Package 'xtvfreg'

Type Package

Title Variance Function Panel Regression

Version 1.1.0

Description Implements an iterative mean-variance panel regression estimator that allows both the mean and variance of the dependent variable to be functions of covariates. The method alternates between estimating a mean equation (using generalized linear models with Gaussian family) and a variance equation (using generalized linear models with Gamma family on squared within-group residuals) until convergence. Based on the methodology in Mooi-Reci and Liao (2025) <doi:10.1093/esr/jcae052>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Depends R (>= 3.6.0)

Imports stats, rlang

Suggests haven, ggplot2, knitr, gt, modelsummary

RoxygenNote 7.3.3

NeedsCompilation no

Author Tim F. Liao [aut, cre]

Maintainer Tim F. Liao <tfliao@illinois.edu>

Repository CRAN

Date/Publication 2025-11-26 20:40:02 UTC

Contents

---

coef.xtvfreg   *Coefficient extraction method for xtvfreg objects*

Description

Coefficient extraction method for xtvfreg objects

Usage

```
## S3 method for class 'xtvfreg'
coef(object, equation = ”mean”, group = NULL, ...)
```

Arguments

object	An object of class "xtvfreg"
equation	Character; "mean" or "variance"
group	Optional; specific group to extract. If NULL, returns all groups
...	Additional arguments (currently unused)

Value

Named vector or list of coefficient vectors

Examples

```
set.seed(456)
n_id <- 30
n_time <- 4
panel_data <- data.frame(
id = factor(rep(1:n_id, each = n_time)),  # Convert to factor here
group = rep(c(”A”, ”B”), length.out = n_id * n_time),
x = rnorm(n_id * n_time)
)
panel_data$m_x <- ave(panel_data$x, panel_data$id, FUN = mean)
panel_data$d_x <- panel_data$x - panel_data$m_x
panel_data$y <- 1 + panel_data$m_x + rnorm(n_id * n_time)

result <- xtvfreg(
  formula = y ~ 1,
  data = panel_data,
  group = ”group”,
  panel_id = ”id”,
```

```
  mean_vars = c("m_x", "d_x"),
  var_vars = "m_x",
  verbose = FALSE
)

# Extract coefficients
coef(result, equation = "mean")
coef(result, equation = "mean", group = "A")
```

---

comparison_table   *Create comparison table across groups*

Description

Create comparison table across groups

Usage

```
comparison_table(object, equation = "mean", output = "data.frame", ...)
```

Arguments

object	An object of class "xtvfreg"
equation	Character; "mean" or "variance"
output	Character; "data.frame", "kable", or "gt"
...	Additional arguments passed to formatting functions

Value

A formatted table (type depends on output parameter)

Examples

```
# Create small simulated dataset
set.seed(456)
n_id <- 30
n_time <- 4
panel_data <- data.frame(
  id = factor(rep(1:n_id, each = n_time)),
  group = factor(rep(c("A", "B"), length.out = n_id * n_time)),
  x = rnorm(n_id * n_time)
)
panel_data$m_x <- ave(panel_data$x, panel_data$id, FUN = mean)
panel_data$d_x <- panel_data$x - panel_data$m_x
panel_data$y <- 1 + panel_data$m_x + rnorm(n_id * n_time)

result <- xtvfreg(
  formula = y ~ 1,
  data = panel_data,
```

```
  group = "group",
  panel_id = "id",
  mean_vars = c("m_x", "d_x"),
  var_vars = "m_x",
  verbose = FALSE
)

# Create comparison table
comparison_table(result, equation = "mean")
```

---

export_table   *Export results to LaTeX or CSV*

Description

Export results to LaTeX or CSV

Usage

```
export_table(object, file, format = "csv", equation = "both")
```

Arguments

object	An object of class "xtvfreg"
file	Character; output file name
format	Character; "latex" or "csv"
equation	Character; "mean", "variance", or "both"

