A Methodological Review on Time Series Panel Data- by Mahima Oli

Email: mahimaaashnaoli@gmail.com

Abstract

It is a general methodological review on Time Series Panel Data. The paper argues that timeseries panel data can be analyzed from more general to the more specific, which is known as a hypothetico-deductive model that includes seven main stages: problem, hypothesis, research design, measurement, data collection data analysis, and generalization. For the interpretation of data, the lean six sigma method can be used to integrate the six sigma concept (focusing on reducing defect products) and the lean concept (focusing on eliminating waste) as the methodology of cycle (DMAIC: define, measure, analyse, improve and control). Multiple regression analysis may suffer from multicollinearity or heteroscedastic, or autocorrelation problem. The variance inflation factors (VIF) is used to check multicollinearity that enabled us to eliminate the problem and select the "best" predictor variable to enter when other independent variables are present. Cronbach's alpha as an index's reliability is a metric used to assess the interval consistency of a set of items from zero (no internal consistency) to unity (perfect internal consistency). The review is applicable for the researchers who want to use time series panel data for the quantitative analysis. It is a general methodological review based on secondary information. To generalize the findings of Time Series Panel Data in the process of generalization, Cronbach's alpha, Hypothetico-deductive model, Lean six sigma method, and Multiple regression analysis are the basic significant statistical attributes.

Keywords: Cronbach's alpha, Hypothetico-deductive model, Lean six sigma method, Multiple regression analysis

Introduction

A methodological review is a type of systematic secondary research (i.e., research synthesis) which focuses on summarising the state-of-the-art methodological practices of research in a substantive field or topic (Chong & Reinders, 2021). Methodological reviews can be performed to examine any methodological issues relating to the design, conduct and review of research studies and also evidence syntheses (Munn et al, 2018). It is a synthesis of Methodological Review on Time Series Panel Data.

This paper introduces the methodology of time series data for interpretation that comes from different sources (e.g., World Bank Data Open Portal)

(https://data.worldbank.org/), prosperity index

(https://www.prosperity.com/rankings), Global Innovation Index,

(https://www.globalinnovationindex.org/home), Worldwide Governance Index,

(https://info.worldbank.org/governance/wgi/), Human Development Index,

(https://hdr.undp.org/), Economic Freedom Index

(https://www.fraserinstitute.org/) Transformation Index,

(https://bti-project.org/en/?&cb=00000), World Population Prospects,

(https://www.un.org/development/desa/pd/) International Migration Stock ,

(https://www.un.org/development/desa/pd/content/international-migration-1)), and among others.

Research Approaches

The study has qualitative approach to assist quantitative studies. The time series panel data study is generally influenced by positivism philosophy, which is reviewed critically by Denzin and Lincoln (2005:24) as philosophical worldviews (see Table 1.1).

Table.1.1: Various Paradigms and their Criteria, Form of Theory and Type of Narration

Paradigm/Theory	Criteria	Form of Theory	Type of Narration
Positivist/ Postpositivist	Internal, external validity	Logical-deductive, grounded	Scientific report
Constructivist	Trustworthiness, credibility, Transferability, conformability	Substantive-formal	Interpretive case studies, ethnographic fiction

Paradigm/Theory	Criteria	Form of Theory	Type of Narration
Feminist	Afrocentric, lived experience, dialogue, caring, accountability, race, class, gender, reflexivity, praxis, emotion, concrete grounding	Critical, standpoint	Essays, stories, experimental writing
Ethnic	Afrocentric, lived experience, dialogue, caring, accountability, race, class, gender	Standpoint, critical, historical	Essays, fables, dramas
Marxist	Emancipatory theory, falsifiability dialogue, race class gender	Critical, historical, economic	Historical, economic, sociocultural analysis
Cultural Studies	Cultural practices, praxis, social texts, subjectivities	Social criticism	Cultural theory-as criticism
Queer Theory	Reflexivity, deconstruction	Social criticism, historical analysis	Theory as criticism, autobiography

(Source: Study, 2023)

In addition, Creswell (2014) views as "This philosophical worldview focuses on the needs of groups and individuals in our society that may be marginalized or disenfranchised" (p. 39). The goal of transformative research is to lead to political or social change to the benefit of marginalized groups.

