

Package ‘endoSwitch’

October 13, 2022

Type Package

Title Endogenous Switching Regression Models

Version 1.0.0

Description Maximum likelihood estimation of endogenous switching regression models from Heckman (1979) <[doi:10.2307/1912352](https://doi.org/10.2307/1912352)> and estimation of treatment effects.

Encoding UTF-8

License GPL (>= 2)

LazyData true

RoxygenNote 7.0.2

Depends R (>= 3.6.0)

URL <https://github.com/cbw1243/endoSwitch>

Imports data.table (>= 1.12.2), maxLik (>= 1.3-6), stats (>= 3.6.1),
msm (>= 1.6.7)

NeedsCompilation no

Author Bowen Chen [aut, cre] (<<https://orcid.org/0000-0003-0370-2756>>),
Seong Yun [aut],
Benjamin Gramig [aut]

Maintainer Bowen Chen <bwchen0719@gmail.com>

Repository CRAN

Date/Publication 2020-02-21 10:20:02 UTC

R topics documented:

calcPar	2
endoSwitch	2
endoSwitch2Stage	4
ImpactData	5
summary.endoSwitch	6
treatmentEffect	7

Index	8
--------------	----------

calcPar

Endogenous switching regression

Description

This R function provides estimates of original distribution parameters (sigma and rho) from the estimates of transformed distribution parameters using the delta method.

Usage

```
calcPar(Results)
```

Arguments

Results A maxLik object that is estimated by the endoSwitch function.

Value

A matrix that reports the estimates of original distribution parameters. sigma is standard deviation, and rho is correlation coefficient.

endoSwitch

Endogenous Switching Regression Models

Description

This is the main interface for the endoSwitch package to estimate the endogenous switching regression models (Heckman, 1979).

Usage

```
endoSwitch(  
  data,  
  OutcomeDep,  
  SelectDep,  
  OutcomeCov,  
  SelectCov,  
  Weight = NA,  
  treatEffect = TRUE,  
  method = "BFGS",  
  start = NULL,  
  verbose = FALSE,  
  ...  
)
```

Arguments

<code>data</code>	a data frame. Data for running the regression analysis.
<code>OutcomeDep</code>	character. Dependent variable in the outcome equation.
<code>SelectDep</code>	character. Dependent variable in the Selection model. The variable must be binary (0 or 1).
<code>OutcomeCov</code>	character vector. Covariates in the outcome equation.
<code>SelectCov</code>	character vector. Covariates in the selection equation.
<code>Weight</code>	optional character. Name of the weight variable in the dataset, or NA (equal weight).
<code>treatEffect</code>	TRUE/FALSE. If TRUE, average treatment effects will be calculated and returned. If FALSE, expected outcome values will be calculated and returned.
<code>method</code>	character. Maximization method to be used. The default is "BFGS" (for Broyden-Fletcher-Goldfarb-Shanno). Other methods can also be used. See maxLik .
<code>start</code>	optional numeric vector. Used as initial values of parameters for maximization purpose. If NULL, the coefficient estimates from the two-stage estimation will be used.
<code>verbose</code>	TRUE/FALSE. Choose to show the status of optimization or not.
<code>...</code>	Other parameters to be passed to the selected maximization routine. See maxLik .

Details

This function estimates the endogenous switching regression model using the full maximum likelihood estimation method. In this model, a selection equation sorts observation units over two different regimes (e.g., treated and not-treated, or adopter and non-adopter), and two outcome equations that determine the outcome. Estimation of the model relies on joint normality of the error terms in the three-equation system (the selection equation plus two outcome equations). The model is estimated by maximizing the joint likelihood function that is provided in Lokshin and Sajaia (2004).

The `endoSwitch` uses the `maxLik` function in the `maxLik` package to do the optimization. The function automatically searches for starting values for maximization using the results from two-stage estimation following Maddala (1986, chapter 8). Though not recommended, users may provide starting values manually. Assume that you have M variables (including the constant) in the selection equation, and N variables (including the constant) in each outcome equation. Then you need $(M + 2*N + 4)$ starting values. The first M values are for the variables in the selection equation (last value for the constant), followed by N values for the outcome equation for the non-treated individuals (`SelectDep = 0`), and another N values for the outcome equation for the treated individuals (`SelectDep = 1`). The last four values are: σ in outcome equation for the non-treated, σ in outcome equation for the treated, ρ in outcome equation for the non-treated, ρ in outcome equation for the treated.

If `treatEffect = TRUE`, the `endoSwitch` function will report average treatment effects (for the treated or untreated) as well as heterogeneity effects. A detailed description of these effects is provided in Di Falco, Veronesi, and Yesuf (2011, p.837). If `treatEffect = FALSE`, the `endoSwitch` function will report expected outcome values in a list of two dataframes: dataframe `EYA1` reports actual (column `EY1.A1`) and counterfactual (column `EY0.A1`) expected outcome values for the treated; dataframe `EYA0` reports actual (column `EY0.A0`) and counterfactual (column `EY1.A0`) expected outcome values for the untreated.

Value

A list containing three elements. The first element is an object of class "maxLik", which includes parameters in the selection equation, parameters in the outcome equations, and the transformed distributional parameters (parameters are transformed to facilitate maximization, as recommended by Lokshin and Sajaia (2004)). The second element contains the estimates of original distributional parameters (transformed back via the delta method). The third element contains a table reporting average treatment effects or a list of expected outcome values, depending on users' choice of treatEffect.

