Package: hetiv (via r-universe)

July 9, 2026

Title Heteroskedasticity and Proxy-Based IV Methods

Version 0.1.0

Description Tools for identifying structural shocks using heteroskedasticity- and proxy-based instrumental variable (IV) methods in event-study settings. Supports local projection impulse response estimation of Jorda (2005) <doi:10.1257/0002828053828518>, generalised weak instrument testing of Lewis and Mertens (2025) <doi:10.1093/restud/rdaf103>, recursive heteroskedasticity-IV identification of Burri and Kaufmann (2026b) <doi:10.1016/j.econlet.2026.113091>, and Kalman-filter shock extraction of Burri and Kaufmann (2026a).

URL https://dankaufmann.github.io/hetiv/, https://github.com/dankaufmann/hetiv

BugReports https://github.com/dankaufmann/hetiv/issues

License MIT + file LICENSE

Encoding UTF-8

Depends R (>= 4.1.0)

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Config/roxygen2/version 7.3.3

VignetteBuilder knitr

Imports dplyr, ggplot2, grid, ivreg, MASS, rlang, sandwich, zoo

Suggests cowplot, knitr, rmarkdown, testthat (>= 3.0.0)

Config/pak/sysreqs cmake make libicu-dev

Repository https://marcburri.r-universe.dev

Date/Publication 2026-07-09 12:25:52 UTC

RemoteUrl https://github.com/dankaufmann/hetiv

RemoteRef main

RemoteSha 9c7c2248ee4f0d7f2075280f63c25348b578b1e6

Contents

`computeirf` *Compute impulse response functions for a VAR(P)*

Description

Recursively computes IRFs given an impact matrix and VAR coefficient array. Optionally cumulates responses for selected variables.

Usage

```
computeirf(Psi, Phi, H, cum)
```

Arguments

`Psi`	Impact matrix, dimension `N x R`.
`Phi`	Array of VAR coefficient matrices, dimension `N x N x P`.
`H`	Horizon (number of periods) for which to compute IRFs.
`cum`	Logical scalar or logical vector of length `N`. If `TRUE` for variable `i`, the IRF is cumulated via `cumsum`. A single value is applied to all variables.

Value

An array:

irf Array of impulse responses, dimension `H x N x R`. Row names are labelled `0` to `H-1`.

Examples

```
Psi <- matrix(c(1, 0.5), nrow = 2)
Phi <- array(0, dim = c(2, 2, 1))
Phi[, , 1] <- matrix(c(0.4, 0.1, 0, 0.3), 2, 2)
computeirf(Psi = Psi, Phi = Phi, H = 4, cum = FALSE)
```

---

filllinear *Linearly interpolate gaps in a data frame*

Description

Fills missing values (NA) in every column of a data frame using linear interpolation, up to a maximum of gap consecutive missing observations. Gaps longer than gap are left as NA.

Usage

```
filllinear(DF, gap)
```

Arguments

DF	A data frame or matrix with numeric columns.
gap	Integer. Maximum number of consecutive NAs to fill.

Value

The data frame DF with short gaps linearly interpolated.

Examples

```
filllinear(data.frame(x = c(1, NA, 3), y = c(2, NA, 4)), gap = 1)
```

---

firstdiff *First-difference transformation*

Description

Computes the first difference of a numeric vector or time series, i.e. x_t - x_{t-1}. Gaps in dates are ignored; the difference is always taken with respect to the immediately preceding observation.

Usage

```
firstdiff(TS)
```

Arguments

`TS`	A numeric vector or time series.

Value

A numeric vector of the same length as `TS`, with `NA` in the first position.

