

**xtnbreg** — Fixed-effects, random-effects, & population-averaged negative binomial models

[Description](#)[Syntax](#)[Remarks and examples](#)[References](#)[Quick start](#)[Options for RE/FE models](#)[Stored results](#)[Also see](#)[Menu](#)[Options for PA model](#)[Methods and formulas](#)

## Description

`xtnbreg` fits random-effects and conditional fixed-effects overdispersion models where the random effects or fixed effects apply to the distribution of the dispersion parameter. The dispersion is the same for all observations in the same panel. In the random-effects model, the dispersion varies randomly from group to group, such that the inverse of one plus the dispersion follows a Beta distribution. In the fixed-effects model, the dispersion parameter in a group can take on any value.

`xtnbreg` also fits a population-averaged negative binomial model for a nonnegative count dependent variable with overdispersion.

## Quick start

Random-effects negative-binomial regression of  $y$  on  $x$  and [indicators](#) for levels of categorical variable  $a$  using `xtset` data

```
xtnbreg y x i.a
```

As above, but report incidence-rate ratios

```
xtnbreg y x i.a, irr
```

Conditional fixed-effects model with exposure variable `evar`

```
xtnbreg y x i.a, fe exposure(evvar)
```

Population-averaged model with robust standard errors

```
xtnbreg y x i.a, pa vce(robust)
```

## Menu

Statistics > Longitudinal/panel data > Count outcomes > Negative binomial regression (FE, RE, PA)

## Syntax

*Random-effects (RE) and conditional fixed-effects (FE) overdispersion models*

```
xtnbreg depvar [indepvars] [if] [in] [weight] [, [re|fe] RE/FE_options]
```

*Population-averaged (PA) model*

```
xtnbreg depvar [indepvars] [if] [in] [weight] , pa [PA_options]
```

<i>RE/FE_options</i>	Description
Model	
<code>noconstant</code>	suppress constant term; not available with <code>fe</code>
<code>re</code>	use random-effects estimator; the default
<code>fe</code>	use fixed-effects estimator
<code>exposure(<i>varname</i>)</code>	include $\ln(\textit{varname})$ in model with coefficient constrained to 1
<code>offset(<i>varname</i>)</code>	include <i>varname</i> in model with coefficient constrained to 1
<code>constraints(<i>constraints</i>)</code>	apply specified linear constraints
<code>collinear</code>	keep collinear variables
SE	
<code>vce(<i>vcetype</i>)</code>	<i>vcetype</i> may be <code>oim</code> , <code>bootstrap</code> , or <code>jackknife</code>
Reporting	
<code>level(#)</code>	set confidence level; default is <code>level(95)</code>
<code>irr</code>	report incidence-rate ratios
<code>noskip</code>	perform overall model test as a likelihood-ratio test
<code>nocnsreport</code>	do not display constraints
<code>display_options</code>	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
<code>maximize_options</code>	control the maximization process; seldom used
<code>coeflegend</code>	display legend instead of statistics

<i>PA_options</i>	Description
Model	
<code>noconstant</code>	suppress constant term
<code>pa</code>	use population-averaged estimator
<code>exposure(varname)</code>	include $\ln(\text{varname})$ in model with coefficient constrained to 1
<code>offset(varname)</code>	include <i>varname</i> in model with coefficient constrained to 1
Correlation	
<code>corr(correlation)</code>	within-panel correlation structure
<code>force</code>	estimate even if observations unequally spaced in time
SE/Robust	
<code>vce(vctype)</code>	<i>vctype</i> may be <code>conventional</code> , <code>robust</code> , <code>bootstrap</code> , or <code>jackknife</code>
<code>nmp</code>	use divisor $N - P$ instead of the default $N$
<code>scale(parm)</code>	overrides the default scale parameter; <i>parm</i> may be <code>x2</code> , <code>dev</code> , <code>phi</code> , or <code>#</code>
Reporting	
<code>level(#)</code>	set confidence level; default is <code>level(95)</code>
<code>irr</code>	report incidence-rate ratios
<code>display_options</code>	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Optimization	
<code>optimize_options</code>	control the optimization process; seldom used
<code>coeflegend</code>	display legend instead of statistics

