University of Brescia

Department of Economics and Management

Scientific Conference on

Statistics

for

Health and Well-being

University of Brescia

Department of Economics and Management

25 – 27 September 2019

ASA CONFERENCE 2019
Statistics for Health and Well-being

BOOK OF SHORT PAPERS

Maurizio Carpita and Luigi Fabbris

Editors

University of Brescia
Department of Economics and Management

Scientific Conference on

Statistics for Health and Well-being

University of Brescia
Department of Economics and Management
25 – 27 September 2019

ASA CONFERENCE 2019
Statistics for Health and Well-being

BOOK OF SHORT PAPERS

Maurizio Carpita and Luigi Fabbris
Editors

Analysis of the financial performance in Italian football championship clubs *via* longitudinal count data and diagnostic test

Anna Crisci[a], Luigi D'Ambra[a]
[a] Department Economic, Management, Institutions, University of Naples Federico II, Naples, Italy.

1. Introduction

Football is undoubtedly the most powerful and most popular sport in Italy, linking communities and stirring emotions. Professional business operators consider football an important industry with enormous potential in terms of growth and also for the indirect benefits gained by investors and management due to the popularity of football teams. In the football world, major consulting companies provide statistical data relating exclusively to athletic performance and sports results. The recipients of such data can be placed in two main categories. The first concerns professional football players, sports clubs, coaches, sports directors, etc. Such information is sold, in some cases, for payment. The second category is represented by media outlets, which release statistical reports to fans and sports people. The main goal of any Football Championship club is to achieve sport results. Nevertheless, football has also become one of the most profitable industries, with a significant economic impact in infrastructure development, sponsorships, TV rights and transfers of players. Very informative is considered the connection between the points in the championship and the resource allocation strategies. The aim of this paper is to give an interpretation of the link between the points in the championship and the resource allocation strategies using the longitudinal count data. In addition to the introduction, this paper consists of two further sections. In Section 2, the panel data approach is described while, in Section 3 a case study is shown.

2. The panel data

We often have data where variables have been measured for the same subjects (or countries, or companies, or whatever) at multiple points in time. These are typically referred to as Panel Data or as Cross-Sectional Time Series Data. With panel data you can include variables at different levels of analysis (i.e. students, schools, districts, states) suitable for multilevel or hierarchical modeling. Why do we use panel data? (Hsiao, 1985).

Benefits:

- They allow to identify the effects that are not identified in the cross-section data (Ben-Porath,1973).
- The panel allows to study the dynamics: while the cross-section allows you to estimate what proportion of the population is unemployed in a unit of time, the panel data show how this share varies over time;
- The panel data contain more information, more variability and therefore less collinearity among the variables and produce estimates more efficient, more precise parameters.
- They allow to control the effect of individual heterogeneity: i.e variables constant over time (individual heterogeneity) not observed (for which no data are available) (Baltagi and Levin, 1992).

Limits:

- Difficulty in the sample design and data collection.
- Distortion of the measurement errors.

CLEUP Coop. Libraria Editrice
ISBN: 978-88-5495-135-8

- Problem of selection, no answers nor dissensions
- Limited dimension of time series.

2.1Fixed and random effects

The fixed effects (FE) explore the relationship between predictor and outcome variables within an entity (persons, teams, company, etc.). Each entity has its own individual characteristics that may or may not influence the predictor variables. Each entity is different, therefore the entity's error term and the constant (which captures individual characteristics) should not be correlated with the others (Stock and Watson, 2012).

The fixed effect model is:

$$\boldsymbol{y}_{it} = \boldsymbol{\beta}'\boldsymbol{x}_{it} + \alpha_i + \boldsymbol{\varepsilon}_{it} \qquad (1)$$

where

α_i (i=1…n) is the unknown intercept for each entity (n entity-specific intercepts).
$\boldsymbol{y}_{it}$ is the vector of dependent variables where $i =$ entity and $t =$ time.
$\boldsymbol{x}_{it}$ represent the vector of covariates.
$\boldsymbol{\varepsilon}_{it}$ is the vector of error terms.

In the random effects the variation across entities is assumed to be random and uncorrelated with the predictor or independent variables included in the model. The crucial distinction between fixed and random effects is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are "stochastic or not". The random effect model is:

$$\boldsymbol{y}_{it} = \boldsymbol{\beta}'\boldsymbol{x}_{it} + \alpha_i + \boldsymbol{\varepsilon}_{it} \qquad (2)$$

where $\boldsymbol{v}_{it} = \alpha_i + \boldsymbol{\varepsilon}_{it}$ is the error of the random effect model.

The generally accepted way of choosing between fixed and random effects is running a Hausman H-test (Hausman, 1978). Statistically, fixed effects are always a reasonable thing to do with panel data (they always give consistent results) but they may not be the most efficient model to run. Random effects will give you better p.values as they are a more efficient estimator, so you should run random effects if it is statistically justifiable to do so. Under the null hypothesis, the random effects is correctly specified, so both the fixed and random effects model are consistent, while under the alternative hypothesis, the random effects are correlated with the regressors, so the random effects model loses its consistency.