Examples

```
firstdiff(c(1, 3, 6))
```

`gweakivtest` *Generalised weak instrument test*

Description

Tests for weak instruments with multiple endogenous regressors using the generalised minimum eigenvalue statistic of Lewis and Mertens (2025). The test is robust to heteroskedasticity and autocorrelation and nests the classical Stock-Yogo (2005) test as a special case. The function is a direct port from the Matlab codes by Lewis and Mertens (2025)

Usage

```
gweakivtest(
  y,
  Y,
  X,
  Z,
  cov_type = "EHW",
  alfa = 0.05,
  tau = 0.1,
  points = 1000L,
  target = "beta",
  crit = "abs",
  seed = 12345L
)
```

Arguments

`y`	Regressand (T x 1 numeric vector or matrix).
`Y`	Endogenous regressors (T x N numeric matrix).
`X`	Exogenous regressors (T x Nx numeric matrix). A constant column is added automatically if one is absent or if `X` has zero columns.
`Z`	Instruments (T x K numeric matrix). Requires K >= N.

`cov_type` HAR covariance estimator: `"EHW"` (Eicker-Huber-White, default) or `"NW"` (Newey-West with Lazarus et al. (2018) bandwidth).

`alfa` Nominal significance level (default `0.05`).

`tau` Bias tolerance: maximum acceptable relative (or absolute, see `crit`) bias of the 2SLS estimator (default `0.10`).

`points` Number of random starting points for the Stiefel manifold optimisation used to compute the sharp critical value when K > N + 1 (default `1000`).

`target` Either `"beta"` (default) to test the full coefficient vector, or a positive integer `j <= N` to target the single coefficient `beta_j`.

`crit` Bias criterion: `"abs"` (absolute bias, default) or `"rel"` (relative bias). `"abs"` requires the error covariance matrix; `"rel"` does not.

`seed` Integer seed used for the random Stiefel starting points when computing the sharp critical value. The default is fixed for reproducible critical values and the caller's RNG state is restored afterwards. Use `NA` to leave the RNG unmanaged.

Value

A named list with the following elements:

`nobs` Number of complete observations used.

`beta_2SLS` 2SLS point estimate(s).

`target` Description of the targeted parameter.

`criterion` Bias criterion used (`"abs"` or `"rel"`).

`gmin_generalized` Generalised minimum eigenvalue test statistic (Lewis-Mertens).

`gmin_generalized_critical_value` Sharp critical value via Stiefel optimisation.

`gmin_generalized_critical_value_simplified` Conservative simplified critical value (closed-form bound).

`stock_yogo_test_statistic` Stock-Yogo test statistic under the Nagar approximation.

`stock_yogo_critical_value_nagar` Stock-Yogo critical value under the Nagar approximation.

References

Lazarus, E., Lewis, D. J., Stock, J. H. and Watson, M. W. (2018). HAR inference: recommendations for practice. *Journal of Business & Economic Statistics*, 36(4), 541-559.

Lewis, D. J. and Mertens, K. (2025). A robust test for weak instruments for 2SLS with multiple endogenous regressors. *Review of Economic Studies*, DOI: 10.1093/restud/rdaf103.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and inference for econometric models: essays in honor of Thomas Rothenberg*, pp. 80-108. Cambridge University Press.

Examples

```
set.seed(1)
n <- 80
Z <- matrix(rnorm(n), ncol = 1)
X <- matrix(1, n, 1)
Y <- 0.8 * Z + rnorm(n)
y <- 1.5 * Y + rnorm(n)
gweakivtest(y = y, Y = Y, X = X, Z = Z, points = 1)
```

---

hetiv — *Estimate impulse responses via heteroskedasticity-based IV local projections*

Description

Estimates impulse response functions (IRFs) using recursive heteroskedasticity-IV identification (Rigobon, 2003; Rigobon and Sack, 2004; Lewis, 2022; Burri and Kaufmann, 2026a, 2026b) combined with local projections (Jorda, 2005). Identification exploits the difference in variance between policy event days and control days to construct instruments for the endogenous variables.

Usage

```
hetiv(
  y,
  O,
  X = NULL,
  Ind,
  P,
  H,
  E = 1,
  norm = 1,
  interact = FALSE,
  cum = FALSE,
  Hstep = 1,
  cov_type = "HC0",
  details = FALSE
)
```

Arguments

`y`	Numeric matrix of stationary outcome variables (T x N). The effect on the first variable in each dimension is normalized to unity at horizon 0. These variables are also used to construct heteroskedasticity-based instruments and for recursive ordering to identify multiple dimensions.
`O`	Numeric matrix of information set variables (T x M). May be identical to `y`. Included as lags 1 through `P`.