<i>correlation</i>	Description
<code>exchangeable</code>	exchangeable
<code>independent</code>	independent
<code>unstructured</code>	unstructured
<code>fixed matname</code>	user-specified
<code>ar #</code>	autoregressive of order <i>#</i>
<code>stationary #</code>	stationary of order <i>#</i>
<code>nonstationary #</code>	nonstationary of order <i>#</i>

A panel variable must be specified. For `xtnbreg`, `pa`, correlation structures other than `exchangeable` and `independent` require that a time variable also be specified. Use `xtset`; see [XT] `xtset`.

`indepvars` may contain factor variables; see [U] 11.4.3 **Factor variables**.

`devar` and `indepvars` may contain time-series operators; see [U] 11.4.4 **Time-series varlists**.

`by`, `mi estimate`, and `statsby` are allowed; see [U] 11.1.10 **Prefix commands**. `fp` is allowed for the random-effects and fixed-effects models.

`vce(bootstrap)` and `vce(jackknife)` are not allowed with the `mi estimate` prefix; see [MI] **mi estimate**.

`weights`, `fweights`, and `pweights` are allowed for the population-averaged model, and `weights` are allowed in the random-effects and fixed-effects models; see [U] 11.1.6 **weight**. Weights must be constant within panel.

`coeflegend` does not appear in the dialog box.

See [U] 20 **Estimation and postestimation commands** for more capabilities of estimation commands.

## Options for RE/FE models

### Model

`noconstant`; see [R] [estimation options](#).

`re` requests the random-effects estimator, which is the default.

`fe` requests the conditional fixed-effects estimator.

`exposure(varname)`, `offset(varname)`, `constraints(constraints)`, `collinear`; see [R] [estimation options](#).

### SE

`vce(vcetype)` specifies the type of standard error reported, which includes types that are derived from asymptotic theory (`oim`) and that use bootstrap or jackknife methods (`bootstrap`, `jackknife`); see [XT] [vce\\_options](#).

### Reporting

`level(#)`; see [R] [estimation options](#).

`irr` reports exponentiated coefficients  $e^b$  rather than coefficients  $b$ . For the negative binomial model, exponentiated coefficients have the interpretation of incidence-rate ratios.

`noskip`; see [R] [estimation options](#).

`nocnsreport`; see [R] [estimation options](#).

`display_options`: `nocl`, `nopvalues`, `noomitted`, `vsquish`, `noemptycells`, `baselevels`, `allbaselevels`, `nofvlabel`, `fvwrap(#)`, `fvwrapon(style)`, `cformat(%fmt)`, `pformat(%fmt)`, `sformat(%fmt)`, and `nolstretch`; see [R] [estimation options](#).

### Maximization

`maximize_options`: `difficult`, `technique(algorithm_spec)`, `iterate(#)`, `[no]log`, `trace`, `gradient`, `showstep`, `hessian`, `showtolerance`, `tolerance(#)`, `ltolerance(#)`, `nrtolerance(#)`, `nonrtolerance`, and `from(init_specs)`; see [R] [maximize](#). These options are seldom used.

The following option is available with `xtnbreg` but is not shown in the dialog box:

`coeflegend`; see [R] [estimation options](#).

## Options for PA model

### Model

`noconstant`; see [R] [estimation options](#).

`pa` requests the population-averaged estimator.

`exposure(varname)`, `offset(varname)`; see [R] [estimation options](#).

### Correlation

`corr(correlation)` specifies the within-panel correlation structure; the default corresponds to the equal-correlation model, `corr(exchangeable)`.

When you specify a correlation structure that requires a lag, you indicate the lag after the structure's name with or without a blank; for example, `corr(ar 1)` or `corr(ar1)`.

