Package: sfadv (via r-universe)

October 14, 2024

Version 1.0.1

Date 2017-06-19

Title Advanced Methods for Stochastic Frontier Analysis

Maintainer Yann Desjeux <yann.desjeux@inra.fr>

Description Stochastic frontier analysis with advanced methods. In
particular, it applies the approach proposed by Latruffe et al.
(2017) <DOI:10.1093/ajae/aaw077> to estimate a stochastic
frontier with technical inefficiency effects when one input is
endogenous.

Depends R (>= 2.10)

Imports gmm, minpack.lm

License GPL-3

Encoding UTF-8

LazyData true

NeedsCompilation no

Author Yann Desjeux [aut, cre], Laure Latruffe [aut], Alain Carpentier
[ctb]

Date/Publication 2017-06-19 11:45:43 UTC

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

RemoteUrl https://github.com/cran/sfadv

RemoteRef HEAD

RemoteSha 25a74149c41e1a70de5e11c7bd8304d6ffc3b002

Contents

Farms   *Data set of farm accountancy data*

Description

The Farms data frame contains simulated data for 2,500 dairy farms.

Usage

    Farms

Format

This data frame contains the following 14 variables:

farm_output	Farm total output, in constant Euros.
agri_land	Farm agricultural area, in hectares.
tot_lab	Total labour used on the farm, in hours.
tot_asset	Total assets of the farm (excluding land), in constant Euros.
LFA	Dummy variable indicating whether the farm is located within a Less Favoured Area (1) or not (0).
hired_lab	Ratio of the farm hired labour to total labour (tot_lab).
rented_land	Ratio of the farm rented land to agricultural area (agri_land).
debt_asset	Ratio of the farm total debts to total assets.
costs	Costs of variable inputs used on the farm, in constant Euros.
subs	Amount of the farm production subsidies received per hectare of agricultural area (agri_land), in constant Euros.
region	Factor variable representing the administrative region of the farm.
milkprice	Average farm milk price per ton of milk, in constant Euros.
price_ind	National yearly price index of variable inputs.
T	Time trend.

Source

Simulated farmers' accountancy data

Examples

```
head(Farms)
str(Farms)
summary(Farms)

lm.output <- lm(farm_output ~ agri_land + tot_lab + tot_asset + costs, data = Farms)
summary(lm.output)
```

---

sfaendog	*Stochastic frontier analysis with technical inefficiency effects and endogeneity of one input*

Description

It implements a Method of Moments (MM) estimation of stochastic production frontier with explanatory variables influencing technical inefficiency (*i.e.* the technical inefficiency effects) and accounting for one single endogenous input.

Usage

```
sfaendog(y, x.exo, x.endo, c.var, ineff, inst, data, nls.algo = c("GN", "LM"),
      gmm.kernel = c("Bartlett", "Quadratic Spectral", "Truncated", "Parzen",
      "Tukey-Hanning"), gmm.optim = c("BFGS", "Nelder-Mead", "CG", "SANN"),
      maxiter = 100)
```

Arguments

y	A character string specifying the name of the dependent variable of the production function (output). Values of the y variable will be internally coerced to their *logged* values.
x.exo	A vector of strings containing the names of the exogenous inputs of the production function. Values of the x.exo variables will be internally coerced to their *logged* values.
x.endo	A character string specifying the name of the input of the production function considered as endogenous. Values of the the x.endo variable will be internally coerced to their *logged* values.
c.var	A vector of strings containing the names of non-input variables influencing the output (production environment conditions, time trend, etc.).
ineff	A vector of strings containing the names of the variables influencing technical inefficiency.
inst	A vector of strings containing the names of the external instrumental variables.
data	A data frame containing the variables called in y, x.exo, x.endo, c.var, ineff and inst.
nls.algo	A character string specifying the algorithm used in the Non-linear Least Squares (NLS) estimation (see Estimation 'recipe'). The default algorithm is the Gauss-Newton algorithm ("GN"). The other possible value is "LM" for the Levenberg-Marquardt algorithm. Levenberg-Marquardt being a more robust variation of Gauss-Newton algorithm, specifiying nls.algo = "LM" may help in case of non convergence.
gmm.kernel	A character string specifying the type of kernel used to compute the covariance matrix of the vector of sample moment conditions in the Generalised Method of Moments (GMM) estimations (see Estimation 'recipe'). The default value is "Bartlett" kernel. Other possible options are "Quadratic Spectral", "Truncated", "Parzen", or "Tukey-Hanning" (see gmm for more details).

gmm.optim	A character string specifying the optimisation method to be applied in the GMM estimations (see Estimation 'recipe'). The default value ("BFGS") is the quasi-Newton method. Other possibles methods are "Nelder-Mead", "CG", "SANN" (see optim for more details).
maxiter	A positive integer specifying the maximum number of iterations allowed for the NLS and GMM estimations. The default value is 100.

Details

The function sfaendog() implements the 4-step 'recipe' detailed in Latruffe *et al.* (2017, p.788).

