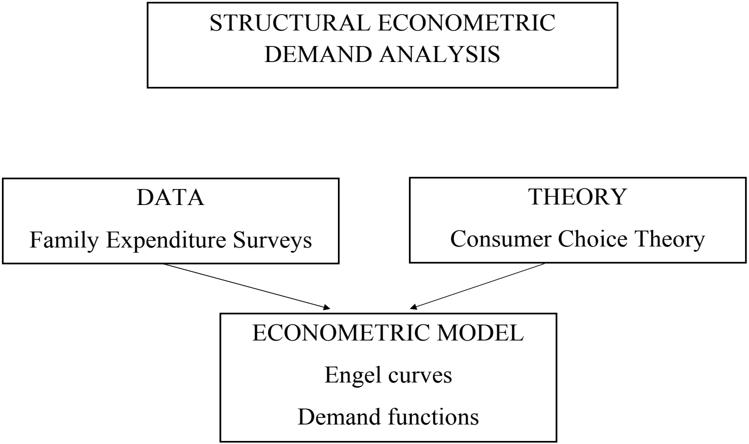
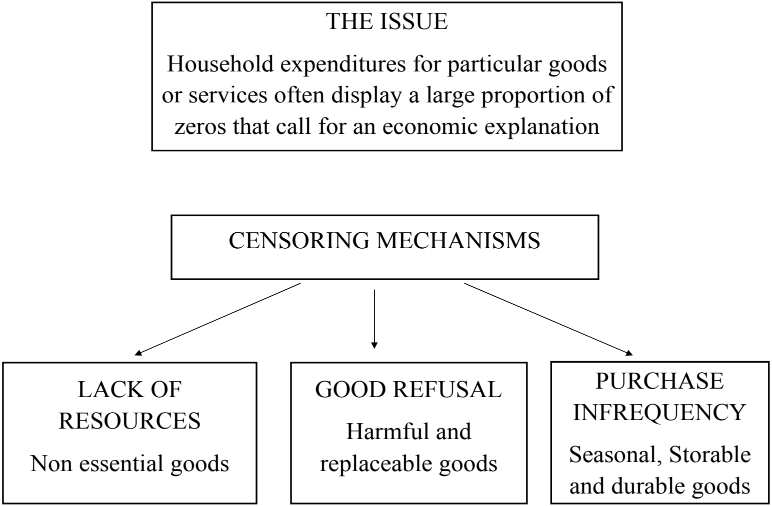
Multiple hurdle models in R : The mhurdle Package

Fabrizio CARLEVARO
University of Geneve

Yves CROISSANT
University of Lyon 2

Stéphane HOAREAU
University of La Réunion


STRUCTURAL ECONOMETRIC DEMAND ANALYSIS
DATA
Family Expenditure Surveys
THEORY
Consumer Choice Theory
ECONOMETRIC MODEL
Engel curves
Demand functions


THE ISSUE
Household expenditures for particular goods or services often display a large proportion of zeros that call for an economic explanation
CENSORING MECHANISMS
LACK OF RESOURCES
Non essential goods
GOOD REFUSAL
Harmful and replaceable goods
PURCHASE INFREQUENCY
Seasonal, Storable and durable goods

Seminal papers

Single equation models

Model	Author	censoring mechanism
Tobit Model	TOBIN (1958)	Lack of resources
Single hurdle model	CRAGG (1971)	Good refusal
Double hurdle model	CRAGG (1971)	Good refusal and
	BLUNDELL (1987)	lack of resources
P-Tobit model	DEATON and IRISH (1984)	Purchase infrequency and
		lack of resources

Systems of demand equations

Author	censoring mechanism
WALES and WOODLAND (1982)	Lack of resources
HANEMANN (1984)	Good refusal
ROBIN and MEGHIR (1992)	Purchase infrequency
BOIZOT, ROBIN and VISSER (2001)	Purchase infrequency

A comprehensive econometric framework : The triple-hurdle model

	Latent variable relation	**censoring rule**
Good selection	$y_1^\star = \beta_1^{'} x_1 + \epsilon_1$	$I_1 = \begin{cases} 1 & \text{if } y_1^\star > 0 \\ 0 & \text{if } y_1^\star \leq 0 \end{cases}$
Good consumption	$y_2^\star = \beta_2^{'} x_2 + \epsilon_2$	$I_2 = \begin{cases} 1 & \text{if } y_2^\star > 0 \\ 0 & \text{if } y_2^\star \leq 0 \end{cases}$
Good purchase	$y_3^\star = \beta_3^{'} x_3 + \epsilon_3$	$I_3 = \begin{cases} 1 & \text{if } y_3^\star > 0 \\ 0 & \text{if } y_3^\star \leq 0 \end{cases}$