If you specify the fixed correlation structure, you specify the name of the matrix containing the assumed correlations following the word `fixed`, for example, `corr(fixed myr)`.

`force` specifies that estimation be forced even though the time variable is not equally spaced. This is relevant only for correlation structures that require knowledge of the time variable. These correlation structures require that observations be equally spaced so that calculations based on lags correspond to a constant time change. If you specify a time variable indicating that observations are not equally spaced, the (time dependent) model will not be fit. If you also specify `force`, the model will be fit, and it will be assumed that the lags based on the data ordered by the time variable are appropriate.

#### SE/Robust

`vce(vctype)` specifies the type of standard error reported, which includes types that are derived from asymptotic theory (`conventional`), that are robust to some kinds of misspecification (`robust`), and that use bootstrap or jackknife methods (`bootstrap`, `jackknife`); see [XT] [vce\\_options](#).

`vce(conventional)`, the default, uses the conventionally derived variance estimator for generalized least-squares regression.

`nmp`, `scale(x2 | dev | phi | #)`; see [XT] [vce\\_options](#).

#### Reporting

`level(#)`; see [R] [estimation options](#).

`irr` reports exponentiated coefficients  $e^b$  rather than coefficients  $b$ . For the negative binomial model, exponentiated coefficients have the interpretation of incidence-rate ratios.

`display_options`: `noci`, `nopvalues`, `noomitted`, `vsquish`, `noemptycells`, `baselevels`, `allbaselevels`, `nofvlabel`, `fvwrap(#)`, `fvwrapon(style)`, `cformat(%fmt)`, `pformat(%fmt)`, `sformat(%fmt)`, and `nolstretch`; see [R] [estimation options](#).

#### Optimization

`optimize_options` control the iterative optimization process. These options are seldom used.

`iterate(#)` specifies the maximum number of iterations. When the number of iterations equals `#`, the optimization stops and presents the current results, even if convergence has not been reached. The default is `iterate(100)`.

`tolerance(#)` specifies the tolerance for the coefficient vector. When the relative change in the coefficient vector from one iteration to the next is less than or equal to `#`, the optimization process is stopped. `tolerance(1e-6)` is the default.

`nolog` suppresses display of the iteration log.

`trace` specifies that the current estimates be printed at each iteration.

The following option is available with `xtnbreg` but is not shown in the dialog box:

`coeflegend`; see [R] [estimation options](#).

## Remarks and examples

`xtnbreg` fits random-effects overdispersion models, conditional fixed-effects overdispersion models, and population-averaged negative binomial models. Here “random effects” and “fixed effects” apply to the distribution of the dispersion parameter, not to the  $x\beta$  term in the model. In the random-effects and fixed-effects overdispersion models, the dispersion is the same for all elements in the same group (that is, elements with the same value of the panel variable). In the random-effects model, the dispersion varies randomly from group to group, such that the inverse of one plus the dispersion follows a  $\text{Beta}(r, s)$  distribution. In the fixed-effects model, the dispersion parameter in a group can take on any value, because a conditional likelihood is used in which the dispersion parameter drops out of the estimation.

By default, the population-averaged model is an equal-correlation model; `xtnbreg`, `pa` assumes `corr(exchangeable)`. Thus, `xtnbreg` is a convenience command for fitting the population-averaged using `xtgee`; see [XT] `xtgee`. Typing

```
. xtnbreg ..., ... pa exposure(time)
```

is equivalent to typing

```
. xtgee ..., ... family(nbinomial) link(log) corr(exchangeable) exposure(time)
```

See also [XT] `xtgee` for information about `xtnbreg`.

By default, or when `re` is specified, `xtnbreg` fits a maximum-likelihood random-effects overdispersion model.