X — Numeric matrix of deterministic variables (T x K). For example, time trend, seasonal dummies or other deterministic controls. Included as is (no lags). A constant is included by default.

Ind — Integer vector of length T, event indicator:

- `0` Control day (no event)
- `1` Policy day (event)
- `2` Contaminated control day (excluded from estimation)

P — Integer. Maximum lag order for the information set. Set to `0` for no lags (regression on deterministic terms only).

H — Integer. Maximum horizon (in periods) up to which IRFs are estimated.

E — Integer. Number of shock dimensions to identify via recursive ordering.

norm — Numeric scalar. Normalize the impact response of the first variable to a specific value. Set to `1` for standard unit-effect normalization.

interact — Logical. If `TRUE`, lagged information set variables are interacted with event/non-event dummies.

cum — Logical vector of length N. For each variable in `y`, whether to report the cumulative impulse response instead of the level response. If only one provided, applied to all impulse responses.

Hstep — Integer. Step size between horizons. The default `1` estimates all horizons 0 through H - 1. Values greater than 1 estimate only the selected horizons.

cov_type — Covariance estimator for local-projection standard errors: `"HC0"` (default) for heteroskedasticity-robust standard errors or `"NW"` for Newey-West HAC standard errors. `"HC0"` is the default because Montiel Olea et al. (2025) show that heteroskedasticity-robust standard errors suffice for local-projection impulse responses under weak conditions, even though multi-step forecast errors are typically serially correlated. `"NW"` remains available as an optional HAC robustness check.

details — Logical. If `TRUE`, code saves detailed IV results, which is slightly slower. if set to `FALSE`, returns only impulse response and standard error (e.g. for bootstrap)

Details

For `E > 1`, identification is recursive and order-dependent: the column order of `y` defines both the shock ordering and the normalization variable for each shock dimension.

Value

A named list with the following elements:

`irf` Array (H x N x E) of estimated impulse responses.

`se` Array (H x N x E) of local-projection standard errors.

`IVRes` List of `ivreg` model objects, one per horizon, variable, and shock dimension.

`Obs` Data frame with observation counts: `Tp` (policy days), `Tc` (control days), `To` (contaminated days), `Tt` (total used).

`Method` Character string `"Heteroscedasticity-IV"`.

`et` Data frame of OLS residuals on event days (used for covariance estimation and shock extraction).

`Sig` Covariance matrix of residuals on event days (used for shock extraction).

`SigR` Covariance matrix of residuals on control days, or `NA` if unavailable (used for shock extraction).

`Psi` Impact matrix (N x E), equal to `irf[1, , ]`. By the package's indexing convention `HSeries` starts at 1, so the first LP uses `lead(y, 0)` (the contemporaneous value) and is labelled horizon 0; `irf[1, , ]` is therefore always the impact response.

`WeakData` Data frame of endogenous variables and instruments for the Lewis-Mertens (2025) weak instrument test.

References

Burri, M. and D. Kaufmann (2026a). Measuring monetary policy shocks. IRENE Working Papers 24-03, IRENE Institute of Economic Research, University of Neuchatel.

Burri, M. and D. Kaufmann (2026b). Multiple monetary policy shocks from daily data: A heteroskedasticity IV approach. *Economics Letters*, 268, 113091. doi:10.1016/j.econlet.2026.113091

Jorda, O. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1), 161-182.

Lewis, D. J. (2022). Robust inference in models identified via heteroskedasticity. *Review of Economics and Statistics*, 104(3), 510-524.

Lewis, D. J. and Mertens, K. (2025). A robust test for weak instruments for 2SLS with multiple endogenous regressors. *Review of Economic Studies*, DOI: 10.1093/restud/rdaf103

Montiel Olea, J. L., M. Plagborg-Moller, E. Qian, and C. K. Wolf (2025). Local projections or VARs? A primer for macroeconomists. *NBER Working Paper* No. 33871.

Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics*, 85(4), 777-792.

