A mixture of ordered probit models with endogenous assignment to two latent classes

J. L. Huismans[1] J. W. Nijenhuis[2] A. Sirchenko[3]

[1]Faculty of Economics and Business Economics
University of Amsterdam
[2]Faculty of Technical Computer Science
University of Twente
[3]School of Business and Economics
Maastricht University

June 10, 2022

Overview

We introduce the new **swopit** command that fits a switching ordered probit model with exogenous or endogenous class assignment to two latent regimes.

- produced as a software package for STATA
- model the decisions of the FOMC to change the federal funds rate

The **swopit** model allows for:

- two latent regimes which are estimated by binary probit
- a seperate ordered probit model for each regime
- custom starting values
- exogenous and endogenous switching
- optionally bootstrapped standard errors

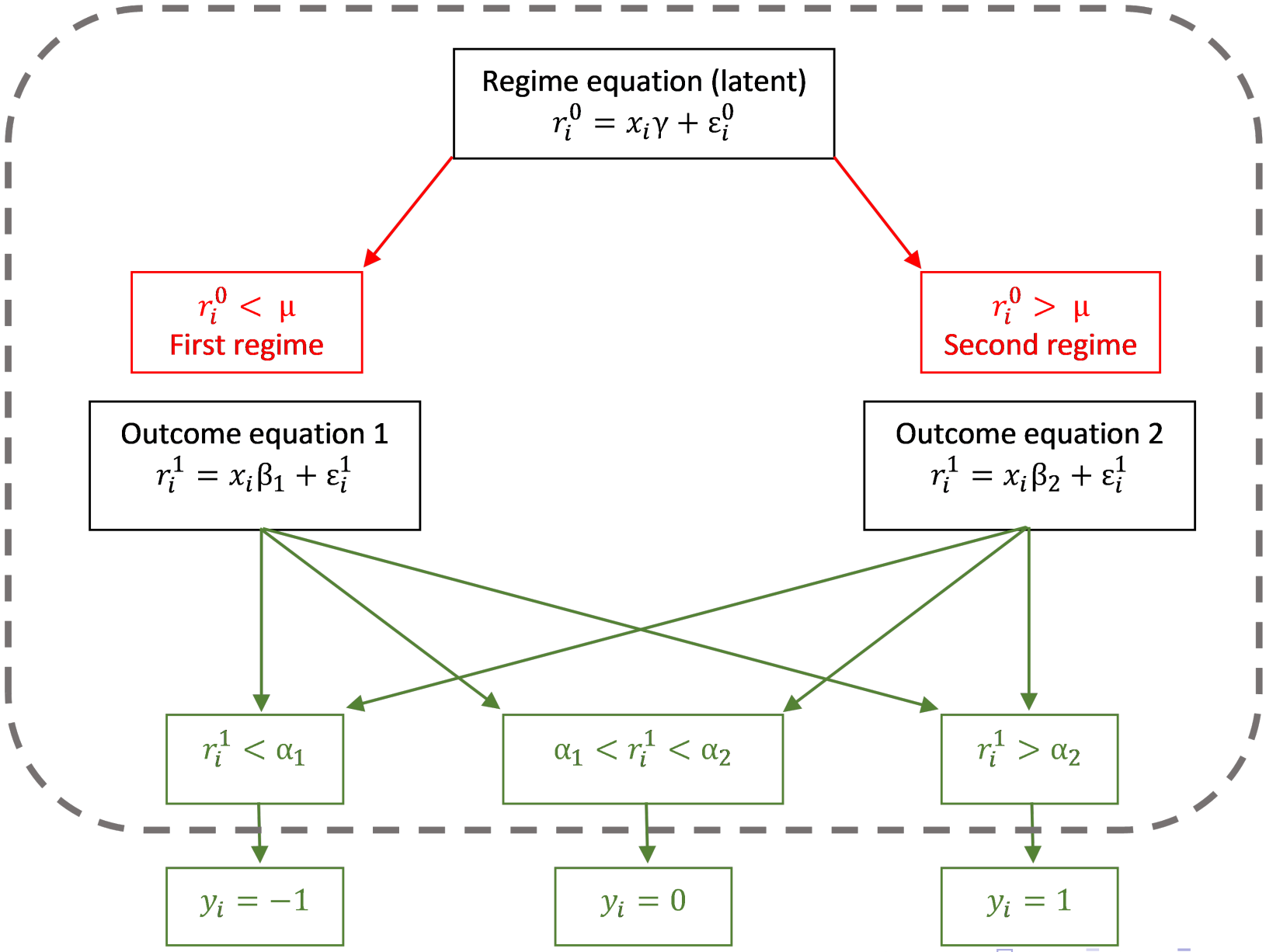
Previous work

The following is observed with regard to advancements in the field of endogenous switching:

- existing endogenous switching models deal mainly with continuous or binary data
- **fmm: oprobit** command can only fit switching ordered probit with exogenous switching
- most finite mixture models use the multinomial logistic distribution to model latent classes, where **swopit** uses binary probit
- zero-inflated model for endogenous class-assignment does not allow for all outcomes in each regime

The **swopit** model fits a switching ordered probit model with either exogenous or endogenous switching. There are no restrictions on outcomes in the two regimes.

