***Appendix I***

***Do-File***

***Appendix II***

***Extra Assumptions and Diagnostic Tests Required For Panel Analysis***

*Test assumptions: Excessive Zeros (Zero-Inflated Consideration)*

From Figure 1 below, we note that the modal patent value is zero. In fact, zero makes up 1175 out of the total 5,136 patent values, representing about 22.88% of total values.

|  |
| --- |
| Chart  Description automatically generated |
| **Figure 1**. *Frequency of Patents Counts* |

The next diagnostic test involves examining whether there exist excess zeros in the dependent variable, warranting the consideration of zero-inflated models.The Vuong test indicates preference for the zero-inflated Poisson over the regular Poisson model (Vuong = 10.452, p = 0.000) and indicates preference for the zero-inflated negative binomial model over the regular negative binomial model (Vuong = 2.866, p = 0.002). We however note that there have been considerable reservations on the use of the Vuong test.

|  |
| --- |
|  |
| **Figure 2.** *Quantile-Quantile Plot of Poisson and Zero-Inflated Poisson Models.* |

|  |
| --- |
| Chart, line chart  Description automatically generated |
| **Figure 3.** *Quantile-Quantile Plot of Negative Binomial and Zero-Inflated Negative Binomial Models.* |

We explore the graphical examination procedure by Wilson (2015) where QQ-plot of the fitted probabilities of data under zero-inflated and non-zero inflated points are examined. From the QQ-plots shown below, the points do not lie on the x=y line, an indication that zero inflation may be indicated. The graphical examination is consistent with the result of the Vuong test.

*Data Diagnostic Tests: Count Data*

When analysing count dependent variables, two (2) important diagnostic tests that should be examined are overdispersion (Cameron and Trivedi, 1990) and excessive zeros (Vuong, 1989).

*Overdispersion*

The overdispersion test examine the null hypothesis of equidispersion where Var(Y|X) = E(Y|X) based on the following equation: . The statistical significance of determine whether Poisson or Negative binomial model is adequate: (equidispersion: Poisson model preferred) and (overdispersion: Negative Binomial model preferred).

Before examining the result of the formal test, we explore the relationship between the mean and the variance statistics of the dependent variable. From the table below, we note that the variance is substantially higher than the mean; an indication of overdispersion. Moreover, examining the mean and variance of every patent group also show substantial difference with the exception of groups with zero(0) counts.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 3**. Summary Statistics of Dependent Variable: Patents | | | |
|  | Mean | Standard Deviation | Variance |
| Overall | 26.30802 | 67.79056 | 4595.56 |
| Between |  | 66.70922 | 4450.1198 |
| Within |  | 12.3088 | 151.50657 |

The test of overdispersion test for both the model without the lagged values of research and development (t = 32.93, p = 0.000) and with the lagged values (t = 20.21, p = 0.000) are statistically significant; an indication of overdispersion. Hence, the test suggests that the negative binomial is preferred over the Poisson in model.

*Information Criteria*

Previous examination suggested the existence of individual firm effect within the model. Based on the results of the excess zeros examination, we investigate the performance of zero-inflated models with panel (random effects) forms of Poisson and Negative Binomial models. Table 4 below shows the performance of these models in terms of log-likelihood, Akaike Information Criterion and Bayesian Information Criterion. From the table, we note that the Panel Poisson and Panel Negative Binomial models had superior information criteria than the inflated-based models. Hence, the panel-based models would be used in estimations.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 4.** Akaike's information criterion and Bayesian information criterion | | | | |
|  | **Ll(model)** | **AIC** | **BIC** |
| Zero-Inflated Poisson | -41400.74 | 82815.470 | 82861.280 |
| Panel Poisson Random Effect Model | -14919.55 | 29849.100 | 29881.820 |
| Zero-Inflated Negative Binomial | -15532.18 | 31080.350 | 31132.710 |
| Panel Negative Binomial Model | -13515.28 | 27042.560 | 27081.830 |