3. Case study

The data used for our case study was obtained from the financial statements filed by the Serie A football teams. The period of study concerned the championship from season 2010/2011 up to 2014/2015.

The focus of the analysis is to verify the impact that some financial indicators have on the points achieved by football teams. We consider the following independent variables: Depreciation Expense of multi-annual player contracts (DEM), Net equity (NE) and Revenue net of player capital gain (RNC) . In addition, we have considered, on the bases a bivariate descriptive analysis, also the square effect of DEM (DEM^2), given the non- linear relationship between Point and DEM. Finally, the interaction between DEM and NE (DEM*NE) also was considered.

In order to explore the panel data, figure 1, shows Point versus Year from 2010 to 2015; a line connects the five observations within each team. These lines represent a change over time.

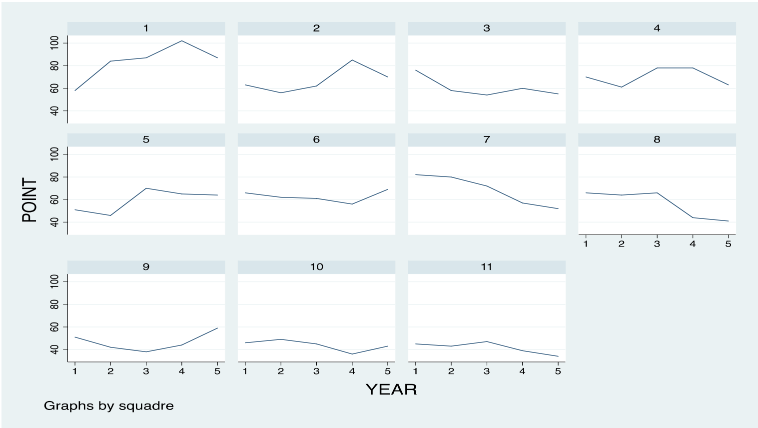

Figure 1: plot Point versus Year from 2010 to 2015

The fixed effects (FE) Poisson model, in table 2, shows a significant overall model (p.value = 0.0397), with only one statistically significant variable: the RNC.

Table 1: Fixed effects Poisson regression

Point	Coef.	Std.Err.	z	p.value
DEM	-0.426	2.864	-0.15	0.882
NE	-0.737	0.801	-0.92	0.358
RNC	0.288	0.104	2.75	0.006
DEM^2	-0.011	0.099	-0.11	0.914
DEM*NE	0.044	0.046	0.96	0.336

The output of the random effects (RE) Poisson model is shown in table 2:

Table 2: Random effects Poisson regression

Point	Coef.	Std.Err.	z	p.value
DEM	3.211	1.546	2.08	0.038
NE	-0.776	0.383	-2.02	0.043
RNC	0.316	0.060	5.25	0.000
DEM^2	-0.120	0.050	-2.3	0.017
DEM*NE	0.048	0.022	2.14	0.033
Cons.	-22.261	12.861	-1.73	0.083
/ln alpha	-7.3223	2.5236		
alpha	0.0006	0.0016		

In the random effects model we have all variables statistically significant. Finally, the Hausman H-test reveals that the random effects estimator is more appropriate (p.value= 0.2851, well above the critical value of 0.05).

Some final consideration should be made. The validity of the RE Poisson depends on very strong distributional assumptions. So, we would just stick to the FE regression. In particular, the choice of dealing with individual effects as fixed or random enough delicate. The fixed effects should be used to estimate the specific effects of the sample (i.e, an exhaustive sample countries, a sample of companies in a particular industry in which the selected sample is representative of the characteristics of the industry). By contrast, the random effects should be used for random samples and to make inference on the population. Then, in our case the choice could be cast on the fixed effects model, as our entity can not really be thought of as random draws from a population. In fact, the inferences that we have drawn are conditioned to the individuals included in the sample as opposed to a random model where the individual characteristics become a component of the population and the inferences are then related to the same population.

References

Baltagi, B. H. and Levin D. (1992). Cigarette Taxation: Raising Revenue and Reducing Consumption. *Structural change and Economics Dynamic,* **3**(2), pp. 321-335.

Ben-Porath, Y. (1973). Labor-Force Participation Rates and the Supply of Labor. *Journal of Political Economy*, **81**(3), pp. 697-704.

Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica*, **46**(6), pp. 1251-1271.

Hsiao, C. (1985). Benefits and Limitations of Panel Data. *Econometric Reviews*, **4**(1) pp.121-174.

Stock, H.J and Watson, M.W. (2012). *Introduction to Econometrics*, 3rd ed., Pearson Addison Wesley. Chapter 10.

ISBN: 978-88-5495-135-8