Switching Ordered Probit Model



The probabilities

$$
\begin{aligned}
Pr(y_i = j | x_i^{all}) \\
&= \Phi_2(\mu - z_i\gamma; \alpha_{1,j} - x_{1,i}\beta_1; \rho_1) \\
&- \Phi_2(\mu - z_i\gamma; \alpha_{1,j-1} - x_{1,i}\beta_1; \rho_1) \\
&+ \Phi_2(-\mu + z_i\gamma; \alpha_{2,j} - x_{2,i}\beta_2; -\rho_2) \\
&- \Phi_2(-\mu + z_i\gamma; \alpha_{2,j-1} - x_{2,i}\beta_2; -\rho_2)
\end{aligned}
$$

Identification issue: By design, the model has two maxima in maximum likelihood estimation with complete overlap. Changing the signs of the correlation coefficients and the parameters in the class assignment models switches the regime labels and results in two identical likelihood values.

Monte Carlo Simulations (1)

In order to assess the performance of the model, the following Monte Carlo setup was used for a situation with two *latent* regimes and three classes:

1. 5 independent variables $g_j \sim 4 * \mathcal{N}(0, 1)$ were generated and used as covariates.
2. We simulated three different scenarios: 'complete overlap', 'partial overlap' and 'no overlap' between the covariates.
3. Parameters were calibrated to ensure that both *latent* classes have the same theoretical probability. The same holds for the outcomes.
4. To study consistency of the estimates, the simulation is applied for 250, 500, 1000 and 2000 observations.
5. For endogenous switching the correlation between unobservables is set to a positive number, for exogenous switching it is fixed at 0.

Monte Carlo Simulations - Parameters

Sample size	Class assignment: Covariates' overlap:	Exogenous none	Exogenous partial	Exogenous complete	Endogenous none	Endogenous partial	Endogenous complete
		Slope parameters γ, β_1 and β_2					
250	Bias, x10	5.5	5.6	5.4	4.8	4.5	4.7
500		1.7	1.7	1.7	1.6	1.6	1.6
1000		0.8	0.7	0.8	0.7	0.7	0.7
2000		0.3	0.3	0.3	0.3	0.3	0.3
250	RMSE, x10	18.2	18.6	16.6	13.3	12.0	12.9
500		4.6	4.3	4.6	4.3	4.1	4.4
1000		2.3	2.2	2.3	2.3	2.2	2.3
2000		1.5	1.4	1.5	1.4	1.4	1.5
250	Coverage rate (at 95% level), %	97.3	97.3	97.4	87.6	86.1	87.2
500		96.3	96.5	96.4	89.9	87.9	89.8
1000		95.5	95.6	95.4	92.3	90.9	92.6
2000		95.3	95.3	95.4	94.3	93.7	94.1
250	Bias of standard error estimator, x100	95.6	111.8	69.5	46.7	55.9	59.7
500		10.5	8.6	9.7	8.2	8.5	9.9
1000		1.7	1.5	1.6	1.9	1.8	1.5
2000		0.4	0.4	0.4	0.3	0.3	0.4

Monte Carlo Simulations - Thresholds

		Threshold parameters μ, α_1 and α_2					
250	Bias, x10	26.7	25.0	27.1	22.4	21.5	23.0
500		13.3	12.7	13.2	12.6	11.7	12.3
1000		8.4	8.2	8.3	8.2	7.9	8.2
2000		5.7	5.5	5.6	5.7	5.5	5.7
250	RMSE, x10	27.7	27.0	26.6	24.2	22.8	25.5
500		9.3	8.4	9.0	8.8	8.3	8.9
1000		4.9	4.7	4.8	4.8	4.7	4.8
2000		3.1	3.0	3.0	3.0	3.0	3.0
250	Coverage rate (at 95% level), %	97.2	97.0	97.4	86.0	87.5	85.3
500		96.2	95.9	96.3	88.2	88.4	88.1
1000		95.6	95.3	95.4	91.4	91.0	91.6
2000		95.2	95.1	95.3	94.0	93.5	93.6
250	Bias of standard error estimator, x100	120.2	123.5	106.5	108.4	100.3	119.1
500		18.2	12.5	15.5	17.1	17.4	20.0
1000		3.1	2.7	3.0	4.0	4.3	3.5
2000		0.9	0.7	0.7	0.6	0.8	0.8

Application

The model is applied to changes in the interest rate as imposed by the FOMC. First, a standard OP model is fitted. Then, the swopit model with exogenous switching and consecutively the swopit model with endogenous switching. The best model is chosen based on the Akaike Information Criterion (AIC). For the FOMC decisions a swopit model with exogenous switching outperforms the other methods.