### ▷ Example 1

You have (fictional) data on injury “incidents” incurred among 20 airlines in each of 4 years. (Incidents range from major injuries to exceedingly minor ones.) The government agency in charge of regulating airlines has run an experimental safety training program, and, in each of the years, some airlines have participated and some have not. You now wish to analyze whether the “incident” rate is affected by the program. You choose to estimate using random-effects negative binomial regression, as the dispersion might vary across the airlines for unidentified airline-specific reasons. Your measure of exposure is passenger miles for each airline in each year.



We may alternatively fit a fixed-effects overdispersion model:

```
. xtnbreg i_cnt inprog, exposure(pmiles) irr fe nolog
Conditional FE negative binomial regression      Number of obs      =          80
Group variable: airline                          Number of groups   =          20
                                                Obs per group:
                                                min =              4
                                                avg =              4.0
                                                max =              4
Wald chi2(1)                                    =          2.11
Prob > chi2                                       =          0.1463
Log likelihood = -174.25143
```

i_cnt	IRR	Std. Err.	z	P> z	[95% Conf. Interval]	
inprog	.9062669	.0613917	-1.45	0.146	.793587	1.034946
_cons	.0329025	.0331262	-3.39	0.001	.0045734	.2367111
ln(pmiles)	1	(exposure)				



► Example 2

We rerun our [previous example](#), but this time we fit a robust equal-correlation population-averaged model:

```
. xtnbreg i_cnt inprog, exposure(pmiles) irr vce(robust) pa
Iteration 1: tolerance = .02499392
Iteration 2: tolerance = .0000482
Iteration 3: tolerance = 2.929e-07
GEE population-averaged model
Group variable:                airline      Number of obs      =          80
Link:                          log          Number of groups   =          20
Family:                         negative binomial(k=1)  Obs per group:
Correlation:                    exchangeable      min =              4
                                                avg =              4.0
                                                max =              4
Wald chi2(1)                    =          1.28
Scale parameter:                1          Prob > chi2        =          0.2571
                                (Std. Err. adjusted for clustering on airline)
```

i_cnt	IRR	Semirobust Std. Err.	z	P> z	[95% Conf. Interval]	
inprog	.927275	.0617857	-1.13	0.257	.8137513	1.056636
_cons	.0080211	.0004117	-94.02	0.000	.0072535	.00887
ln(pmiles)	1	(exposure)				



We compare this with a pooled estimator with clustered robust-variance estimates:

```
. nbgreg i_cnt inprog, exposure(pmiles) irr vce(cluster airline)
Fitting Poisson model:
Iteration 0: log pseudolikelihood = -293.57997
Iteration 1: log pseudolikelihood = -293.57997
Fitting constant-only model:
Iteration 0: log pseudolikelihood = -335.13615
Iteration 1: log pseudolikelihood = -279.43327
Iteration 2: log pseudolikelihood = -276.09296
Iteration 3: log pseudolikelihood = -274.84036
Iteration 4: log pseudolikelihood = -274.81076
Iteration 5: log pseudolikelihood = -274.81075
Fitting full model:
Iteration 0: log pseudolikelihood = -274.56985
Iteration 1: log pseudolikelihood = -274.55077
Iteration 2: log pseudolikelihood = -274.55077
Negative binomial regression          Number of obs   =          80
Wald chi2(1)                         =           0.60
Dispersion = mean                    Prob > chi2     =          0.4369
Log pseudolikelihood = -274.55077    Pseudo R2      =          0.0009
                                     (Std. Err. adjusted for 20 clusters in airline)
```

i_cnt	IRR	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
inprog	.9429015	.0713091	-0.78	0.437	.8130032	1.093555
_cons	.007956	.0004237	-90.77	0.000	.0071674	.0088314
ln(pmiles)	1	(exposure)				
/lnalpha	-2.835089	.3351784			-3.492027	-2.178151
alpha	.0587133	.0196794			.0304391	.1132507

## Stored results

`xtnbreg`, `re` stores the following in `e()`:

### Scalars

<code>e(N)</code>	number of observations
<code>e(N_g)</code>	number of groups
<code>e(k)</code>	number of parameters
<code>e(k_aux)</code>	number of auxiliary parameters
<code>e(k_eq)</code>	number of equations in <code>e(b)</code>
<code>e(k_eq_model)</code>	number of equations in overall model test
<code>e(k_dv)</code>	number of dependent variables
<code>e(df_m)</code>	model degrees of freedom
<code>e(ll)</code>	log likelihood
<code>e(ll_0)</code>	log likelihood, constant-only model
<code>e(ll_c)</code>	log likelihood, comparison model
<code>e(chi2)</code>	$\chi^2$
<code>e(chi2_c)</code>	$\chi^2$ for comparison test
<code>e(g_min)</code>	smallest group size
<code>e(g_avg)</code>	average group size
<code>e(g_max)</code>	largest group size
<code>e(r)</code>	value of $r$ in $\text{Beta}(r, s)$
<code>e(s)</code>	value of $s$ in $\text{Beta}(r, s)$
<code>e(p)</code>	significance
<code>e(rank)</code>	rank of <code>e(V)</code>
<code>e(rank0)</code>	rank of <code>e(V)</code> for constant-only model
<code>e(ic)</code>	number of iterations
<code>e(rc)</code>	return code
<code>e(converged)</code>	1 if converged, 0 otherwise

### Macros

<code>e(cmd)</code>	<code>xtnbreg</code>
<code>e(cmdline)</code>	command as typed
<code>e(depvar)</code>	name of dependent variable
<code>e(ivar)</code>	variable denoting groups
<code>e(model)</code>	<code>re</code>
<code>e(wtype)</code>	weight type
<code>e(wexp)</code>	weight expression
<code>e(title)</code>	title in estimation output
<code>e(offset)</code>	linear offset variable
<code>e(chi2type)</code>	Wald or LR; type of model $\chi^2$ test
<code>e(chi2_ct)</code>	Wald or LR; type of model $\chi^2$ test corresponding to <code>e(chi2_c)</code>
<code>e(vce)</code>	<code>vcetype</code> specified in <code>vce()</code>
<code>e(vcetype)</code>	title used to label Std. Err.
<code>e(method)</code>	estimation method
<code>e(distrib)</code>	$\text{Beta}$ ; the distribution of the random effect
<code>e(opt)</code>	type of optimization
<code>e(which)</code>	<code>max</code> or <code>min</code> ; whether optimizer is to perform maximization or minimization
<code>e(ml_method)</code>	type of <code>ml</code> method
<code>e(user)</code>	name of likelihood-evaluator program
<code>e(technique)</code>	maximization technique
<code>e(properties)</code>	<code>b V</code>
<code>e(predict)</code>	program used to implement <code>predict</code>
<code>e(asbalanced)</code>	factor variables <code>fvset</code> as <code>asbalanced</code>
<code>e(asobserved)</code>	factor variables <code>fvset</code> as <code>asobserved</code>

### Matrices

<code>e(b)</code>	coefficient vector
<code>e(Cns)</code>	constraints matrix
<code>e(ilog)</code>	iteration log
<code>e(gradient)</code>	gradient vector
<code>e(V)</code>	variance-covariance matrix of the estimators

### Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

xtnbreg, fe stores the following in  $e()$ :

#### Scalars

<code>e(N)</code>	number of observations
<code>e(N_g)</code>	number of groups
<code>e(k)</code>	number of parameters
<code>e(k_eq)</code>	number of equations in $e(b)$
<code>e(k_eq_model)</code>	number of equations in overall model test
<code>e(k_dv)</code>	number of dependent variables
<code>e(df_m)</code>	model degrees of freedom
<code>e(r2_p)</code>	pseudo $R$ -squared
<code>e(ll)</code>	log likelihood
<code>e(ll_0)</code>	log likelihood, constant-only model
<code>e(chi2)</code>	$\chi^2$
<code>e(g_min)</code>	smallest group size
<code>e(g_avg)</code>	average group size
<code>e(g_max)</code>	largest group size
<code>e(p)</code>	significance
<code>e(rank)</code>	rank of $e(V)$
<code>e(ic)</code>	number of iterations
<code>e(rc)</code>	return code
<code>e(converged)</code>	1 if converged, 0 otherwise