Value

Invisibly returns NULL; called for side effects

Examples

```
# Create temporary file
set.seed(456)
n_id <- 30
n_time <- 4
panel_data <- data.frame(
  id = rep(1:n_id, each = n_time),
  group = rep(c("A", "B"), length.out = n_id * n_time),
  x = rnorm(n_id * n_time)
)
panel_data$m_x <- ave(panel_data$x, panel_data$id, FUN = mean)
panel_data$d_x <- panel_data$x - panel_data$m_x
panel_data$y <- 1 + panel_data$m_x + rnorm(n_id * n_time)

result <- xtvfreg(
  formula = y ~ 1,
```

```
    data = panel_data,
    group = "group",
    panel_id = "id",
    mean_vars = c("m_x", "d_x"),
    var_vars = "m_x",
    verbose = FALSE
)

# Export to CSV
temp_file <- tempfile(fileext = ".csv")
export_table(result, file = temp_file, format = "csv", equation = "mean")

# Clean up
unlink(temp_file)
```

---

get_variance_decomp   *Get variance decomposition*

Description

Get variance decomposition

Usage

```
get_variance_decomp(object, group = NULL)
```

Arguments

object	An object of class "xtvfreg"
group	Optional; specific group to extract. If NULL, returns all groups

Value

A data frame with variance decomposition results

Examples

```
set.seed(456)
n_id <- 30
n_time <- 4
panel_data <- data.frame(
  id = factor(rep(1:n_id, each = n_time)),
  group = factor(rep(c("A", "B"), length.out = n_id * n_time)),
  x = rnorm(n_id * n_time)
)
panel_data$m_x <- ave(panel_data$x, panel_data$id, FUN = mean)
panel_data$d_x <- panel_data$x - panel_data$m_x
panel_data$y <- 1 + panel_data$m_x + rnorm(n_id * n_time)
```

```
result <- xtvfreg(
  formula = y ~ 1,
  data = panel_data,
  group = "group",
  panel_id = "id",
  mean_vars = c("m_x", "d_x"),
  var_vars = "m_x",
  verbose = FALSE
)

# Get variance decomposition
get_variance_decomp(result)
```

---

nlswork_subset   *National Longitudinal Survey of Young Women (Subset)*

Description

A subset of 300 randomly sampled women from the National Longitudinal Survey of Young Women, 1968-1988. This is a subsample of the full nlswork dataset commonly used in Stata examples. The data contains labor market information for young women tracked over multiple years.

Usage

```
nlswork_subset
```

Format

A data frame with approximately 2,400-2,700 observations (depending on sampling) and the following variables:

idcode Individual identifier (numeric)

year Survey year (numeric)

birth_yr Year of birth (numeric)

age Age in current year (numeric)

race Race: 1=white, 2=black, 3=other (numeric or labeled)

msp Marital status: 1=never married, 2=married, 3=separated/divorced/widowed (numeric or labeled)

nev_mar 1 if never married (numeric)

grade Current grade completed (numeric)

collgrad 1 if college graduate (numeric)

not_smsa 1 if not in SMSA (Standard Metropolitan Statistical Area) (numeric)

c_city 1 if in central city (numeric)

south 1 if in south (numeric)

ind_code Industry code (numeric)

occ_code Occupation code (numeric)

union 1 if union member (numeric)

wks_ue Weeks unemployed last year (numeric)

ttl_exp Total work experience (years) (numeric)

tenure Job tenure in years (numeric)

hours Usual hours worked per week (numeric)

wks_work Weeks worked last year (numeric)

ln_wage Natural log of hourly wage (numeric)

Details

This dataset is a subset of the nlswork data available from Stata Press. It contains 300 randomly sampled individuals from the original 5,159 women, preserving all time periods for the selected individuals. The data is an unbalanced panel with varying numbers of observations per individual.

The subset was created using:

```
set.seed(123)
unique_ids <- unique(nlswork$idcode)
sampled_ids <- sample(unique_ids, size = 300, replace = FALSE)
nlswork_subset <- nlswork[nlswork$idcode %in% sampled_ids, ]
```

Source

Original data from Stata Press: https://www.stata-press.com/data/r19/nlswork.dta

National Longitudinal Survey of Young Women, 1968-1988. U.S. Bureau of Labor Statistics.

References

Center for Human Resource Research. (2002). NLS Handbook 2001. Columbus, OH: The Ohio State University.