References

Lokshin, Michael, and Roger B. Newson. "Impact of Interventions on Discrete Outcomes: Maximum Likelihood Estimation of the Binary Choice Models with Binary Endogenous Regressors." *Stata Journal* 11, no. 3 (2011): 368–85.

Heckman, James J. "Sample Selection Bias as a Specification Error." *Econometrica* 47, no. 1 (1979): 153–61. <https://doi.org/10.2307/1912352>.

Maddala, G. S. "Limited-Dependent and Qualitative Variables in Econometrics." Cambridge Books. Cambridge University Press, 1986.

Di Falco, Salvatore, Marcella Veronesi, and Mahmud Yesuf. "Does Adaptation to Climate Change Provide Food Security? A Micro-Perspective from Ethiopia." *American Journal of Agricultural Economics* 93, no. 3 (2011): 829–46. <https://doi.org/10.1093/ajae/aar006>.

Abdulai, Abdul Nafeo. "Impact of Conservation Agriculture Technology on Household Welfare in Zambia." *Agricultural Economics* 47, no. 6 (2016): 729–41. <https://doi.org/10.1111/agec.12269>.

Examples

```
data(ImpactData) # Data are from Abdulai (2016)
OutcomeDep <- 'Output'
SelectDep <- 'CA'
OutcomeCov <- c('Age')
SelectCov <- c('Age', 'Perception')
endoReg <- endoSwitch(ImpactData, OutcomeDep, SelectDep, OutcomeCov, SelectCov)

summary(endoReg) # Summarize the regression results
```

endoSwitch2Stage *Endogenous switching regression*

Description

This function estimates the endogenous switching regression model via two-stage estimations (Maddala, 1986)

Usage

```
endoSwitch2Stage(data, OutcomeDep, SelectDep, OutcomeCov, SelectCov)
```

Arguments

data	a data frame. Data for running the regression analysis.
OutcomeDep	character. Dependent variable in the outcome equation.
SelectDep	character. Dependent variable in the Selection model. The variable must be binary (0 or 1).
OutcomeCov	character vector. Covariates in the outcome equation.
SelectCov	character vector. Covariates in the selection equation.

Details

The first stage uses a probit model to estimate the selection equation. The second stage uses ordinary least squares including the inverse mills ratios computed from the first stage estimation results to estimate the outcome equations.

Value

A list containing regression results.

References

Maddala, G. S. "Limited-Dependent and Qualitative Variables in Econometrics." Cambridge Books. Cambridge University Press, 1986.

Examples

```
data(ImpactData)
OutcomeDep <- 'Output'
SelectDep <- 'CA'
OutcomeCov <- c('Age')
SelectCov <- c('Age', 'Perception')
Results <- endoSwitch2Stage(ImpactData, OutcomeDep, SelectDep, OutcomeCov, SelectCov)
# First stage regression results
summary(Results$FirstStageReg)
# Second stage regression results -- non-adopter
summary(Results$SecondStageReg.0)
# Second stage regression results -- adopter
summary(Results$SecondStageReg.1)
```

ImpactData

A dataset on adoption of conservation agriculture in Zambia

Description

A dataset on adoption of conservation agriculture in Zambia

Usage

```
data(ImpactData)
```

Format

An object of class `data.table` (inherits from `data.frame`) with 408 rows and 31 columns.

Source

Abdulai, Abdul Nafeo. "Impact of Conservation Agriculture Technology on Household Welfare in Zambia." *Agricultural Economics* 47, no. 6 (2016): 729–41. <https://doi.org/10.1111/agec.12269>.

References

Abdulai, Abdul Nafeo. "Impact of Conservation Agriculture Technology on Household Welfare in Zambia." *Agricultural Economics* 47, no. 6 (2016): 729–41. <https://doi.org/10.1111/agec.12269>.

Examples

```
data(ImpactData)
summary(ImpactData)
```

<code>summary.endoSwitch</code>	<i>Summarize the endogenous switching regression results.</i>
---------------------------------	---

Description

Summarize the endogenous switching regression results.

Usage

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

Arguments

<code>object</code>	Estimated endogenous switching regression model.
<code>...</code>	Other elements.

Value

A list containing the key regression results.

treatmentEffect	<i>Endogenous switching regression</i>
-----------------	--

Description

This function calculates average treatment effects and heterogeneity effects from an estimated endogenous switching regression model.

Usage

```
treatmentEffect(  
  Results,  
  data,  
  OutcomeDep,  
  SelectDep,  
  OutcomeCov,  
  SelectCov,  
  treatEffect  
)
```

Arguments

Results	Estimated endogenous switching regression model.
data	a data frame. Data for running the regression analysis.
OutcomeDep	character. Dependent variable in the outcome equation.
SelectDep	character. Dependent variable in the Selection model. The variable must be binary (0 or 1).
OutcomeCov	character vector. Covariates in the outcome equation.
SelectCov	character vector. Covariates in the selection equation.
treatEffect	TRUE/FALSE. Show average treatment effects or expected outcome values.

Value

A data table that reports the average treatment effects or a list of two tables reporting expected outcome values.

Index

* datasets

ImpactData, [5](#)

calcPar, [2](#)

endoSwitch, [2](#)

endoSwitch2Stage, [4](#)

ImpactData, [5](#)

maxLik, [3](#)

summary.endoSwitch, [6](#)

treatmentEffect, [7](#)