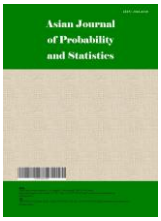

Asian Journal of Probability and Statistics

6(4): 24-33, 2020; Article no.AJPAS.54899
ISSN: 2582-0230

Estimating the Number of Patents in the World Using Count Panel Data Models

Ahmed H. Youssef[1], Mohamed R. Abonazel[1*] and Elsayed G. Ahmed[1]

[1]Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2020/v6i430167
Editor(s):
(1) Dr. S. M. Aqil Burney, University of Karachi, Pakistan.
Reviewers:
(1) Chipo Zidana, Botswana International University of Science and Technology, Botswana.
(2) Zimeras Stelios, University of the Aegean, Greece.
Complete Peer review History: http://www.sdiarticle4.com/review-history/54899

Original Research Article

Received: 09 January 2020
Accepted: 17 March 2020
Published: 19 March 2020

Abstract

In this paper, we review some estimators of count regression (Poisson and negative binomial) models in panel data modeling. These estimators based on the type of the panel data model (the model with fixed or random effects). Moreover, we study and compare the performance of these estimators based on a real dataset application. In our application, we study the effect of some economic variables on the number of patents for seventeen high-income countries in the world over the period from 2005 to 2016. The results indicate that the negative binomial model with fixed effects is the better and suitable for data, and the important (statistically significant) variables that effect on the number of patents in high-income countries are research and development (R&D) expenditures and gross domestic product (GDP) per capita.

Keywords: Conditional maximum likelihood estimation; fixed effects model; Hausman test; negative binomial regression; Poisson regression; random effects model.

1 Introduction

In econometrics literature, the panel data refer to the pooling of observations on a cross section of households, countries, firms, etc., over several time periods. Panel data are now widely used to estimate

**Corresponding author: E-mail: mabonazel@hotmail.com;*

dynamic econometric models. There are several advantages of panel data compared with either purely cross-sectional or purely time series data, such as [1] controlling for individual heterogeneity, panel data gives more informative data, and it is better able to study the dynamics of adjustment. And when the response variable of panel data model is non-negative integer value, in this case the model is called the count panel data (CPD) model. Actually, the use of CPD models has become popular in many economic applications for example, health economics, firm's productivity, patents, transportation and education. In panel data modeling, the most commonly estimated models in are probably fixed effects and random effects models.

In general, the fixed effects model allows each cross-sectional (i) unit to has a different intercept term though all slopes are the same. The fixed effects model, in the general form, can write as [2,3,4]:

$$y_{it} = \alpha_i + \mathrm{x}_{it}^{'}\beta + u_{it},\ i = 1, 2, \dots, N;\ t = 1, 2, \dots, T, \tag{1}$$

where y_{it} is the response variable for individual i at time t, x_{it} is the it^{th} observation on explanatory variables, α_i is the intercept, β is the vector of the regression coefficients, and u_{it} is the error term of the model. While the random effects model assumes that there is one constant term (α) for all across unites, and the differences of the intercept term can be captured in the error term, hence the error term become have new assumptions, so the rationale behind random effects model is that, unlike the fixed effects model, the variation across entities is assumed to be random, in addition to random effects assumes that the unit's error term is not correlated with the predictors, see [2,5]. The random effects model is given by:

$$y_{it} = \alpha + \mathrm{x}_{it}^{'}\beta + u_{it},\ i = 1, 2, \dots, N;\ t = 1, 2, \dots, T \tag{2}$$

where $u_{it} = \nu_i + \varepsilon_{it}$; this means that the error term of the model consists two components; where ν_i denotes the unobservable individual-specific effects, and ε_{it} denotes the disturbances which varies with units and time.

There are many economic studies, e.g. [6,7,8,9,10,11,12], on research and development (R&D) activities indicate an increasing interest in the relationship between firms' R&D investment and patent applications.

The purpose of this paper is to explore the main variables that effect on the number of patent applications in high-income countries by applying CPD models on seventeen high-income countries over the period from 2005 to 2016.

