problem: Increase of fixed or random effects in nonlinear probability models
- A solution: Rescaling of fixed and random effects
- Example of implementation in Stata

The problem

- Adding a random intercept or variables with fixed effects to a logistic or probit model may increase effects of earlier included variables.

Logistic regression models for taking a science subject (Snijders & Bosker, 1999, p. 266 f.)

	Single Level		ML Model 1		ML Model 2		ML Model 3	
Fixed Effect	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
γ_0 Intercept	2.246	0.090	2.440	0.109	1.448	0.058	2.487	0.110
γ_1 Gender	-1.397	0.102	-1.507	0.102			-1.515	0.102
γ_2 Minority Status					-0.644	0.174	-0.727	0.195
Random Effect	VComp	SE	VComp	SE	VComp	SE	VComp	SE
$\tau_0^2 = var(u_{0j})$			0.514	0.084	0.293	0.043	0.481	0.082
Deviance	3345.2		3251.9		3476.1		3238.3	

The problem

- Adding a random intercept or variables with fixed effects to a logistic or probit model may increase effects of earlier included variables.

Logistic regression models for taking a science subject (Snijders & Bosker, 1999, p. 266 f.)

	Single Level		ML Model 1		ML Model 2		ML Model 3	
Fixed Effect	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
γ_0 Intercept	2.246	0.090	2.440	0.109	1.448	0.058	2.487	0.110
γ_1 Gender	-1.397	0.102	-1.507	0.102			-1.515	0.102
γ_2 Minority Status					-0.644	0.174	-0.727	0.195
Random Effect	VComp	SE	VComp	SE	VComp	SE	VComp	SE
$\tau_0^2 = var(u_{0j})$			0.514	0.084	0.293	0.043	0.481	0.082
Deviance	3345.2		3251.9		3476.1		3238.3	

increase

The problem

- Adding a random intercept or variables with fixed effects to a logistic or probit model may increase effects of earlier included variables.

Logistic regression models for taking a science subject (Snijders & Bosker, 1999, p. 266 f.)

	Single Level		ML Model 1		ML Model 2		ML Model 3	
Fixed Effect	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
γ_0 Intercept	2.246	0.090	2.440	0.109	1.448	0.058	2.487	0.110
γ_1 Gender	-1.397	0.102	-1.507	0.102			-1.515	0.102
γ_2 Minority Status					-0.644	0.174	-0.727	0.195
Random Effect	VComp	SE	VComp	SE	VComp	SE	VComp	SE
$\tau_0^2 = var(u_{0j})$			0.514	0.084	0.293	0.043	0.481	0.082
Deviance	3345.2		3251.9		3476.1		3238.3	

increase

The problem

- The ultimate reason for this phenomenon is the fact that in nonlinear probability models the variance of the residual variance (on the individual level) is fixed at a constant (see Long & Freese, 2006):
  - $\pi^2/3 = 3.29$ (logistic regression models)
  - 1.0 (probit regression models)
- Therefore, the residual variance cannot decrease when adding fixed effects of other variables to the model. Instead, the estimates of other regression coefficients (and random effects) will become larger in absolute value.
- As a consequence, changes (or a lack of change) of fixed effects of earlier included variables may not be interpreted as in OLS or multilevel linear regression models: Decreases due to a correlation of the independent variables are obscured by increases due to this phenomenon.

A solution

- Hox (2010) (based on Fielding, 2004) suggests to rescale the fixed and random effects so that real changes in parameter values can be assessed.
- By using the extent of real changes in the level 1 variance when moving from one model to the next, a scaling factor is computed which effectively holds the implicit scaling of the response constant to that of a base model.
- The procedure includes to
  - calculate the total variance of the null model σ_0^2
  - calculate the total variance of model *m* including the first level predictor variables σ_m^2
  - rescale the fixed effects and random effects by using the scale correction factor $\sqrt{\frac{\sigma_0^2}{\sigma_m^2}}$