#### Macros

<code>e(cmd)</code>	xtnbreg
<code>e(cmdline)</code>	command as typed
<code>e(depvar)</code>	name of dependent variable
<code>e(ivar)</code>	variable denoting groups
<code>e(model)</code>	fe
<code>e(wtype)</code>	weight type
<code>e(wexp)</code>	weight expression
<code>e(title)</code>	title in estimation output
<code>e(offset)</code>	linear offset variable
<code>e(chi2type)</code>	LR; type of model $\chi^2$ test
<code>e(vce)</code>	<i>vctype</i> specified in <code>vce()</code>
<code>e(vctype)</code>	title used to label Std. Err.
<code>e(method)</code>	requested estimation method
<code>e(opt)</code>	type of optimization
<code>e(which)</code>	max or min; whether optimizer is to perform maximization or minimization
<code>e(ml_method)</code>	type of ml method
<code>e(user)</code>	name of likelihood-evaluator program
<code>e(technique)</code>	maximization technique
<code>e(properties)</code>	<b>b V</b>
<code>e(predict)</code>	program used to implement predict
<code>e(asbalanced)</code>	factor variables <code>fvset</code> as <code>asbalanced</code>
<code>e(asobserved)</code>	factor variables <code>fvset</code> as <code>asobserved</code>

#### Matrices

<code>e(b)</code>	coefficient vector
<code>e(Cns)</code>	constraints matrix
<code>e(ilog)</code>	iteration log
<code>e(gradient)</code>	gradient vector
<code>e(V)</code>	variance-covariance matrix of the estimators

#### Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

`xtnbreg`, `pa` stores the following in `e()`:

#### Scalars

<code>e(N)</code>	number of observations
<code>e(N_g)</code>	number of groups
<code>e(df_m)</code>	model degrees of freedom
<code>e(chi2)</code>	$\chi^2$
<code>e(p)</code>	significance
<code>e(df_pear)</code>	degrees of freedom for Pearson $\chi^2$
<code>e(chi2_dev)</code>	$\chi^2$ test of deviance
<code>e(chi2_dis)</code>	$\chi^2$ test of deviance dispersion
<code>e(deviance)</code>	deviance
<code>e(dispers)</code>	deviance dispersion
<code>e(phi)</code>	scale parameter
<code>e(g_min)</code>	smallest group size
<code>e(g_avg)</code>	average group size
<code>e(g_max)</code>	largest group size
<code>e(rank)</code>	rank of <code>e(V)</code>
<code>e(tol)</code>	target tolerance
<code>e(dif)</code>	achieved tolerance
<code>e(rc)</code>	return code

#### Macros

<code>e(cmd)</code>	<code>xtgee</code>
<code>e(cmd2)</code>	<code>xtnbreg</code>
<code>e(cmdline)</code>	command as typed
<code>e(depvar)</code>	name of dependent variable
<code>e(ivar)</code>	variable denoting groups
<code>e(tvar)</code>	variable denoting time within groups
<code>e(model)</code>	<code>pa</code>
<code>e(family)</code>	negative binomial( $k=1$ )
<code>e(link)</code>	<code>log</code> ; link function
<code>e(corr)</code>	correlation structure
<code>e(scale)</code>	<code>x2</code> , <code>dev</code> , <code>phi</code> , or <code>#</code> ; scale parameter
<code>e(wtype)</code>	weight type
<code>e(wexp)</code>	weight expression
<code>e(offset)</code>	linear offset variable
<code>e(chi2type)</code>	Wald; type of model $\chi^2$ test
<code>e(vce)</code>	<code>vcetype</code> specified in <code>vce()</code>
<code>e(vcetype)</code>	title used to label Std. Err.
<code>e(nmp)</code>	<code>nmp</code> , if specified
<code>e(nbalph)</code>	$\alpha$
<code>e(properties)</code>	<code>b V</code>
<code>e(predict)</code>	program used to implement <code>predict</code>
<code>e(marginsnotok)</code>	predictions disallowed by <code>margins</code>
<code>e(asbalanced)</code>	factor variables <code>fvset</code> as <code>asbalanced</code>
<code>e(asobserved)</code>	factor variables <code>fvset</code> as <code>asobserved</code>