The rest of the article is organized as follows: section 2 presents Poisson and negative binomial models in both cases (fixed and random effects), and the proposed estimators of these models. In section 3 the empirical study on patents in the world is presented. Finally, section 4 offers the concluding remarks.

2 Count Panel Data Models

If the response variable in panel data models is not normally distributed, specifically; the response variable takes nonnegative integer values (count data). For example, the number of accidents in several areas, the number of days of absence for many persons over several years, the number of protests in each of several different countries over several years and number of doctor visits, the number of occurrences of a specific health event for each of many patients in multiple time periods. In econometrics literature, commonly used models that fit this data are Poisson and negative binomial models, see e.g. [13,14,15].

2.1 Poisson model

The Poisson model assumes that the response variable (y_{it}) has a Poisson distribution with a probability density function:

$$f(y_{it}; \lambda_{it}) = \frac{[exp(-\lambda_{it})]\ (\lambda_{it})^{y_{it}}}{y_{it!}},\ i = 1, 2, \dots, N;\ t = 1, 2, \dots, T, \tag{3}$$

where λ_{it} is the expected or the predicted mean of y_{it}. In this model, the mean and the variance of y_{it} must be equal, i.e. $E(y_{it}) = var\ (y_{it}) = \lambda_{it}$.

The Poisson model has one parameter to be estimated (λ_{it}), which is sometimes referred to as the location parameter and must be positive. It is convenient to specify λ_{it} as an exponential function of a linear index of the explanatory variables. The exponential form ensures that λ_{it} remains positive for all possible combinations of parameters and explanatory variables.

For the fixed effects Poisson (FEP) model, all characteristics that are not time-varying are captured by the individual heterogeneity term α_i. The intercept is merged into α_i, hence the regressors x_{it} do not include an intercept, see [16]. The conditional probability function of the FEP model as

$$f(y_{it}|x_{it}, \alpha_i, \beta) = \frac{[exp(-\alpha_i\lambda_{it})]\ (\alpha_i\lambda_{it})^{y_{it}}}{y_{it}!},\ i = 1, 2, \dots, N;\ t = 1, 2, \dots, T, \quad (4)$$

where $\lambda_{it} = exp(x_{it}'\beta)$. To estimate the parameters of this model, it can use the conditional maximum likelihood (CML) method that developed by Hausman et al [6]. Since y_{it} and $\sum_{t=1}^{T} y_{it}$ are follow the Poisson distribution, then the conditional joint density function (CJDF) for the i^{th} observation is

$$f(y_{i1}, \dots, y_{iT}|\sum_{t=1}^{T} y_{it}) = \frac{(\sum_{t=1}^{T} y_{it})!}{(\sum_{t=1}^{T} \lambda_{it})^{\sum_{t=1}^{T} y_{it}}} \prod_{t=1}^{T} \frac{\lambda_{it}{}^{y_{it}}}{y_{it}!},$$

taking the logarithm of CJDF and summing over all individuals, the conditional log-likelihood is

$$lnL = \sum_{i=1}^{N}\{ ln(\sum_{t=1}^{T} y_{it})! - \sum_{t=1}^{T} lny_{it}! + \sum_{t=1}^{T}[\, y_{it}x_{it}'\beta - y_{it} ln \sum_{t=1}^{T} exp(x_{it}'\beta)]\},$$

it can obtain the estimated parameters for the FEP model by solving

$$\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}' \left(y_{it} - \frac{\sum_{t=1}^{T} y_{it}}{\sum_{t=1}^{T} \lambda_{it}} \lambda_{it}\right) = 0.$$

In the random effects Poisson (REP) model, the individual-specific effect v_i must have a specified distribution to estimate the parameters of this model. Many papers, e.g. [6,17,3,18,19], assumed that the individual-specific effect in this model has a gamma distribution with parameters (γ, γ). They used the maximum likelihood estimation (MLE) method to estimate the parameters of this model. The maximum likelihood function for the i^{th} observation is

$$f(y_{it}|v_i, x_{it}) = \prod_{t=1}^{T} \left(\frac{\lambda_{it}{}^{y_{it}}}{y_{it}!}\right) \left[\frac{\gamma}{\gamma + \sum_{t=1}^{T} \lambda_{it}}\right]^{\gamma} \left[\frac{\Gamma(\sum_{t=1}^{T} y_{it} + \gamma)}{\Gamma(\gamma)}\right] [\gamma + \sum_{t=1}^{T} \lambda_{it}\ ]^{-\sum_{t=1}^{T} y_{it}}.$$