The pragmatic worldview is often associated with mixed research because it "is not committed to any one system of philosophy and reality" (Creswell, 2014, p.39). Its concern is with the research problem itself, and then employing whatever is needed to understand the problem. However, pragmatic approach is beyond the scope of the study.

Objective

To generalize the findings of Time Series Panel Data in the process of generalization

Materials and Methods

It is a general methodological review on Time Series Panel Data. Panel data or longitudinal data typically refer to data containing time series observations of a number of individuals. Therefore, observations in panel data involve at least two dimensions; a cross-sectional dimension, indicated by subscript i, and a time series dimension, indicated by subscript t. However, panel data could have a more complicated clustering or hierarchical structure (Hsiao, 2005). Thus the review is applicable for the researchers who want to use time series panel data for the quantitative analysis. It is a general methodological review based on secondary information.

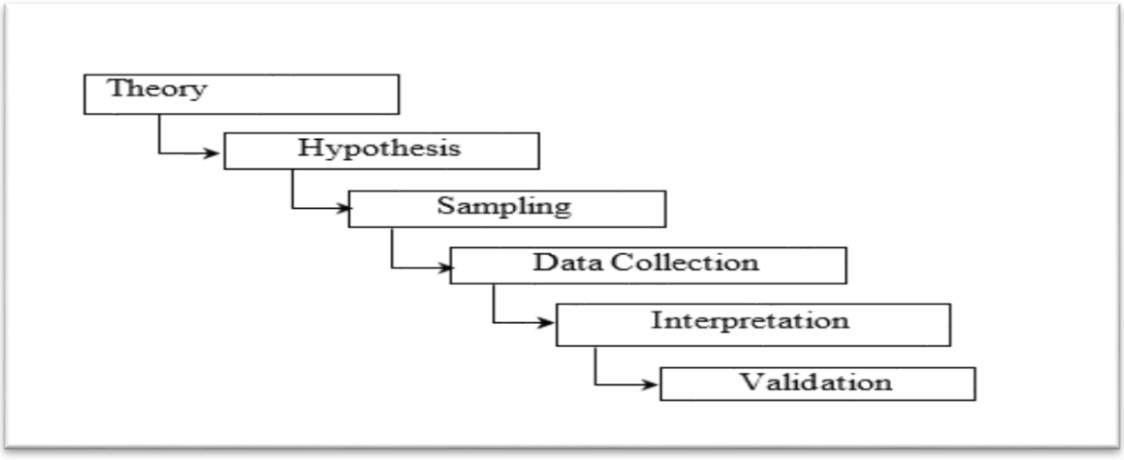
Results and Discussion

The most characteristic feature of the research process is its cycle nature. It usually starts with a problem and ends with tentative empirical generalizations (Frankfort-Nachmais &Nachamai, 2008). Thus it is based on Hypothetico-deductive model, Lean six sigma method, Multiple regression analysis and Cronbach's alpha.

Hypothetico-Deductive Model

For these time series panel data, analysis comes from the more general to the more specific. It might start thinking of a theory about the topic of interest. It then narrows that down into more specific hypothesis that it can test. It narrows down even further when it employs the techniques of sampling and collect information to test the hypotheses or a validation (or not) of the original theories as shown in Figure 1.1.

Figure 1.1`: Model of Hypothetico-Deductive Method



(Study, 2023)

The most characteristic feature of the research process is its cycle nature. It usually starts with a problem and ends with tentative empirical generalizations (Frankfort-Nachmais &Nachamai, 2008). As illustrated in Figure 1.1, the research process consists of seven main stages: problem, hypothesis, research design, measurement, data collection data analysis, and generalization. Each stage affects theory and is affected by it as well.