Examples

```
# Load the data
data(nlswork_subset)

# Examine structure
str(nlswork_subset)

# Summary statistics
summary(nlswork_subset$ln_wage)

# Panel structure
table(table(nlswork_subset$idcode))  # Distribution of obs per individual

## Not run:
```

```
# Example analysis with xtvfreg
# Create race groups
nlswork_subset$race_group <- factor(nlswork_subset$race,
                                    levels = 1:2,
                                    labels = c("white", "black"))

# Create within and between components for tenure
nlswork_subset$m_tenure <- ave(nlswork_subset$tenure,
                               nlswork_subset$idcode,
                               FUN = function(x) mean(x, na.rm = TRUE))
nlswork_subset$d_tenure <- nlswork_subset$tenure - nlswork_subset$m_tenure

# Estimate varying effects model
result <- xtvfreg(
  formula = ln_wage ~ 1,
  data = subset(nlswork_subset, !is.na(ln_wage) & race %in% 1:2),
  group = "race_group",
  panel_id = "idcode",
  mean_vars = c("m_tenure", "d_tenure", "age"),
  var_vars = c("m_tenure", "age"),
  verbose = TRUE
)

# View results
summary(result)

## End(Not run)
```

---

plot_coefficients   *Plot coefficient comparison across groups*

Description

Plot coefficient comparison across groups

Usage

```
plot_coefficients(object, equation = "mean", variable = NULL)
```

Arguments

object	An object of class "xtvfreg"
equation	Character; "mean" or "variance"
variable	Optional; specific variable to plot. If NULL, plots all

Value

A ggplot2 object

Examples

```
# Requires ggplot2 package
if (requireNamespace("ggplot2", quietly = TRUE)) {
  set.seed(456)
  n_id <- 30
  n_time <- 4
  panel_data <- data.frame(
    id = rep(1:n_id, each = n_time),
    group = rep(c("A", "B"), length.out = n_id * n_time),
    x = rnorm(n_id * n_time)
  )
  panel_data$m_x <- ave(panel_data$x, panel_data$id, FUN = mean)
  panel_data$d_x <- panel_data$x - panel_data$m_x
  panel_data$y <- 1 + panel_data$m_x + rnorm(n_id * n_time)

  result <- xtvfreg(
    formula = y ~ 1,
    data = panel_data,
    group = "group",
    panel_id = "id",
    mean_vars = c("m_x", "d_x"),
    var_vars = "m_x",
    verbose = FALSE
  )

  # Plot coefficients
  plot_coefficients(result, equation = "mean")
}
```

---

print.xtvfreg   *Print method for xtvfreg objects*

Description

Print method for xtvfreg objects

Usage

```
## S3 method for class 'xtvfreg'
print(x, ...)
```

Arguments

x	An object of class "xtvfreg"
...	Additional arguments (currently unused)

Value

Invisibly returns the input object

Examples

```
set.seed(456)
n_id <- 30
n_time <- 4
panel_data <- data.frame(
  id = rep(1:n_id, each = n_time),
  group = rep(c("A", "B"), length.out = n_id * n_time),
  x = rnorm(n_id * n_time)
)
panel_data$m_x <- ave(panel_data$x, panel_data$id, FUN = mean)
panel_data$d_x <- panel_data$x - panel_data$m_x
panel_data$y <- 1 + panel_data$m_x + rnorm(n_id * n_time)

result <- xtvfreg(
  formula = y ~ 1,
  data = panel_data,
  group = "group",
  panel_id = "id",
  mean_vars = c("m_x", "d_x"),
  var_vars = "m_x",
  verbose = FALSE
)

print(result)
```

---

summary.xtvfreg	*Summary method for xtvfreg objects*

Description

Summary method for xtvfreg objects

Usage

```
## S3 method for class 'xtvfreg'
summary(object, ...)
```

Arguments

object	An object of class "xtvfreg"
...	Additional arguments (currently unused)

Value

Invisibly returns the input object

Examples

```
set.seed(456)
n_id <- 30
n_time <- 4
panel_data <- data.frame(
  id = rep(1:n_id, each = n_time),
  group = rep(c("A", "B"), length.out = n_id * n_time),
  x = rnorm(n_id * n_time)
)
panel_data$m_x <- ave(panel_data$x, panel_data$id, FUN = mean)
panel_data$d_x <- panel_data$x - panel_data$m_x
panel_data$y <- 1 + panel_data$m_x + rnorm(n_id * n_time)

result <- xtvfreg(
  formula = y ~ 1,
  data = panel_data,
  group = "group",
  panel_id = "id",
  mean_vars = c("m_x", "d_x"),
  var_vars = "m_x",
  verbose = FALSE
)

summary(result)
```