#### Matrices

<code>e(b)</code>	coefficient vector
<code>e(R)</code>	estimated working correlation matrix
<code>e(V)</code>	variance-covariance matrix of the estimators
<code>e(V_modelbased)</code>	model-based variance

#### Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

## Methods and formulas

xtnbreg, pa reports the population-averaged results obtained by using xtgee, family(nbinomial) link(log) to obtain estimates. See [XT] xtgee for details on the methods and formulas.

For the random-effects and fixed-effects overdispersion models, let  $y_{it}$  be the count for the  $t$ th observation in the  $i$ th group. We begin with the model  $y_{it} \mid \gamma_{it} \sim \text{Poisson}(\gamma_{it})$ , where  $\gamma_{it} \mid \delta_i \sim \text{gamma}(\lambda_{it}, \delta_i)$  with  $\lambda_{it} = \exp(\mathbf{x}_{it}\boldsymbol{\beta} + \text{offset}_{it})$  and  $\delta_i$  is the dispersion parameter. This yields the model

$$\Pr(Y_{it} = y_{it} \mid \mathbf{x}_{it}, \delta_i) = \frac{\Gamma(\lambda_{it} + y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it} + 1)} \left(\frac{1}{1 + \delta_i}\right)^{\lambda_{it}} \left(\frac{\delta_i}{1 + \delta_i}\right)^{y_{it}}$$

(See Hausman, Hall, and Griliches [1984, eq. 3.1, 922]; our  $\delta$  is the inverse of their  $\delta$ .) Looking at within-panel effects only, we find that this specification yields a negative binomial model for the  $i$ th group with dispersion (variance divided by the mean) equal to  $1 + \delta_i$ , that is, constant dispersion within group. This parameterization of the negative binomial model differs from the default parameterization of nbreg, which has dispersion equal to  $1 + \alpha \exp(\mathbf{x}\boldsymbol{\beta} + \text{offset})$ ; see [R] nbreg.

For a random-effects overdispersion model, we allow  $\delta_i$  to vary randomly across groups; namely, we assume that  $1/(1 + \delta_i) \sim \text{Beta}(r, s)$ . The joint probability of the counts for the  $i$ th group is

$$\begin{aligned} \Pr(Y_{i1} = y_{i1}, \dots, Y_{in_i} = y_{in_i} \mid \mathbf{X}_i) &= \int_0^\infty \prod_{t=1}^{n_i} \Pr(Y_{it} = y_{it} \mid \mathbf{x}_{it}, \delta_i) f(\delta_i) d\delta_i \\ &= \frac{\Gamma(r + s)\Gamma(r + \sum_{t=1}^{n_i} \lambda_{it})\Gamma(s + \sum_{t=1}^{n_i} y_{it})}{\Gamma(r)\Gamma(s)\Gamma(r + s + \sum_{t=1}^{n_i} \lambda_{it} + \sum_{t=1}^{n_i} y_{it})} \prod_{t=1}^{n_i} \frac{\Gamma(\lambda_{it} + y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it} + 1)} \end{aligned}$$

for  $\mathbf{X}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{in_i})$  and where  $f$  is the probability density function for  $\delta_i$ . The resulting log likelihood is