Lean Six-Sigma Methodology

A researcher can adopt the lean six sigma method which can be used to integrate the six sigma concept (focusing on reducing defect products) and the lean concept (focusing on eliminating waste) for the interpretation of data. According to the work of Furterer and Elshennawy (2005), it concentrates on improving quality, reducing variety, and eliminating waste in research organization/industry. Using lean six sigma as a business strategy can significantly increase growth and profitability while reducing expenses (Jeyaraman & Kee, 2010). The majority of lean six sigma research (Kumar et al., 2006; Su et al., 2006; Souraj et al., 2010) described it as a continuous improvement methodology that focuses on product and process performance.

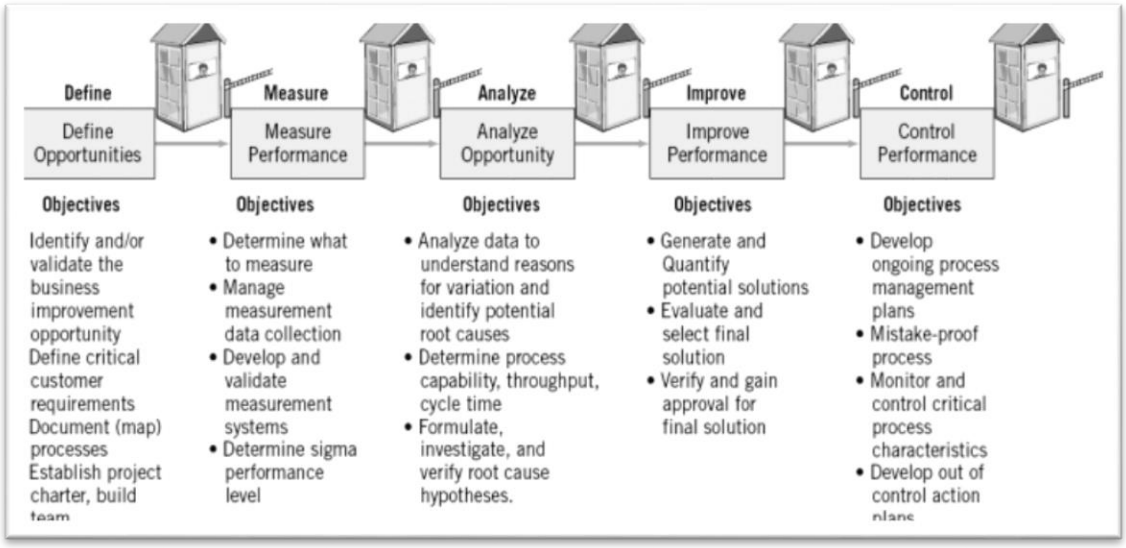
Most of the lean six sigma literature shown that combining the two techniques might properly support one another and result in a powerful approach to increase the quality of the product and process while also shortening the time. When used separately, the approaches of lean and six sigmas do not offer the possibility of progress, but when combined, they complement each other effectively (Pepper & Spedding, 2010).

The integration of the lean six sigma concept creates an effective methodology to improve products, process and increase customer satisfaction (Kumar et al., 2006; Su et al., 2006, Souraj et al., 2010). The integration of lean six sigma can create a dramatic improvement rather than when lean or six sigma is used separately (Cudney et al., 2006). This integration can reduce cost, time and improve quality and organisational performance (Lee & Wei, 2010; Kumar et al., 2006). Lean six sigma leads to synergies for quality improvement in industry and promote the combination of human and process aspects as a part of process improvement that it reinforces each other simultaneously (Su et al., 2006; Snee, 2010) Furthermore, the integration of lean and six sigma principles may lead a significant result by improving the productivity and quality of providing financial services (Furterer & Elshennawy, 2005).

Generally, two types of common framework are supported by the six sigma concept that is to be suitable for the lean six sigma framework too. Firstly, the DMAIC cycle (define, measure, analyse, improve and control) can be used for existing products or processes. Secondly, the DFLS cycle (design for lean sigma) can be applied for the design of new products or processes (Aggogeri & Mazzola, 2008).

