Local Maxima in the Estimation of the ZINB and Sample Selection models

J.M.C. Santos Silva
School of Economics, University of Surrey

23rd London Stata Users Group Meeting
7 September 2017

- Maximum likelihood (ML) estimators have many desirable properties.
- However, ML estimators also have problems:
  1. The ML estimator may not exist;
  2. The likelihood function may have multiple maxima.
- Stata makes available many ML estimators to users that may not be aware of these potential problems.

- Non-existence issues are reasonably well understood and solutions are available.
- For example:
    1. Stata deals well with non-existence issues in the logit/probit;
    2. The user-written `ppml` command deals with non-existence issues in Poisson regression;
    3. A similar issue exists with other estimators (e.g., Tobit) and `ppml` can be used to address some of these.

- The existence of multiple optima has received less attention.
- This is perhaps because the issue does not arise in some leading cases (Poisson, logit, probit, Tobit).
- However, the existence of multiple (local) maxima is a problem for many frequently used estimators.
- In this presentation I'll focus on two important examples, but there may be many others.

2. The `heckman` command

- This is one of the most used (abused?) estimators in applied economics.
- Olsen (1982) shows that the log-likelihood function of the sample selection estimator has a unique maximum for fixed values of ρ.
- However, when ρ has to be estimated, the log-likelihood is not globally concave and multiple maxima may exist.
- Olsen (1982) suggested that estimation should start with a grid search over ρ; I believe Stata does not do that.

- Consider the following DGP:

$$y = 15 + x_1 - x_2 + (\kappa u_1 + u_2) \,/\, \left(1 + \kappa^2\right)^{0.5}$$
$$y \text{ is observed if } \left(1 + x_1 - x_2 + u_1\right) > 0$$
$$x_1 \sim U(0,1), \quad x_2 \sim B(1, 0.3), \quad u_i \sim N(0,1)$$

- The parameter κ controls the correlation between the errors of the two equations: $\rho = \kappa / \sqrt{(1 + \kappa^2)}$.
- I performed some simulations for different sample sizes and for different values of κ.
- Estimation was performed either using the default method or using as the starting values the ML estimates with the sign of ρ switched.

Table 1: Simulation results for the `heckman` command

n	250			1000		
κ	-2	0	2	-2	0	2
Both converged	959	999	951	1000	1000	1000
Alternative is better	151	37	125	58	9	58
Default is better	325	123	290	456	75	449

NB: results are considered different if the log-likelihoods differ by more than 0.1.

- Results based on 1000 replicas.
- None of the methods dominates the other.
- The existence of multiple maxima is an issue, especially with small samples.
- The differences between the results can be substantial.

3. The `zinb` command

- The zero-inflated negative binomial estimator is also very popular.
- Part of the reason for its popularity is due to misconceptions about overdispersion and to results of Vuong's test reported by Stata.
- Unfortunately:
  - `zinb` often converges to local maxima of the likelihood function.
  - Vuong's test as reported by Stata is not valid in this context.
- Next I use a small simulation to illustrate the existence of multiple maxima in the `zinb`.

- Consider the following DGP:

$$\begin{aligned} y^* &\sim \text{Poisson}(\mu) \\ \mu &= \exp(1 + x_1 - x_2)\,\eta \\ y &= y^* \times I\left(u > \frac{\exp(\kappa + x_1 - x_2)}{1 + \exp(\kappa + x_1 - x_2)}\right) \\ x_1 &\sim U(0,1), \quad x_2 \sim B(1, 0.3), \\ \eta &\sim \Gamma(1,1), \quad u \sim U(0,1) \end{aligned}$$

- So, y is generated by a ZINB and the probability of zero inflation increases with κ.
- I performed 1000 simulations for $\kappa \in \{-\infty, -2, -1\}$; these correspond to zero-inflation probabilities of about 0, 0.15, and 0.32.

- Estimation is performed using two different approaches:
  1. The default (start by estimating a model where μ is constant and then estimate the full model);
  2. Estimate the ZINB starting form the `nbreg` estimates.

Table 2: Simulation results for the `zinb` command

n	250			1000		
κ	$-\infty$	-2	-1	$-\infty$	-2	-1
Both converged	747	871	957	764	924	990
Alternative is better	103	179	50	133	271	9
Default is better	46	17	3	49	0	0

NB: results are considered different if the log-likelihoods differ by more than 0.1.

- Like before, no method dominates and the existence of multiple maxima is an issue.
- Again, in some cases the differences are substantial.

4. Vuong's test

- Vuong (1989) presents model selection tests that can be applied to nested, non-nested, and overlapping models.
  - For nested models, Vuong's test coincides with the classical LR test.
  - For overlapping models, Vuong's test is based on a statistic that under the null is distributed as a weighted sum of χ^2 random variables.
  - For strictly non-nested models, Vuong's test is directional and is based on a statistic that under the null has a normal distribution.
- For non-nested models, Vuong's test is very different from the tests for non-nested hypotheses inspired by Cox (1961).

- Stata implements Vuong's test for non-nested model to test for zero-inflation (ZINB vs NB and ZIP vs Poisson).
- However, the competing models are overlapping, not non-nested.
- This problem has been noted by Santos Silva, Tenreyro, and Windmeijer (2015) and Wilson (2015).
- The results of the test can be very misleading.
- For example, if the data is generated by a NB process, the test of the Poisson vs the ZIP will never favour the Poisson model and generally favours the ZIP.

5. Concluding remarks

- Multiple maxima in ML can be a serious problem.
- It would be great if Stata could do more to deal with this.
- At least `tnbr` is also affected by this problem.
- The current `vuong` option should be removed from `zip` and `zinb`.

- Cox, D.R. (1961). "Tests of Separate Families of Hypotheses." In Proc. 4th Berkeley Symp. Mathematical Statistics and Probability, vol. 1, 105–123. Berkeley: University of California Press.
- Olsen, R.J. (1982). "Distributional Tests for Selectivity Bias and a More Robust Likelihood Estimator," *International Economic Review* 23, 223–240.
- Santos Silva, J.M.C., Tenreyro, S., and Windmeijer, F. (2015). "Testing Competing Models for Non-Negative Data with Many Zeros," *Journal of Econometric Methods*, 4, 29–46.
- Vuong, Q.H. (1989). "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57, 307–333.
- Wilson, P. (2015). "The Misuse of the Vuong Test for Non-Nested Models to Test for Zero-Inflation," *Economics Letters*, 127, 51–53.