Note that the intercept is included in this model and merged into x_{it}. The log-maximum likelihood function is:

$$lnL = \sum_{i=1}^{N}\{ \sum_{t=1}^{T}(y_{it}x_{it}'\beta - lny_{it}!) + \gamma ln\gamma - \gamma\, ln[\gamma + \sum_{t=1}^{T} exp(x_{it}'\beta)] + l\, n[\Gamma(\sum_{t=1}^{T} y_{it} + \gamma)] - ln[\Gamma(\gamma)] - \sum_{t=1}^{T} y_{it}\, ln[\gamma + \sum_{t=1}^{T} exp(x_{it}'\beta)]\},$$

it can obtain the estimated parameters of this model by solving

$$\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}' \left[y_{it} - \lambda_{it} \left(\frac{\bar{y}_i + \gamma/T}{\bar{\lambda}_i + \gamma/T}\right)\right] = 0.$$

2.2 Negative binomial model

In general, the NB model is used as a good alternative to the Poisson model when the data has the over dispersion problem; this problem appears when $var\ (y_{it}) > E(y_{it})$. Since the NB model has a dispersion

parameter ϕ_i, it allows the variance to be greater than mean, as the dispersion parameter provides a wider shape of the distribution of counts than the Poisson model.

For the fixed effects negative binomial (FENB) model, Hausman et al [6] showed that the CJDF for the i^{th} observation is

$$f(y_{i1},\dots,y_{iT}|\sum_{t=1}^{T}y_{it}) = \frac{\Gamma(\sum_{t=1}^{T}\lambda_{it})\Gamma(\sum_{t=1}^{T}y_{it}+1)}{\Gamma(\sum_{t=1}^{T}\lambda_{it}+\sum_{t=1}^{T}y_{it})} \times \left[\prod_{t=1}^{T}\frac{\Gamma(\lambda_{it}+y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it}+1)}\right],$$

where $\sum_{t=1}^{T}y_{it} \sim NB\left(\theta_i\sum_{t=1}^{T}\lambda_{it}\ , (\theta_i\sum_{t=1}^{T}\lambda_{it})(1+\theta_i)\right); \theta_i = \alpha_i/\phi_i$, and $\Gamma(\cdot)$ is the gamma function. It can get the CML estimation of the FENB model, by maximizing the following log-conditional maximum likelihood function:

$$lnL = \sum_{i=1}^{N}\{ln\Gamma(\sum_{t=1}^{T}\lambda_{it}) + ln\Gamma(\sum_{t=1}^{T}y_{it}+1) - ln\Gamma(\sum_{t=1}^{T}\lambda_{it}+\sum_{t=1}^{T}y_{it}) + \sum_{i=1}^{T}[\ln\Gamma(\lambda_{it}+y_{it}) - \ln\Gamma(\lambda_{it}) - \ln\Gamma(y_{it}+1)]\}.$$

While in the random effects negative binomial (RENB) model, Hausman et al [6] assumed that y_{it} specified to be independent and identically distributed negative binomial, and $1/(1+\delta_i)$, where $\delta_i = \nu_i/\phi_i$, is distributed as beta with parameters (a, b). The mean and the variance of y_{it} are $\lambda_{it}\delta_i$ and $\lambda_{it}\delta_i(1+\delta_i)$, respectively. Then the CJDF for the i^{th} observation is

$$f(y_{it}|x_{it}) = \frac{\Gamma(\alpha+b)\,\Gamma(\alpha+\sum_{t=1}^{T}\lambda_{it})\,\Gamma(b+\sum_{t=1}^{T}y_{it})}{\Gamma(\alpha)\,\Gamma(b)\,\Gamma(\alpha+b+\sum_{t=1}^{T}\lambda_{it}+\sum_{t=1}^{T}y_{it})} \times \prod_{t=1}^{T}\left[\frac{\Gamma(\lambda_{it}+y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it}+1)}\right].$$