The use of basic formula operators generally involved in model formulae, such as ":", "*", or "^", are allowed in x.exo, x.endo, c.var, ineff and inst.
As in function formula, the function I() can also be used to inhibit the interpretation of operators such as "+", "−", "*" and "^" as formula operators, so that they are used as arithmetical operators.

- Stochastic production frontier model with a single endogenous input:

sfaendog() assumes a Cobb-Douglas functional form for the production frontier.

In this case, the stochastic frontier production model is written as:

$$\ln y = \boldsymbol{\alpha'}_0 \ln \mathbf{x} - \eta \exp(\boldsymbol{\theta'}_0 \mathbf{z}) + v$$

where ln is the log; y is the observed output; $\boldsymbol{\alpha}_0$ and $\boldsymbol{\theta}_0$ are vectors of parameters to be estimated; $\mathbf{x}$ is a vector containing the inputs as well as a constant term one; v is a random term which accounts for the effects of unobserved heterogeneity across observations and for stochastic events affecting the production process; $\eta \exp(\boldsymbol{\theta'}_0 \mathbf{z})$ is a non-negative term accounting for the presence of technical inefficiency; $\mathbf{z}$ is a vector of variables influencing technical inefficiency (the inefficiency effects), including a constant term one; and η is a positive random term with mean one.

If all inputs are exogenous, the above equation can be estimated by NLS.

sfaendog() accounts for the endogeneity of one input, with a MM estimator based on Chamberlain's (1987) 'efficient instruments'. In the case of endogeneity of one input, the stochastic frontier production model defined above can be rewritten as:

$$\ln y = \boldsymbol{\alpha}_{x,0} \ln \mathbf{x}_x + \alpha_{e,0} \ln x_e - \eta \exp(\boldsymbol{\theta'}_0 \mathbf{z}) + v$$

where $\mathbf{x}_x$ is the vector containing the exogenous inputs and the constant term one; x_e is the endogenous input; and the subscript 0 denotes the 'true' parameters value. The vector of exogenous variables is denoted by $\mathbf{w} = (\ln \mathbf{x}_x, \mathbf{q}, \mathbf{z})$ where $\mathbf{q}$ is the vector of external instrumental variables. Assuming that $E[v|\mathbf{w}] = 0$, and that η and $(v, \mathbf{w})$ are independent, the stochastic frontier production model can be rewritten as:

$$\ln y = \boldsymbol{\alpha}_{x,0} \ln \mathbf{x}_x + \alpha_{e,0} \ln x_e - \exp(\boldsymbol{\theta'}_0 \mathbf{z}) + e \quad \text{with} \quad E[e|\mathbf{w}] = 0$$

where the error term e is defined as $e \equiv e(\boldsymbol{\delta}_0) = \ln y - \boldsymbol{\alpha}_{x,0} \ln \mathbf{x}_x - \alpha_{e,0} \ln x_e + \exp(\boldsymbol{\theta'}_0 \mathbf{z})$

- Estimation 'recipe':

The estimation 'recipe' detailed in Latruffe *et al.* (2017, p.788) and implemented through `sfaendog()` consists in the following four steps:

- **Step1:** Ordinary Least Squares (OLS) estimation of $\ln x_e$ on a set of external instrumental variables (**q**) and all exogenous variables included in the stochastic frontier production model. This generates the predicted values of $\ln x_e$ to be used as instruments in `Step3`.
  The strength of the external instrumental variables (**q**) is measured by testing, with a Fisher test, the nullity of the parameters related to these external instrumental variables.
  This `Step1` corresponds to *Step 1* of the estimation 'recipe' presented in Latruffe *et al.* (2017, p.788).

- **Step2:** NLS estimation of the stochastic frontier production model, to compute a **non consistent and non efficient** estimator to be used in `Step3`.
  `Step2` consists in three sub-steps (not detailed as such in Latruffe *et al.*'s (2017, p.788) recipe):
  - *Step2a.i*: OLS estimation of a production model with the output as the dependent variable, and the explanatory variables being the inputs and the non-input variables influencing the output.
    This provides predicted residuals to be used in `Step2a.ii`, and parameters to be used as starting values for the variables in the production part of the stochastic frontier production model in `Step2b`.
  - *Step2a.ii*: OLS estimation of the predicted residuals of `Step2a.i` on the variables influencing technical inefficiency.
    This provides parameters to be used as starting values for the variables in the inefficiency effect part of the stochastic frontier production model in `Step2b`.
  - *Step2b*: NLS estimation of the stochastic frontier production model, using starting values obtained from `Step2a.i` and `Step2a.ii`.
    This sub-step corresponds to *Step 2* of the estimation 'recipe' presented in Latruffe *et al.* (2017, p.788).

- **Step3:** Estimation of the stochastic frontier production model, using the predicted values of the endogenous input obtained from `Step1` and using the parameters obtained from `Step2b` as starting values. The estimation is done with GMM.
  `Step3` computes a **consistent but non efficient** estimator to be used in `Step4`.
  This `Step3` corresponds to *Step 3* of the estimation 'recipe' presented in Latruffe *et al.* (2017, p.788).

