**Supplementary file 2 - Description of the conditional moment test**

By noting that the first process of the hurdle model involves only the random variable 𝑀, while the second process uses the variable $X\_{i}$, the conditional moment test is constructed on the following statements. If the separation hypothesis is respected, the following moment conditions hold for the first and second processes:

$$E\left(I\_{Y>0}\right)=P\left(M=1\right)P\left(X>0\right)= E\left(M\right)$$

$$E\left(Y∣Y>0\right)= E\left(M∣M>0\right) E\left(X∣X>0\right)= E\left(X\right) $$

The first equality cannot hold for the zero-inflated distribution since $P(X>0)\ne 1$. The second equality holds for both zero-inflated and hurdle models because $E\left(M∣M>0\right)$ is identically 1, since $M\~Bernoulli$. However, $\left(X∣X>0\right)$ = $\left(X\right) $ is only correct for a hurdle distribution and cannot hold for a zero-inflated distribution. In this situation, the parameters of the first and second processes of the hurdle distribution can be consistently estimated with the condition $G=(G\_{1},G\_{2})$, by the generalized method of moment (GMM):

$E\left\{I\_{\left(Y>0\right)}-E\left(M\right)\right\}=G\_{1}=\frac{1}{n}\sum\_{}^{}g\_{1}(α)=0$ (3)

$E\left\{Y-E\left(X\right)∣Y>0\right\}=G\_{2}=\frac{1}{n}\sum\_{}^{}g\_{2}(β)=0$ (4)

In our numerical example, $g\_{1}(α)$ is a (23 × 1) vector and $g\_{2}(β)$ a (23 × 1) vector. Therefore, having evaluated the parameters of the hurdle distribution by the last two equations using a generalized method of moments, the test can be done by checking the following equality :

$$E\left\{Y-E\left(M\right)E\left(X\right)\right\}=D=\frac{1}{n}\sum\_{}^{}d(α,β)=0$$

Using $θ=(α,β)$, under the null hypothesis that the model is correctly specified and that the separation hypothesis holds, explicit equations of this test are, following the notations of Prieger (2003):

$$T\_{CM}=\lim\_{n\to \infty }nD'\sum\_{0}^{-1}D$$

with:

$$\sum\_{0}^{}= V\_{dd^{'}}+D\_{0}J\_{0}^{-1}V\_{gd^{'}}+V\_{dg^{'}}J\_{0}^{-1}D\_{0}+D\_{0}J\_{0}^{-1}V\_{dd^{'}}J\_{0}^{-1}D\_{0}^{'}$$

$$J=-\lim\_{n\to \infty }\frac{1}{n}\frac{δG}{δθ}=-\lim\_{n\to \infty }\frac{1}{n}\left[\begin{matrix}\frac{δG\_{1}}{δα}&0\\0&\frac{δG\_{2}}{δβ}\end{matrix}\right]$$

$$D\_{0}=\lim\_{n\to \infty }\frac{1}{n}\frac{δD}{δθ}$$

$$V\_{dd^{'}}=D×D'$$

$$V\_{gd^{'}}=G×D'$$

$$V\_{dg^{'}}=D×G'$$

$$V\_{gg^{'}}=G×G'$$

The conditional moment test $T\_{CM}$ is asymptotically $χ^{2}(r)$ and the condition is rejected at significance level 𝛿 when $T\_{CM}>χ^{2}(r;δ)$, with 𝑟 the number of restrictions tested in the model. Rejection of the test indicates model misspecification, where the parameters of the distributions are not correctly estimated by equations (3) and (4).