The ML estimation of the RENB model can be obtained by maximizing the following log-maximum likelihood function:

$$lnL = \sum_{i=1}^{N}\{ln\,\Gamma(\alpha+b) + ln\,\Gamma[\alpha+\sum_{t=1}^{T}\lambda_{it}] + ln\,\Gamma(b+\sum_{t=1}^{T}y_{it}) - ln\,\Gamma(\alpha) - ln\,\Gamma(b) - ln\,\Gamma[\alpha+b+\sum_{t=1}^{T}\lambda_{it}+\sum_{t=1}^{T}y_{it}] + \sum_{i=1}^{T}[ln\,\Gamma(\lambda_{it}+y_{it}) - ln\,\Gamma(\lambda_{it}) - ln\,\Gamma(y_{it}+1)]\}.$$

3 Empirical Study

In this application, the sample was chosen based on the available data on the number of patents in high-income countries in the World Bank website; the sample contains 17 high-income countries over the period from 2005 to 2016. *R-software* (*pglm* package) was used to perform CPD models in this application.

In our study, the response variable is the number of patent applications. the patent application, according to World Intellectual Property Organization (WIPO), is a product or process that provides a new way of doing something or offers a new technical solution to a problem, a patent provides protection for the invention to the owner of the patent for a limited period, generally 20 years (WIPO Patent Report: Statistics on Worldwide Patent Activity). While the explanatory variables, in our study, include the R&D expenditures, the number of researchers in R&D, the gross domestic product (GDP) per capita, the information and communication technology (ICT) goods imports, and unemployment rate. A description of the variables selected in our study is presented in Table 1.

Table 1. Data descriptions

Variable	Description	Unit
PATE	Number of patents **(*response variable*)**	Count
NURD	Number of researchers engaged in R&D	Count (per 1000 people)
RDEX	The logarithm of the R&D expenditures	U.S. Dollar
GDPC	The logarithm of the GDP per capita	U.S. Dollar
IMPO	The logarithm of the ICT goods imports	U.S. Dollar
UNEM	Unemployment rate	Percentage of total labor force

Fig. 1 shows boxplots of the number of patent applications for countries under study. This figure shows that the variation of the number of patents between the countries, but within each country the distribution of the number of patents almost symmetric and it has not outlier values.[1]