- **Step4:** Estimation of the stochastic frontier production model, using the predicted values of the endogenous input obtained from `Step1` the parameters obtained from `Step3` as starting values. This estimation is done with GMM.
  `Step4` computes a **consistent and efficient** estimator, and returns the final results of the estimation of the stochastic frontier production model.
  This `Step4` corresponds to *Step 4* of the estimation 'recipe' presented in Latruffe *et al.* (2017, p.788).

Please note that, the applicability of default options in `nls.algo`, `gmm.kernel`, and `gmm.optim` is

highly data-dependent and the user may have to play around with different options.

Value

sfaendog returns a list of class 'sfaendog'.

The object of class 'sfaendog' is a list containing at least the following components:

Step1	Results of Step1 (see Estimation 'recipe'), inheriting from class 'lm' for which details on components are given in lm.
FishTest	The list of external instrumental variables (inst) used in the specification, as well as Fisher test results showing the strength of these external instrumental variables.
Step2	Results of Step2b (see Estimation 'recipe'), inheriting from class 'nls' for which details on components are given in nls.
Step3	Results of Step3 (see Estimation 'recipe'), inheriting from class 'gmm' for which details on components are given in gmm.
Step4	Results of Step4 (see Estimation 'recipe'), inheriting from class 'gmm' for which details on components are given in gmm.

The function summary is used to obtain and print a summary of the results.

Author(s)

Yann Desjeux, Laure Latruffe

References

Chamberlain G. (1987). Asymptotic Efficiency in Estimation with Conditional Moment Restrictions. *Journal of Econometrics*, **34**(3), 305–334. https://doi.org/10.1016/0304-4076(87)90015-7

Latruffe L., Bravo-Ureta B.E., Carpentier A., Desjeux Y., and Moreira V.H. (2017). Subsidies and Technical Efficiency in Agriculture: Evidence from European Dairy Farms. *American Journal of Agricultural Economics*, **99**(3), 783–799. https://doi.org/10.1093/ajae/aaw077

See Also

summary for creating and printing summary results.

Examples

```
## Not run:
  y <- "farm_output"
  x.exo <- c("agri_land", "tot_lab", "tot_asset")
  x.endo <-  "costs"
  c.var <- c("LFA", "T", "I(T^2)")
  ineff <- c("hired_lab", "rented_land", "debt_asset", "subs", "region", "region:T")
  inst <- c("milkprice", "I(milkprice^2)", "price_ind", "milkprice:region")

  RES <- sfaendog(y, x.exo, x.endo, c.var, ineff, inst, data=Farms)
```

```
    summary(RES)
## End(Not run)
```

---

summary   *Summary method for object of class "sfaendog"*

Description

It creates and displays the results of an estimation returned by sfaendog.

Usage

```
## S3 method for class 'sfaendog'
summary(object, all = FALSE, ...)

## S3 method for class 'summary.sfaendog'
print(x, digits = max(3, getOption("digits") - 3),
     signif.stars = getOption("show.signif.stars"), ...)
```

Arguments

object	An object of class 'sfaendog' returned by the function sfaendog.
x	An object of class 'summary.sfaendog' returned by the function summary.
all	Logical. If FALSE (default), only 'Step1' results, Fisher test results for the strength of external instrumental variables, and 'Step4' results are returned. If all = TRUE, summary of all 'Steps' in sfaendog are returned.
signif.stars	Logical. If TRUE, significance stars are printed.
digits	The minimum number of significant digits to be printed in values.
...	Currently not used.

Details

print.summary.sfaendog tries to be smart about formatting the coefficients, standard errors, etc. and additionally gives significance stars if signif.stars is TRUE.

Value

A list of class 'summary.sfaendog' is returned, containing the number of observations, the parameter estimates, standard errors, t-values and p-values.

More precisely:

Step1	The list of summary statistics of 'Step1' OLS estimation of the endogenous input.
FishTest	A list containing the names of the variables used as external instrumental variables in the estimation, and the results (*i.e.* F-value and Pr(>F)) of the Fisher test on the strength of these external instrumental variables.

Step2	The list of summary statistics of 'Step2' NLS estimation of the stochastic frontier production model.
Step3	The list of summary statistics of 'Step3' GMM estimation of the stochastic frontier production model. The name of the endogenous input (wrapped with '*') is also returned.
Step4	The list of summary statistics of 'Step4' GMM estimation of the stochastic frontier production model. In addition, it also returns the name of the endogenous input (wrapped with '*' in the result table) indicating that the predicted value from Step1 is used.

Step2 and Step3 are returned only if all = TRUE.

In Step2 to Step4 results, InterceptX is the constant term in the production part of the stochastic frontier production model, while InterceptZ is the constant term in the inefficiency effect part.

Information on convergence is also provided with Step3 and Step4 results, where Convergence code = 0 indicates successful completion (see optim).

Author(s)

Yann Desjeux, Laure Latruffe

See Also

sfaendog, for the model fitting function.

Index