Rigobon, R. and Sack, B. (2004). The impact of monetary policy on asset prices. *Journal of Monetary Economics*, 51(8), 1553-1575.

Examples

```
set.seed(1)
y <- matrix(rnorm(80), ncol = 2)
Ind <- rep(0L, nrow(y))
Ind[seq(5, nrow(y), by = 5)] <- 1L
res <- hetiv(y = y, O = y, Ind = Ind, P = 1, H = 3, details = TRUE)
dim(res$irf)
```

`kfpredict` *Extract monetary policy shocks via Kalman filter*

Description

Extracts structural shocks from reduced-form residuals using the Kalman filter formula. Given the impact matrix `Psi` and the covariance matrices of residuals on event and control days, the function recovers the latent shocks for each observation. Optionally rescales shocks to unit variance.

Usage

```
kfpredict(Sig, SigR, Psi, et, tol = sqrt(.Machine$double.eps), scale = TRUE)
```

Arguments

`Sig`	Numeric matrix (N x N). Covariance matrix of reduced-form residuals on policy event days (e.g. `et` from `hetiv()`).
`SigR`	Numeric matrix (N x N). Covariance matrix of reduced-form residuals on control (non-event) days. Used to back out the implied shock variances when `scale = TRUE`.
`Psi`	Numeric matrix (N x E). Impact matrix, i.e. the contemporaneous responses of all N variables to the E structural shocks (e.g. `Psi` from `hetiv()`).
`et`	Numeric matrix or data frame (T x N). Reduced-form residuals on event days (e.g. `et` from `hetiv()`). Rows correspond to time periods, columns to variables.
`tol`	Numeric scalar. Tolerance for the generalized inverse (`MASS::ginv()`). Clipped from below at `sqrt(.Machine$double.eps)`.
`scale`	Logical. If `TRUE` (default), shocks are rescaled to unit variance using the implied shock variances recovered from `Sig` and `SigR`. This scaling assumes that structural shock variances are diagonal, so that `Sig - SigR = sum_i sigma_i Psi_i Psi_i'`. If `FALSE`, the raw Kalman filter projection is returned.

Value

A numeric matrix (T x E) of extracted structural shocks, with the same number of rows as `et` and one column per shock dimension.

References

Burri, M. and Kaufmann, D. (2026a). Measuring monetary policy shocks. IRENE Working Papers 24-03, IRENE Institute of Economic Research, University of Neuchatel.

Examples

```
Sig <- diag(2)
SigR <- diag(c(0.5, 0.5))
Psi <- matrix(c(1, 0.5), nrow = 2)
et <- matrix(rnorm(20), ncol = 2)
kfpredict(Sig = Sig, SigR = SigR, Psi = Psi, et = et)
```

`logdiff` *Log-difference transformation*

Description

Computes the first log-difference of a numeric vector or time series, i.e. `log(x_t) - log(x_{t-1})`. Gaps in dates are ignored; the difference is always taken with respect to the immediately preceding observation.

Usage

```
logdiff(TS)
```

Arguments

`TS`	A numeric vector or time series.

Value

A numeric vector of the same length as `TS`, with `NA` in the first position.

Examples

```
logdiff(c(1, 2, 4))
```

`normalize` *Normalize a time series*

Description

Standardizes a numeric vector to zero mean and unit standard deviation (z-score normalization), ignoring missing values.

Usage

```
normalize(x)
```

Arguments

`x`	A numeric vector or time series.

Value

A numeric vector of the same length as `x`, with mean 0 and standard deviation 1.

Examples

```
normalize(c(1, 2, 3, 4, 5))
```

`plot2irf` *Plot and compare impulse responses from two estimation approaches*

Description

Produces a panel of IRF plots comparing two sets of impulse responses. Each panel shows the point estimates and confidence bands at level `ci` for both approaches. Returns a list of `ggplot` objects (one per variable-shock combination) that can be arranged with e.g. `cowplot::plot_grid()`.