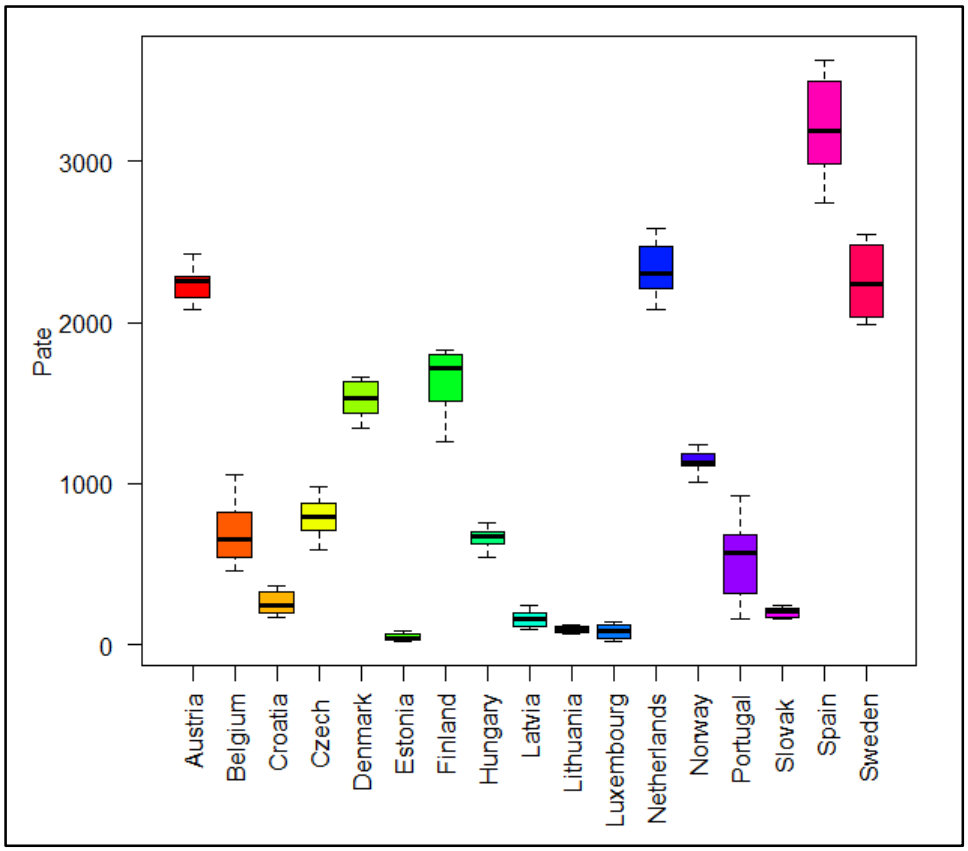

Fig. 1. Boxplots of the number of patents for each country

Table 2 resents some descriptive statistics of the six variables. It shows that the coefficient of variation (CV) of all variables less than one, then the data not have large variation.

Table 2. Descriptive statistics

Variable	Mean	Std. Dev.	CV	Min.	Max.
PATE	1049.922	963.934	0.918	15	3632
NURD	3.824	1.717	0.449	1.308	7.846
RDEX	21.674	1.581	0.073	18.310	23.787
GDPC	10.304	0.701	0.068	8.930	11.685
IMPO	22.432	1.221	0.054	19.950	24.829
UNEM	0.087	0.045	0.517	0.025	0.261

Table 3 presents the pairwise correlation coefficients between all variables. It can note that the correlation between PATE and RDEX is the higher correlation, while the smallest correlation is between PATE and

[1]*If the data contains outlier values, then the classical (non-robust) estimator is not efficient, a robust estimator must be used to estimate the regression parameters. Many robust estimators are discussed by many papers in several regression models, see e.g. [20,21,22,23,24].*

UNEM. Also, the results of Table 3 indicate that the data not have multicollinearity problem[2] because all the values of the variance inflation factor (VIF) are less than 10.

Table 3. Correlation matrix and VIF

	PATE	NURD	RDEX	GDPC	IMPO	UNEM
PATE	1					
NURD	0.3954	1				
RDEX	0.8459	0.6075	1			
GDPC	0.4507	0.7638	0.6607	1		
IMPO	0.7091	0.2258	0.7929	0.3045	1	
UNEM	-0.0179	-0.4187	-0.2342	-0.4994	-0.1842	1
VIF	-------	**3.1884**	**7.9014**	**3.8203**	**4.5574**	**1.9502**

For selecting the appropriate CPD model for this data, we will follow the methodology presented in Fig. 2. This figure summarizes the estimation steps and how to select the appropriate (efficient) model for the data. According to our methodology, the four CPD models will be estimating, and conducting the Hausman [25] test[3] to compare the fixed and random effects models. In the final step, the selection criteria (goodness-of-fit measures) will be used to select the appropriate CPD model.

Table 4 presents the results of FEP and REP models. We estimated the parameters in fixed effects using CML method, while the MLE method was used to estimate the random effects model. The two (FEP and REP) models are statistically significant because the P-value of Wald test is less than 0.05. Based on the results of Hausman test, the P-value of chi-squared is less than 0.05, then we can reject the null hypothesis, this means that FEP model is more appropriate.