---

vcov.xtvfreg   *Variance-covariance matrix extraction method*

Description

Variance-covariance matrix extraction method

Usage

```
## S3 method for class 'xtvfreg'
vcov(object, equation = "mean", group = NULL, ...)
```

Arguments

object	An object of class "xtvfreg"
equation	Character; "mean" or "variance"
group	Optional; specific group to extract. If NULL, returns all groups
...	Additional arguments (currently unused)

Value

Matrix or list of matrices

Examples

```
set.seed(456)
n_id <- 30
n_time <- 4
panel_data <- data.frame(
  id = rep(1:n_id, each = n_time),
  group = rep(c("A", "B"), length.out = n_id * n_time),
  x = rnorm(n_id * n_time)
)
panel_data$m_x <- ave(panel_data$x, panel_data$id, FUN = mean)
panel_data$d_x <- panel_data$x - panel_data$m_x
panel_data$y <- 1 + panel_data$m_x + rnorm(n_id * n_time)

result <- xtvfreg(
  formula = y ~ 1,
  data = panel_data,
  group = "group",
  panel_id = "id",
  mean_vars = c("m_x", "d_x"),
  var_vars = "m_x",
  verbose = FALSE
)

# Extract variance-covariance matrix
vcov(result, equation = "mean", group = "A")
```

---

  xtvfreg   *Variance Function Panel Regression*

Description

Implements an iterative mean-variance panel regression estimator that allows both the mean and variance of the dependent variable to be functions of covariates. Based on Mooi-Reci and Liao (2025).

Usage

```
xtvfreg(
  formula,
  data,
  group,
  panel_id,
  mean_vars,
  var_vars,
  weights = NULL,
  subset = NULL,
  converge = 1e-06,
  max_iter = 100,
  verbose = TRUE
)
```

Arguments

`formula`	A formula specifying the dependent variable
`data`	A data frame containing the variables
`group`	A character string naming the grouping variable
`panel_id`	A character string naming the panel identifier
`mean_vars`	A character vector of variable names for the mean equation
`var_vars`	A character vector of variable names for the variance equation
`weights`	Optional character string naming the weight variable
`subset`	Optional logical vector for subsetting
`converge`	Convergence tolerance (default: 1e-6)
`max_iter`	Maximum iterations (default: 100)
`verbose`	Logical; print iteration history? (default: TRUE)

Value

An object of class "xtvfreg" containing:

`results`	List of results for each group
`groups`	Vector of group values
`call`	The matched call
`convergence`	Convergence information for each group
`variance_decomp`	
	Variance decomposition for each group
`depvar`	Name of dependent variable
`panel_id`	Name of panel identifier
`group_var`	Name of grouping variable
`mean_vars`	Variables in mean equation
`var_vars`	Variables in variance equation

References

Mooi-Reci, I., and Liao, T. F. (2025). Unemployment: a hidden source of wage inequality? European Sociological Review, 41(3), 382-401. doi:10.1093/esr/jcae052

Examples

```
# Example using nlswork subset data
data(nlswork_subset)

# Prepare the data
# Keep only observations with complete wage data and white/black race
analysis_data <- subset(nlswork_subset,
                        !is.na(ln_wage) &
                        !is.na(tenure) &
```

```
                                race %in% 1:2)

# Create race grouping variable
analysis_data$race_group <- factor(analysis_data$race,
                                   levels = 1:2,
                                   labels = c("white", "black"))

# Create within and between components for tenure
analysis_data$m_tenure <- ave(analysis_data$tenure,
                              analysis_data$idcode,
                              FUN = function(x) mean(x, na.rm = TRUE))
analysis_data$d_tenure <- analysis_data$tenure - analysis_data$m_tenure

# Estimate varying effects model
result <- xtvfreg(
  formula = ln_wage ~ 1,
  data = analysis_data,
  group = "race_group",
  panel_id = "idcode",
  mean_vars = c("m_tenure", "d_tenure", "age"),
  var_vars = c("m_tenure"),
  verbose = FALSE
)

# View a summary of results
summary(result)

# Extract coefficients for white group if needed
coef(result, equation = "mean", group = "white")
```

Index