Usage

```
plot2irf(IRF1, IRF1se, IRF2, IRF2se, HTick, Labels, ci = 0.9)
```

Arguments

`IRF1`	Array (H x N x E) of impulse responses for the first approach. Dimnames on the first dimension must be the horizon labels (numeric).
`IRF1se`	Array (H x N x E) of standard errors for the first approach (set to 0 to suppress confidence intervals).
`IRF2`	Array (H x N x E) of impulse responses for the second approach.
`IRF2se`	Array (H x N x E) of standard errors for the second approach (set to 0 to suppress confidence intervals).
`HTick`	Integer. Step size for x-axis tick marks.
`Labels`	Character vector of length N. Variable names used as panel titles.
`ci`	Numeric scalar in (0, 1). Confidence level for the shaded bands. Defaults to `0.90` (90% CI).

Value

A list of `ggplot` objects, one per variable-shock combination, ordered by shock (outer loop) then variable (inner loop).

Examples

```
irf1 <- array(c(1, 0.5, 0.2, 0.1), dim = c(4, 1, 1))
dimnames(irf1)[[1]] <- 0:3
irf2 <- irf1 * 0.8
se <- array(0.1, dim = dim(irf1), dimnames = dimnames(irf1))
plot2irf(irf1, se, irf2, se * 0, HTick = 1, Labels = "Output")
```

---

plotirf *Plot impulse responses from a single estimation approach*

Description

Produces a panel of IRF plots for a single set of impulse responses, with optional confidence bands. Returns a list of `ggplot` objects (one per variable-shock combination) that can be arranged with e.g. `cowplot::plot_grid()`.

Usage

```
plotirf(IRFest, IRFse = NULL, HTick, Labels, ci = c(0.9, 0.95))
```

Arguments

`IRFest`	Array (H x N x E) of impulse response point estimates. Dimnames on the first dimension must be the horizon labels (numeric).
`IRFse`	Array (H x N x E) of standard errors, or `NULL` to suppress confidence bands.
`HTick`	Integer. Step size for x-axis tick marks.
`Labels`	Character vector of length N. Variable names used as panel titles.
`ci`	Numeric vector of confidence levels in (0, 1). One shaded band is drawn per level, with wider bands rendered at lower opacity. Defaults to `c(0.90, 0.95)`.

Value

A list of `ggplot` objects, one per variable-shock combination, ordered by shock (outer loop) then variable (inner loop).

Examples

```
irf <- array(c(1, 0.5, 0.2, 0.1), dim = c(4, 1, 1))
dimnames(irf)[[1]] <- 0:3
se <- array(0.1, dim = dim(irf), dimnames = dimnames(irf))
plotirf(irf, se, HTick = 1, Labels = "Output")
```

plotpval — *Plot p-values across horizons*

Description

Produces a panel of p-value plots across IRF horizons, with optional horizontal reference lines at user-supplied significance levels. Returns a list of `ggplot` objects (one per variable-shock combination) that can be arranged with e.g. `cowplot::plot_grid()`.

Usage

```
plotpval(pvals, HTick, Labels, sigLevels = c(0.05, 0.1))
```

Arguments

`pvals`	Array (H x N x E) of p-values. Dimnames on the first dimension must be the horizon labels (numeric).
`HTick`	Integer. Step size for x-axis tick marks.
`Labels`	Character vector of length N. Variable names used as panel titles.
`sigLevels`	Numeric vector of significance levels at which to draw horizontal reference lines. The first value is drawn as a solid line, the second as dashed, and any further values as dotted. Defaults to `c(0.05, 0.10)`.

Value

A list of `ggplot` objects, one per variable-shock combination, ordered by shock (outer loop) then variable (inner loop).

Examples

```
pvals <- array(c(0.2, 0.08, 0.04), dim = c(3, 1, 1))
dimnames(pvals)[[1]] <- 0:2
plotpval(pvals, HTick = 1, Labels = "Output")
```

proxyiv — *Estimate impulse responses via proxy-IV local projections*

Description

Estimates impulse response functions (IRFs) using user-provided external instruments (proxies) combined with local projections (Jorda, 2005). The proxy variables serve directly as instruments for the endogenous shock variables. Optionally imposes recursive zero restrictions across shock dimensions and supports deterministic controls following the same conventions as `hetiv()`.