A solution

- For a multilevel logistic regression model with a random intercept, the scale correction factor is calculated by
  - $$\sigma_0^2 = \sigma_{u0}^2 + \sigma_R^2 = \sigma_{u0}^2 + 3.29$$

    with σ_{u0}^2 = second level intercept variance
    and σ_R^2 = lowest level residual variance
  - $$\sigma_m^2 = \sigma_F^2 + \sigma_{u0}^2 + \sigma_R^2 = \sigma_F^2 + \sigma_{u0}^2 + 3.29$$

    with σ_F^2 = variance of the linear predictor of model *m*, using the coefficients of the predictors of the fixed part of the equation
  - $$SCF = \sqrt{\frac{\sigma_0^2}{\sigma_m^2}}$$

    with *SCF* = scale correction factor

Implementation in Stata

- The rescaling procedure is implemented in Stata by –meresc– (Enzmann & Kohler, 2012), available on SSC:

Title

meresc Rescaled results for nonlinear mixed models

Syntax

```
meresc [ , verbose ]
```

Description

meresc rescales the results of mixed nonlinear probability models such as xtmelogit, xtlogit, or xtprobit to the same scale as the intercept-only model. The technique applied is described in chapter 6.5 of Hox (2010: 133--139).

The technique rescales all random and fixed effects of a multilevel model. The *variance scale correction factor* for random effect parameters is the total variance of the intercept only model devided by the total variance of the model with lowest level variables only. The fixed effects are rescaled using the square root of the variance scale correction factor (i.e. using the *scale correction factor*).

Implementation in Stata

Saved Results

```
    meresc keeps most returned results of the user defined estimation command in
    memory. However, it stores the rescaled coefficient vector in e(b), and the
    rescaled variance-covariance matrix in e(V). Moreover it adds the follwing
    results to the stored results:

    Scalars
      e(SCF)                    Scale correction factor
      e(VCF)                    Variance scale correction factor
      e(Var_Flevel1)            Linear Predictor Variance using first level vars only
      e(Var_u#)                 Variance of Level-# random effect
      e(Var_R)                  Variance of residuals
      e(Var_u0)                 Variance of random effects of constant only model
      e(Var_u#resc)             Variance of Level-# random effect, rescaled
      e(Var_Rresc)              Variance of residuals, rescaled
      e(r2_mz)                  McKelvy & Zavoina's R2
      e(deviance)               Model Deviance

    Macros
      e(cmd)                    meresc
      e(cmdline)                command-line of previous estimation
```

Replication of Hox's example using –meresc–

- Stata syntax (excerpt):

```
// (Re)Production of table 6.9 (p. 138) using -meresc-:

xtmelogit rep1 || schoolid:, var intp(20)
r2_mz
est sto m0
xtmelogit rep1 male pped || schoolid:, var intp(20)
r2_mz
est sto m1
meresc
est sto m1sc
xtmelogit rep1 male pped mses || schoolid:, var intp(20)
r2_mz
est sto m2
meresc
est sto m2sc
```

ado available on SSC to calculate and save in e() model deviance, total of variance of random effects, and McKelvey & Zavoina's pseudo R^2

Save model with results rescaled by -meresc-

Replication of Hox's example using –meresc–

- Stata syntax (excerpt continued):

ado written by Ben Jann (available on SSC) to produce regression table from stored estimates

```
set linesize 94
esttab m0 m1 m1sc m2 m2sc,                                   ///
   se drop(lns1_1_1:_cons)                                   ///
   stat(Var_Rresc Var_u1 Var_u1resc deviance r2_mz,          ///
   labels("fe_var" "var_u1" "var_u1r" "Deviance" "Pseudo R²") ///
   fmt(%5.3f %5.3f %5.3f %8.3f %6.4f))                       ///
   mtitles("M0" "M1" "SC M1" "M2" "SC M2")                   ///
   title("Table 6.9: Logistic regression estimates with rescaling (corrected)")

esttab m0 m1 m1sc m2 m2sc, eform                             ///
   se drop(lns1_1_1:_cons)                                   ///
   stat(Var_Rresc Var_u1 Var_u1resc deviance r2_mz,          ///
   labels("fe_var" "var_u1" "var_u1r" "Deviance" "Pseudo R²") ///
   fmt(%5.3f %5.3f %5.3f %8.3f %6.4f))                       ///
   mtitles("M0" "M1" "SC M1" "M2" "SC M2")                   ///
   title("Table 6.9: Logistic regression estimates with rescaling (corrected, Odds 
Ratios)")
set linesize 82
```