Table 4. Estimates of Poisson models

Variable	Fixed effects model			Random effects model		
	Estimate	**T-value**	**P-value**	**Estimate**	**T-value**	**P-value**
NUMRD	0.0170	2.69	0.007	-0.0428	-7.85	<0.001
RDEX	0.6671	21.13	<0.001	0.5299	18.80	<0.001
GDPC	-0.2068	-3.14	0.002	-0.4818	-12.05	<0.001
IMPO	-0.1547	-8.18	<0.001	0.0157	1.02	0.310
UNEMP	0.5125	4.27	<0.001	-0.3769	-4.20	<0.001
Intercept	--------	--------	--------	-0.1123	-0.25	0.800
Wald Test	$\chi^2 = 684.93$, df = 5, P-value (χ^2) < 0.001			$\chi^2 = 2285.09$, df = 5, P-value (χ^2) < 0.001		
Hausman Test	$\chi^2 = 450.06$, df = 5, P-value (χ^2) < 0.001					

Table 5 presents the results of CML estimates of FENB model and MLE estimates of RENB model. The two (FENB and RENB) models are statistically significant because the P-value of Wald test is less than 0.05. Since the P-value of Hausman test is less than 0.05, then the FENB model is more appropriate.

Based on the results in Tables 4 and 5, we concluded that FEP and FENB models are better than REP and RENB models. Then we should use the Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) to determine the appropriate model (FEP or FENB).[4] Table 6 shows that the FENB model has minimum AIC and BIC values, then it is the best model to fit the data.

[2]This problem arises when the explanatory variables are highly inter-correlated. Then it becomes difficult to disentangle the separate effects of each of the explanatory variables on the response variable. As a result, the estimated regression parameters may be statistically insignificant and/or have, unexpectedly, different signs. Thus, conducting a meaningful statistical inference would be difficult for the researcher. See e.g. [26,27,28,29,30,31] for handling and solving this problem in several regression models.

[3]In panel data modeling, Hausman test is used to determine the appropriate model for the data (fixed effects model or random effects model). The null hypothesis of this test is the random effects model is appropriate.

[4]AIC and BIC are introduced by Akaike [32] and Schwarz [33], respectively, to analyze the performance of the statistical models and identify the best model among the various models, where the best model is the model has minimum value of AIC or BIC.