Application FOMC decisions - Setup

- y_t: decision of the FOMC to decrease (-1), increase (1) or leave the federal funds rate target unchanged (0). This is the target variable.
- $house_t$: the Greenbook projection of the total number of new privately owned housing units started for the current quarter. Used as a measure of the housing market.
- gdp_t: the Greenbook projection of quarterly growth in the nominal gross domestic (before 1992: national) product for the current quarter, annualized percentage points.
- $spread_t$: the difference between the one-year treasury constant maturity rate and the effective federal funds rate, three business-day moving average.
- $bias_t$: the indicator that we constructed from the policy bias statements at the previous FOMC meeting: it equals 1 if the statement was asymmetric toward tightening, 0 if the statement was symmetric, and -1 if the statement was asymmetric toward easing.

Application FOMC decisions - Results

A mixture of ordered probit models with two latent classes

Latent class switching	= Exogenous
SE method	= OIM
Optimization method	= Newton Raphson
Number of observations	= 279
Log likelihood	= -132.6306
McFadden pseudo R2	= 0.4653
LR chi2(9)	= 230.8291
Prob > chi2	= 0.0000
AIC	= 287.2612
BIC	= 327.2046

y	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
Class membership model						
house	4.911729	1.386313	3.54	0.000	2.194606	7.628852
gdp	.6790068	.202835	3.35	0.001	.2814574	1.076556
/cut1	11.96884	2.980525	4.02	0.000	6.127117	17.81056
Outcome model (Class 1)						
bias	.6798781	.1525716	4.46	0.000	.3808433	.9789129
spread	1.819521	.240331	7.57	0.000	1.348481	2.290561
/cut1	-1.251123	.1463508	-8.55	0.000	-1.537965	-.9642805
/cut2	2.318764	.2520935	9.20	0.000	1.82467	2.812858
Outcome model (Class 2)						
bias	3.954417	1.183389	3.34	0.001	1.635017	6.273817
spread	14.12324	3.432332	4.11	0.000	7.395996	20.85049
/cut1	-16.28826	4.240216	-3.84	0.000	-24.59893	-7.977591
/cut2	1.448369	.6057136	2.39	0.017	.2611923	2.635546

(Post)estimation command(s)

Within the **swopit** command the user has the option to choose which covariates should be used in the regime equation, as well as both outcome equations. The user can choose his/her own starting values, which type of switching and limitations on the optimization algorithm.
After estimation, the following postestimation commands are available:

- swopitpredict
- swopitprobabilities
- swopitmargins
- swopitclassification

(Post)estimation command(s)

```
. swopitmargins, at(house=1.56 gdp=5.9 bias=1 spread=-0.41)

Evaluated at:
   house     gdp     bias     spread
  1.5600 5.9000   1.0000    -0.4100

Marginal effects of all variables on the probabilities of different outcomes
           |    Pr(y=-1)    Pr(y=0)    Pr(y=1)
-----------+-----------------------------------
     house |     -0.2210     0.2361    -0.0150
       gdp |     -0.0306     0.0326    -0.0021
      bias |     -0.0831     0.0706     0.0125
    spread |     -0.2223     0.1864     0.0359

Delta-method standard errors of marginal effects
           |    Pr(y=-1)    Pr(y=0)    Pr(y=1)
-----------+-----------------------------------
     house |      0.1140     0.1137     0.0143
       gdp |      0.0163     0.0163     0.0020
      bias |      0.0198     0.0228     0.0130
    spread |      0.0899     0.0994     0.0390
```

```
. swopitpredict, regimes tabstat

    Variable |         N      Mean        SD  Variance       Max        Min
-------------+--------------------------------------------------------------
  swopit_r_0 |       279     .8363     .2967     .08801         1   .0000455
  swopit_r_1 |       279     .1637     .2967     .08801         1          0
-------------------------------------------------------------------------------
```

(Post)estimation command(s)

```
. swopitprobabilities, at(house=1.5 gdp=8.9 bias=1 spread=-0.0633333)

Evaluated at:
   house     gdp     bias     spread
  1.5000  8.9000   1.0000    -0.0633

Predicted probabilities of different outcomes
   Pr(y=-1)     Pr(y=0)     Pr(y=1)
     0.0026      0.1186      0.8788

Delta-method standard errors of probabilities
   Pr(y=-1)     Pr(y=0)     Pr(y=1)
     0.0034      0.1241      0.1260
```

Contact & Questions

- janwillemnijenhuis@gmail.com
- `https://github.com/janwillemnijenhuis/Swopit`
- Paper is accepted for STATA Journal
- Additional options are in the help files