Usage

```
proxyiv(
  y,
  O,
  Z,
  X = NULL,
  Ind,
  P,
  H,
  E = 1,
  norm = 1,
  cum = FALSE,
  Hstep = 1,
  cov_type = "HC0",
  recursive = FALSE,
  details = FALSE
)
```

Arguments

`y` Numeric matrix of stationary outcome variables (T x N). The effect on the first variable in each dimension is normalized to `norm` at horizon 0. These variables are also used as the endogenous regressors instrumented by the columns of `Z`.

`O` Numeric matrix of information set variables (T x M). May be identical to `y`. Included as lags 1 through `P`.

`Z` Numeric matrix of external instruments (T x E). Column `e` is used as the proxy for shock dimension `e`. Missing values on control days (`Ind == 0`) are treated as contaminated and excluded from estimation. Missing values on policy days (`Ind == 1`) remain labelled as events for residual-based shock prediction, but are necessarily dropped from IV regressions that use the missing proxy.

`X` Numeric matrix of deterministic variables (T x K), or `NULL` (default). May include a constant, time trend, or seasonal dummies. Included as-is (no lags).

`Ind` Integer vector of length T, event indicator:

- `0` Control day (no event)
- `1` Policy day (event)
- `2` Contaminated control day (excluded from estimation)

`P` Integer. Maximum lag order for the information set. Set to `0` for no lags (regression on constant only).

`H` Integer. Maximum horizon (in periods) up to which IRFs are estimated.

`E` Integer. Number of shock dimensions to identify. Default `1`.

`norm` Numeric scalar. Normalize the impact response of the first variable to this value. Set to `1` for standard unit-effect normalization.

`cum` Logical scalar or vector of length N. If `TRUE` for variable `i`, the cumulative IRF is reported. A single value is recycled to all variables. Default `FALSE`.

`Hstep`	Integer. Step size between horizons. The default `1` estimates all horizons 0 through H - 1. Values greater than 1 estimate only the selected horizons. Default `1`.
`cov_type`	Covariance estimator for local-projection standard errors: `"HC0"` (default) for heteroskedasticity-robust standard errors or `"NW"` for Newey-West HAC standard errors. `"HC0"` is the default because Montiel Olea et al. (2025) show that heteroskedasticity-robust standard errors suffice for local-projection impulse responses under weak conditions, even though multi-step forecast errors are typically serially correlated. `"NW"` remains available as an optional HAC robustness check.
`recursive`	Logical. If `TRUE`, imposes recursive zero restrictions across shock dimensions: for shock `e > 1`, the variables and instruments from dimensions `1, ..., e-1` are added as controls. Default `FALSE`.
`details`	Logical. If `TRUE`, returns detailed results including IV model objects, OLS residuals, and covariance matrices. If `FALSE` (default), returns only impulse responses and standard errors (faster; use for bootstrap).

Details

For `E > 1`, identification can be order-dependent: the column order of `y` defines the endogenous shock variables and normalizations, the column order of `Z` assigns proxies to shock dimensions, and `recursive = TRUE` imposes restrictions in that order.

Value

A named list. Always contains:

`irf` Array (H x N x E) of estimated impulse responses.

`se` Array (H x N x E) of local-projection standard errors.

`Method` Character string `"Proxy-IV"`.

With `details = TRUE`, additionally contains:

`IVRes` List of `ivreg` model objects, one per horizon, variable, and shock dimension.

`OLSRes` List of OLS model objects used for residual-based covariance estimation, one per outcome variable.

`Obs` Data frame with observation counts: `Tp` (policy days), `Tc` (control days), `To` (contaminated days), `Tt` (labelled event or control days), and `Tiv` (complete observations available to the IV regression at the impact horizon).

`et` Data frame of OLS residuals on event days.

`Sig` Covariance matrix of residuals on event days.

`SigR` Covariance matrix of residuals on control days, or `NA` if unavailable.