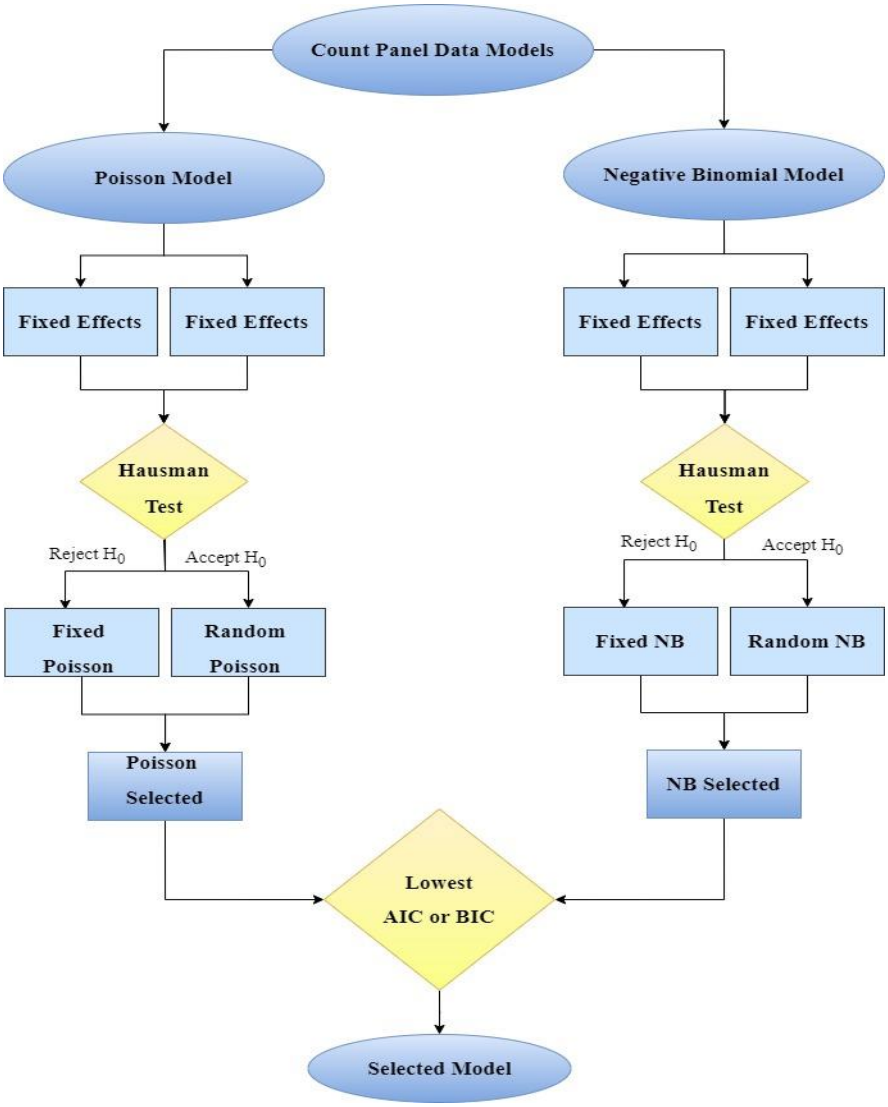

Fig. 2. The methodology for selecting the appropriate count model

Note: H_0: The random effects model is appropriate; AIC: Akaike's information criterion; BIC: Bayesian information criterion

Table 5. Estimates of negative binomial models

Variable	Fixed effects model			Random effects model		
	Estimate	**T-value**	**P-value**	**Estimate**	**T-value**	**P-value**
NUMRD	0.0119	0.39	0.696	0.0197	2.31	0.021
RDEX	0.9277	7.83	<0.001	0.8301	18.77	<0.001
GDPC	-0.6581	-4.13	<0.001	-0.4197	-6.57	<0.001
IMPO	-0.1418	-1.50	0.113	0.0365	1.68	0.093
UNEMP	0.2842	0.56	0.579	-0.3150	-2.42	0.016
Intercept	-6.8310	-3.09	0.002	-8.1280	-17.04	<0.001
Wald Test	$\chi^2 = 761.31$, df = 5, P-value (χ^2) < 0.001			$\chi^2 = 8017.53$, df = 5, P-value (χ^2) < 0.001		
Hausman Test	$\chi^2 = 15.81$, df = 5, P-value (χ^2) = 0.0074					

In the FENB model, we find that RDEX and GDPC variables are significant because the P-values of the two variables are less than 0.05, while the other variables are not significant.

Table 6. Goodness-of-fit measures of fixed effects models

Measure	Fixed effects poisson	Fixed effects negative binomial
Log likelihood	-2347.701	-1125.471
AIC	4727.403	2284.942
BIC	4732.356	2290.206

4 Conclusion

In this paper, we examined the effect of some economic variables on the number of patent applications in 17 high-income countries over the period from 2005 to 2016 by applying four CPD models. The Hausman test has been conducted to compare fixed and random effects models; the results of the Hausman test indicate that fixed effects models are better than random effects models. Using selection criteria (AIC and BIC), we find that the FENB model is the appropriate for this data, because it has minimum AIC and BIC values.

Based on the results of the FENB model, we concluded that the R&D expenditures have positive significant effect on the number of patents. However, the GDP per capita has negative significant effect on the number of patents in high-income countries. While the number of researchers and the technology goods imports are not have significant effect on the number of patents.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Gujarati DN, Porter DC. Basic econometrics. McGraw-Hill Education; 2009.

[2] Hsiao C. Analysis of panel data. 2nd edition, Cambridge University Press; 2003.

[3] Cameron AC, Trivedi PK. Regression analysis of count data. Cambridge University PRESS. 2013;53.

[4] Biørn E. Econometrics of panel data: Methods and applications. Oxford University Press; 2016.

[5] Wooldridge JM. Econometric analysis of cross section and panel data; 2010.

[6] Hausman JA, Hall BH, Griliches Z. Econometric models for count data with an application to the patents-R&D relationship; 1984.

[7] Cincera M. Patents, R&D and technological spillovers at the firm level: some evidence from econometric count models for panel data. Journal of Applied Econometrics. 1997;12(3):265-280.

[8] Blundell R, Griffith R, Reenen JV. Dynamic count data models of technological innovation. The Economic Journal. 1995;105(429):333-344.