$$\begin{aligned} \ln L &= \sum_{i=1}^n w_i \left[ \ln \Gamma(r + s) + \ln \Gamma\left(r + \sum_{k=1}^{n_i} \lambda_{ik}\right) + \ln \Gamma\left(s + \sum_{k=1}^{n_i} y_{ik}\right) - \ln \Gamma(r) - \ln \Gamma(s) \right. \\ &\quad \left. - \ln \Gamma\left(r + s + \sum_{k=1}^{n_i} \lambda_{ik} + \sum_{k=1}^{n_i} y_{ik}\right) + \sum_{t=1}^{n_i} \left\{ \ln \Gamma(\lambda_{it} + y_{it}) - \ln \Gamma(\lambda_{it}) - \ln \Gamma(y_{it} + 1) \right\} \right] \end{aligned}$$

where  $\lambda_{it} = \exp(\mathbf{x}_{it}\boldsymbol{\beta} + \text{offset}_{it})$  and  $w_i$  is the weight for the  $i$ th group (Hausman, Hall, and Griliches 1984, eq. 3.5, 927).

For the fixed-effects overdispersion model, we condition the joint probability of the counts for each group on the sum of the counts for the group (that is, the observed  $\sum_{t=1}^{n_i} y_{it}$ ). This yields

$$\begin{aligned} \Pr(Y_{i1} = y_{i1}, \dots, Y_{in_i} = y_{in_i} \mid \mathbf{X}_i, \sum_{t=1}^{n_i} Y_{it} = \sum_{t=1}^{n_i} y_{it}) \\ = \frac{\Gamma(\sum_{t=1}^{n_i} \lambda_{it})\Gamma(\sum_{t=1}^{n_i} y_{it} + 1)}{\Gamma(\sum_{t=1}^{n_i} \lambda_{it} + \sum_{t=1}^{n_i} y_{it})} \prod_{t=1}^{n_i} \frac{\Gamma(\lambda_{it} + y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it} + 1)} \end{aligned}$$

The conditional log likelihood is

$$\ln L = \sum_{i=1}^n w_i \left[ \ln \Gamma \left( \sum_{t=1}^{n_i} \lambda_{it} \right) + \ln \Gamma \left( \sum_{t=1}^{n_i} y_{it} + 1 \right) - \ln \Gamma \left( \sum_{t=1}^{n_i} \lambda_{it} + \sum_{t=1}^{n_i} y_{it} \right) + \sum_{t=1}^{n_i} \left\{ \ln \Gamma(\lambda_{it} + y_{it}) - \ln \Gamma(\lambda_{it}) - \ln \Gamma(y_{it} + 1) \right\} \right]$$

See [Hausman, Hall, and Griliches \(1984\)](#) for a more thorough development of the random-effects and fixed-effects models. Also see [Cameron and Trivedi \(2013\)](#) for a good textbook treatment of this model.

## References

- Cameron, A. C., and P. K. Trivedi. 2013. *Regression Analysis of Count Data*. 2nd ed. New York: Cambridge University Press.
- Guimarães, P. 2005. A simple approach to fit the beta-binomial model. *Stata Journal* 5: 385–394.
- Hausman, J. A., B. H. Hall, and Z. Griliches. 1984. Econometric models for count data with an application to the patents–R & D relationship. *Econometrica* 52: 909–938.
- Liang, K.-Y., and S. L. Zeger. 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73: 13–22.

## Also see

- [XT] **xtnbreg postestimation** — Postestimation tools for xtnbreg
- [XT] **xtgee** — Fit population-averaged panel-data models by using GEE
- [XT] **xtpoisson** — Fixed-effects, random-effects, and population-averaged Poisson models
- [XT] **xtset** — Declare data to be panel data
- [ME] **menbreg** — Multilevel mixed-effects negative binomial regression
- [MI] **estimation** — Estimation commands for use with mi estimate
- [R] **nbreg** — Negative binomial regression
- [U] **20 Estimation and postestimation commands**