(1)

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} ; \begin{bmatrix} 1 & \sigma\rho_{12} & \rho_{13} \\ \sigma\rho_{12} & \sigma^2 & \sigma\rho_{23} \\ \rho_{13} & \sigma\rho_{23} & 1 \end{bmatrix} \right) \qquad (2)$$

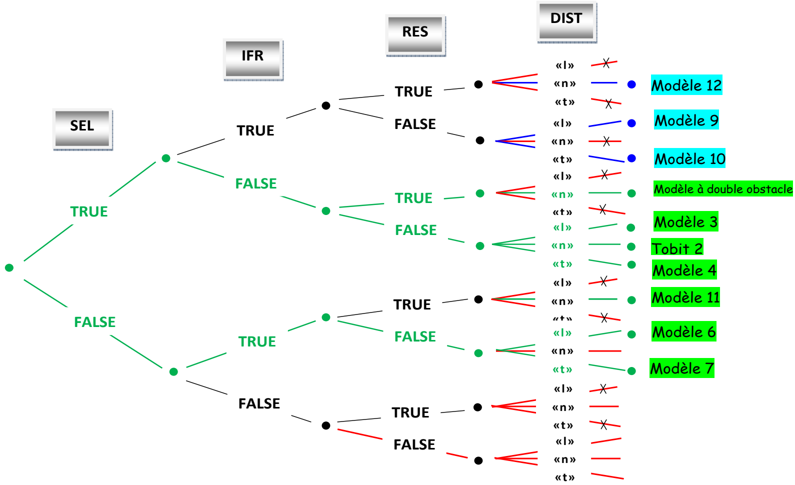
Observation equation : $y = \frac{y_2^\star}{P_{\{I_1 I_2 I_3 = 1\}}} I_1 I_2 I_3$

Using a priori information to specify particular hurdle models

- One or more censoring mechanisms are ineffective. If the “lack of resources” mechanism is inoperative, the desired consumption equation is respecified to enforce non negative consumption levels according to the following specifications :
  - log-normal : $\ln y_2^* \sim N(\beta_2' x_2, \sigma^2)$,
  - truncated normal : $y_2^* \sim NT_{R_+}(\beta_2' x_2, \sigma^2)$
- some or all correlation coefficients ρ_{12}, ρ_{13}, ρ_{23} may be set equal to zero entailing a partial or total independance between the censoring mechanisms

The full set of mhurdle models

× logical inconsistent model



mhurdle rationale

- Syntax :

```
mhurdle(formula, data, subset, weights, na.action,
      start = NULL, dist = c("l","t","n"), res = FALSE,
      sel = TRUE, ifr = FALSE, corr = FALSE, ...)
```

  two-parts formula `y ~ x1 + x2 | s1 + s2`

- Starting values
- Numerical optimisation methods : `maxLik`

An expenditure model for cigarettes in France (2001)

Survey time length excludes purchase infrequency as a relevant censoring mechanism

A priori relevant censoring mechanisms

- ✓ Good selection mechanism (SEL),
- ✓ Lack of ressources mechanism (RES),
- ✓ Good selection and lack of ressources mechanisms (SEL/RES).

Choice of explanatory variables

Good selection equation :

- Socio-professional status,
- Socio-demographic characteristics of family head
- Stress factors
- Health

Good consumption equation :

- Income
- Socio-professional status,
- Socio-demographic characteristics of family head
- Education-Training
- Financial situation
- Sports practice

Model validation and selection

Quality of fit measures

	RES	RES/SEL	SEL
Censored observations	0.34	0.32	0.35
Uncensored observations	0.03	0.06	0.13

Vuong test

	RES	RES/SEL	SEL
RES	♦		
RES/SEL	6.82	♦	
SEL	9.87	7.05	♦

$$H_0 : A \sim B \Leftrightarrow t_\alpha < V < t_{1-\alpha}$$
$$H_1 : A \succ B \Leftrightarrow V \rightarrow \infty$$
$$H_2 : A \prec B \Leftrightarrow V \rightarrow -\infty$$