[9] Crépon B, Duguet E. Estimating the innovation function from patent numbers: GMM on count panel data. Journal of Applied Econometrics. 1997;12(3):243-263.

[10] Wang P, Cockburn LM, Puterman ML. Analysis of patent data—A mixed-Poisson-regression-model approach. Journal of Business & Economic Statistics. 1998;16(1):27-41.

[11] Blundell R, Griffith R, Windmeijer F. Individual effects and dynamics in count data models. Journal of Econometrics. 2002;108(1):113-131.

[12] Gurmu S, Pérez-Sebastián F. Patents, R&D and lag effects: Evidence from flexible methods for count panel data on manufacturing firms. Empirical Economics. 2008;35(3):507-526.

[13] Kumara SSP, Chin HC. Study of fatal traffic accidents in Asia Pacific countries. Transportation Research Record. 2004;1897(1):43-47.

[14] Noland R, Quddus M. Analysis of pedestrian and bicycle casualties with regional panel data. Transportation Research Record: Journal of the Transportation Research Board. 2004;1897(1):28-33.

[15] Angers JF, Desjardins D, Dionne G, Guertin F. Modelling and estimating individual and firm effects with count panel data. ASTIN Bulletin: The Journal of the IAA. 2018;48(3):1049-1078.

[16] Boucher JP, Denuit M. Fixed versus random effects in Poisson regression models for claim counts: A case study with motor insurance. ASTIN Bulletin: The Journal of the IAA. 2006;36(1):285-301.

[17] Winkelmann R. Econometric analysis of count data. Springer Science & Business Media; 2008.

[18] Baltagi BH. (Ed.). The Oxford handbook of panel data. Oxford Handbooks; 2015.

[19] Croissant Y, Millo G. Panel data econometrics with R. John Wiley and Sons, Incorporated; 2019.

[20] Alma Ö. Comparison of robust regression methods in linear regression. International Journal of Contemporary Mathematical Sciences. 2011;6(9):409-21.

[21] Abonazel M, Gad A. Robust partial residuals estimation in semiparametric partially linear model. Communications in Statistics-Simulation and Computation, Accepted paper 2018; 2018.

[22] Abonazel MR. Advanced statistical techniques using R: Outliers and missing data, annual conference on statistics, computer sciences and operations research. Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt. 2019;54.

[23] Abonazel M, Rabie A. The impact for using robust estimations in regression models: An application on the Egyptian economy. Journal of Advanced Research in Applied Mathematics and Statistics. 2019;4(2):8-16.

[24] Abonazel M, Saber O. A comparative study of robust estimators for Poisson regression model with outliers. Working paper, Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt; 2019.

[25] Hausman JA. Specification tests in econometrics. Econometrica. 1978;46(6):1251–1271.

[26] Kibria BG. Performance of some new ridge regression estimators. Communications in Statistics-Simulation and Computation. 2003;32(2):419-435.

[27] Abonazel MR. A practical guide for creating Monte Carlo simulation studies using R. International Journal of Mathematics and Computational Science. 2018;4(1):18-33.

[28] Abonazel MR. New ridge estimators of SUR model when the errors are serially correlated. International Journal of Mathematical Archive. 2019;10(7):53-62.

[29] Rady EA, Abonazel MR, Taha IM. Ridge estimators for the negative binomial regression model with application. The 53rd Annual Conference on Statistics, Computer Science, and Operation Research 3-5 Dec, 2018. ISSR, Cairo University, Egypt; 2018.

[30] Rady EA, Abonazel MR, Taha IM. New shrinkage parameters for Liu-type zero inflated negative binomial estimator. The 54th Annual Conference on Statistics, Computer Science, and Operation Research. Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt; 2019.

[31] Rady El-Housainy A, Abonazel Mohamed R, Taha IM. A new biased estimator for zero-inflated count regression models. Journal of Modern Applied Statistical Methods, Accepted paper 2019; 2019.

[32] Akaike H. A new look at the statistical model identification. In Selected Papers of Hirotugu Akaike. Springer, New York. 1974;215-222.

[33] Schwarz G. Estimating the dimension of a model. Annals of Statistics. 1978;6(2):421–464.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/54899