The six sigma methodology can be deployed as the DMAIC for process improvement and the DMADV (define, measure, analyse, design and verify) for new product and service development (Montgomery and Woodall, 2008; Pusporini et al., 2013). The usage of DMAIC and DFLS (Design for Lean Sigma) or the DMADV cycle provides a generic framework that can be used as guidance to apply the lean six sigma concept in industry. The DMAIC process can be seen in Figure. 1.2.

Figure 1.2: The DMAIC process



(Source: Montgomery and Woodall, 2008)

Throughout the literature the DMAIC cycle is found a useful framework for lean six sigma implementation (Souraj et al., 2010; Snee, 2010; Pusporini et al., 2013). Every step in DMAIC cycle has a significant correlation with improvement performance and each step needs to be identified clearly.

The use of the DMAIC cycle based on the six sigma methodology could systematically improve production quality and reduce cost in the manufacturing shop-floor (Sameer and Michael, 2009).

The application of the lean six sigma concept by using the DMAIC model could also provide a dramatic improvement in the manufacturing process (Dario et al., 2008; Pusporini et al., 2013)

Multiple Regression Analysis

Cohen (1982) developed a generalization of multiple regression and multiple correlation and later discussed by the work of Cohen et al. (2003). Set correlation is a multivariate generalization of multiple regression and estimates the amount of variance shared between two sets of variables. Set correlation also allows for examining the relationship between two sets when controlling for a third set. Set correlation is

$$R^2 = 1 - \prod_{i=1}^{n}(1 - \lambda_i)$$

where λ_i is the ith eigen value of the eigen value decomposition of the matrix

$$R = R_{xx}^{-1} R_{xy} R_{xxR}^{-1} R_{xy}^{-1}$$

Unfortunately, there are several cases where set correlation will give results that are much too high. This will happen if some variables from the first set are highly related to those in the second set, even though most are not. In this case, although the set correlation can be very high, the degree of relationship between the sets is not as high. In this case, an alternative statistic, based upon the average canonical correlation might be more appropriate.

Problems in Time Series Multiple Regression

There are four plausible statistical problems that may occur in time series multiple regression. They are Multicollinearity, Heteroscedasticy, Autocorrelation.

Multicollinearity

Multicollinearity can occur when two predictors are highly correlated with each other. It can also occur when a linear combination of predictors is highly correlated with another linear combination of predictors. There are two major statistical outputs associated with the data not a model itself.

a. a high coefficient of determination (R^2) with low values for the t statistics,
b. high values for simple correlation coefficients (close to but equal to + 1 or -1) between the independent variables, and

To remove multicollinearity, any one of the following techniques can be used.

a. combine variables,
b. eliminate one of the variables or, possibly,
c. increase the sample size.

Heteroscedasticy

The error terms are assumed to have a constant variance, but if the variance depends upon the value of X, then heteroscedasticy exists. The Goldfeld-Quandt test for heteroscedasticy may be used. For this correction, any one of the following methods are required.

a. weighted least squares,
b. transformed variables,
c. respecification of the model.

Autocorrelation

Error terms are assumed to be independent over time. Autocorrelated error terms are the most important consideration for a practicing forecaster. The Durbin-Watson statistic can be used to test for autocorrelation problem that include following statistical procedures.

a. identifying an important omitted variable
b. transforming the variables based upon generalized least squares and
c. introducing a lagged value of the dependent variable on the right-hand side of the equation.

Method of Checking Multicollinearity

The paper considers a variance inflation factors (VIF) to check multicollinearity that enabled us to eliminate the problem and select the "best" predictor variable to enter when other independent variables are present.

Consider the following linear model with I independent variables:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots\ldots + \beta_i X_i + \varepsilon$$

If the predictors are correlated with each other, the standard errors of the coefficient estimates will be bigger than if the predictors were uncorrelated. If the predictors were uncorrelated, the variance of $\widehat{\beta}_\iota$ would be

$$Var[\widehat{\beta}_i] = \frac{\sigma^2}{(n-1)\,Var(X_i)}$$

In the case of perfect correlation (i.e., a correlation of +1 or -1), it is not possible to estimate the regression model.