`Psi` Impact matrix (N x E), equal to `irf[1, , ]`. By the package's indexing convention `HSeries` starts at 1, so the first LP uses `lead(y, 0)` (the contemporaneous value) and is labelled horizon 0; `irf[1, , ]` is therefore always the impact response.

`WeakData` Data frame of endogenous variables and instruments for the Lewis-Mertens (2025) weak instrument test.

References

Jorda, O. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1), 161-182.

Lewis, D. J. and Mertens, K. (2025). A robust test for weak instruments for 2SLS with multiple endogenous regressors. *Review of Economic Studies*, DOI: 10.1093/restud/rdaf103

Mertens, K. and Ravn, M. O. (2013). The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review*, 103(4), 1212-1247.

Montiel Olea, J. L., M. Plagborg-Moller, E. Qian, and C. K. Wolf (2025). Local projections or VARs? A primer for macroeconomists. *NBER Working Paper* No. 33871.

Stock, J. H. and Watson, M. W. (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *Economic Journal*, 128(610), 917-948.

Examples

```
set.seed(1)
y <- matrix(rnorm(80), ncol = 2)
Ind <- rep(0L, nrow(y))
Ind[seq(5, nrow(y), by = 5)] <- 1L
Z <- matrix(Ind * y[, 1] + rnorm(nrow(y)), ncol = 1)
res <- proxyiv(
  y = y, O = y, Z = Z, Ind = Ind, P = 1, H = 3,
  details = TRUE
)
dim(res$irf)
```

`simulatedata` *Simulate VAR data with heteroskedastic shocks and given parameters*

Description

Simulates data from a VAR(P) model with two types of structural shocks: regular shocks (always present) and event shocks (occurring every `Nevn` periods). The shock distribution can be normal, Student-t, or GARCH(1,1).

Usage

```
simulatedata(Phi, SigE, PsiE, PsiR, Nobs, Nbin, N, R, E, Nevn, P, eDist, seed)
```

Arguments

`Phi`	Array of VAR coefficient matrices, dimension `N x N x P`.
`SigE`	Variance of the event shocks. This applies to all E shocks.
`PsiE`	Impact matrix for event shocks, dimension `N x E`.
`PsiR`	Impact matrix for regular shocks, dimension `N x R`.

Nobs	Number of observations to retain after burn-in. The returned data have exactly `Nobs` rows.
Nbin	Number of burn-in observations. These are simulated to initialise the VAR but are discarded before returning.
N	Number of variables in the VAR.
R	Number of regular shocks.
E	Number of event shocks. For GARCH shock distributions (`eDist = c(alpha, beta)`), only `E = 1` is supported.
Nevn	Event frequency: an event shock occurs every `Nevn` periods. Set to `0` to suppress event shocks entirely (no heteroskedasticity).
P	VAR lag order.
eDist	Shock distribution. Use `0` for standard normal; a positive integer for Student-t with that many degrees of freedom; or a numeric vector `c(alpha, beta)` for GARCH(1,1) with ARCH effect `alpha` and persistence `beta`.
seed	Integer seed passed to `set.seed()` for reproducibility. Use `NA` to skip seeding.

Value

A named list with components:

y Simulated VAR data, dimension `Nobs x N` (burn-in discarded).

IndE Event indicator matrix, dimension `Nobs x 1`.

eR Simulated regular shocks, dimension `Nobs x R`.

eE Simulated event shocks, dimension `Nobs x E`.

e Composite structural shocks, dimension `Nobs x N`.

Phi VAR coefficient array (returned unchanged).

PsiE Event shock impact matrix (returned unchanged).

PsiR Regular shock impact matrix (returned unchanged).

SigE Event shock variance (returned unchanged).

Examples

```
N <- 2
Phi <- array(0, dim = c(N, N, 1))
Phi[, , 1] <- diag(c(0.4, 0.2))
simulatedata(
  Phi = Phi, SigE = 2, PsiE = matrix(c(1, 0.5), N, 1),
  PsiR = diag(N), Nobs = 20, Nbin = 5, N = N, R = N, E = 1,
  Nevn = 5, P = 1, eDist = 0, seed = 1
)
```

Index