display exponentiated coefficients (here: odds ratios)

Replication of Hox's example using –meresc–

- Output produced by –esttab– (1):

Table 6.9: Logistic regression estimates with rescaling (corrected)

	(1) M0	(2) M1	(3) SC M1	(4) M2	(5) SC M2
eq1					
male		0.536***	0.527***	0.535***	0.526***
		(0.0760)	(0.0747)	(0.0760)	(0.0747)
pped		-0.642***	-0.631***	-0.627***	-0.616***
		(0.0996)	(0.0979)	(0.100)	(0.0985)
mses				-0.296	-0.291
				(0.217)	(0.213)
_cons	-2.234***	-2.237***	-2.198***	-2.242***	-2.203***
	(0.0878)	(0.107)	(0.105)	(0.107)	(0.105)
fe_var			3.178		3.178
var_u1	1.726	1.697	1.697	1.686	1.686
var_u1r			1.640		1.629
Deviance	5537.444	5443.518	5443.518	5441.660	5441.660
Pseudo R^2	0.0000	0.0342	0.0342	0.0389	0.0389

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Replication of Hox's example using –meresc–

- Output produced by –esttab– (2):

Table 6.9: Logistic regression estimates with rescaling (corrected, Odds Ratios)

	(1) M0	(2) M1	(3) SC M1	(4) M2	(5) SC M2
eq1					
male		1.709***	1.694***	1.708***	1.692***
		(0.130)	(0.127)	(0.130)	(0.126)
pped		0.526***	0.532***	0.534***	0.540***
		(0.0524)	(0.0521)	(0.0536)	(0.0532)
mses				0.744	0.748
				(0.161)	(0.160)
_cons	0.107***	0.107***	0.111***	0.106***	0.110***
	(0.00941)	(0.0114)	(0.0116)	(0.0113)	(0.0116)
fe_var			3.178		3.178
var_u1	1.726	1.697	1.697	1.686	1.686
var_u1r			1.640		1.629
Deviance	5537.444	5443.518	5443.518	5441.660	5441.660
Pseudo R^2	0.0000	0.0342	0.0342	0.0389	0.0389

Exponentiated coefficients; Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Replication of Hox's example using –meresc–

- Note that the results shown are correct and differ from the results in Hox (2010) because of a mistake in the course of calculations (mixing up a squared and a non-squared scaling factor).
- The moral of the story: Use Stata ados to automate calculations that are error prone.
- Note that although the effect of rescaling is rather small in the example given by Hox, in other instances rescaling of fixed and random effects may change the results quite substantially!

Thanks for your attention!

References

Enzmann, D. & Kohler, U. (2012). MERESC: Stata module to rescale the results of mixed nonlinear probability models. *Statistical Software Components*, Boston College Department of Economics (http://ideas.repec.org/c/boc/bocode/s457400.html)

Fielding, A. (2004). Scaling for residual variance components of ordered category responses in generalized linear mixed multilevel models. *Quality & Quantity*, *38*, 425-433.

Hox, J. J. (2010). *Multilevel Analysis: Techniques and Applications*. New York (2nd ed.): Routledge.

Long, J. S. & Freese, J. (2006). *Regression Models for Categorical Dependent Variables Using Stata*. College Station, TX (2nd ed.): Stata Press.

Snijders, T. & Bosker, R. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks, CA: SAGE.