If there is high correlation (close to but not equal to +1 or -1), then the estimation of the regression coefficients is computationally difficult. The computer software, for example, R/SPSS/SAS/Stata use algorithms to limit the effect of multicollinearity on the coefficient estimates.

When multicollinearity is present, the uncertainty associated with individual regression coefficients will be large. This is because they are difficult to estimate. Consequently, statistical tests (e.g., t-tests) on regression coefficients are unreliable. (In forecasting we are rarely interested in such tests.) Also, it will not be possible to make accurate statements about the contribution of each separate predictor to the forecast (Kim, 2019).

Reliability Check: Cronbach's Alpha

A scale is frequently used in survey questions to grade responses; one such example is the measurement of customer satisfaction on a scale of 1 to 5, where 1 represents low customer satisfaction and 5 great customer happiness. These measures are referred to as Likert scales which were created in 1932 to quantify respondents' sentiments about particular situations in a way that was accepted and validated by science (Joshi et al. 2015). The Likert method is known as a summated rating scale because the responses from each item are added up (or averaged) to determine the respondent's score on the scale (Vaske et al. 2017).

Cronbach's alpha is a statistic that encapsulates an index's reliability (Cronbach, 1951). From zero (no internal consistency) to unity (perfect internal consistency), alpha is a metric used to assess the interval consistency of a set of items. For simplicity's sake, we will not offer the standard calculating formula for alpha.

However, if the items that will be added together to produce an index have equal or nearly equal variances, as they will if all of the items utilize the same format for response categories, alpha can be reduced to a relatively straightforward calculation (and possibly divided by the total number of items to get an average response across items). If all elements are converted to Z scores before summing, this formula may also be employed. It is

$$\alpha = \frac{k\bar{r}}{1+(k-1)\bar{r}}$$

where

k = the number of indicators in the index

$\bar{r}$ = the average intercorrelation among the k items comprising the index

Cronbach's alpha measures the extent to which item responses (answers to survey questions) correlate with each other. In other words, Cronbach's alpha estimates the proportion of variance that is systematic or consistent in a set of survey responses. The statistic alpha "typically" ranges from 0.00 to 1.00, but a negative alpha value can occur when the items are not positively correlated among themselves, (Joshi et al 2015.)

The final Cronbach's alpha value is a function of the number of items in the scale, and although alpha is a function of the item inter-correlation, it needs to be interpreted with the number of items in mind; it is accepted that an ideal Likert scale will have a Cronbach's alpha of greater than 0.70, (Cortina 1993), however, the more items utilised in the scale the less accurate as a measure of reliability and inter-correlation the Cronbach's alpha becomes. For example, with 20 items on the scale, the alpha can still be greater than 0.70 but have a small correlation coefficient between items. Thus Cortina (1993) recommended using the Cronbach's alpha carefully when designing complex, multidimensional multi-item time series panel data analysis.

Conclusion

The paper described and discussed a best method of time series data for interpretation. For this purpose, philosophy of quantitative is positivism, which is further elaborated through a lean six sigma as a DMAIC methodology. To generalize the findings of Time Series Panel Data in the process of generalization, Cronbach's alpha, Hypothetico-deductive model, Lean six sigma method, and Multiple regression analysis are the basic significant statistical attributes.

References

Aggogeri, F., & Mazzola, M. (2008). Combining six sigma with lean production to increase the performance level of a manufacturing system. In *Proc. ASME International Mechanical Engineering Congress and Exposition*. Massachusetts.

Chong, S. W., & Reinders, H. (2021). A methodological review of qualitative research

syntheses in CALL: The state-of-the-art. *System*, *103*, 102646.

Cohen J (1982). Set correlation as a general multivariate data-analytic method. *Multivariate Behav Res.* 17(3), 301-41. doi: 10.1207/s15327906mbr1703_2. PMID: 26800754.

Cortina, J. (1993). What is coefficient alpha? An Examination of theory and applications. *Journal of Applied Psychology*, 78 (1), 98-104.

Creswell, J. W. (2014). *Research design qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), pp.297-334.

Cudney, E. A., Mehta, M., & Monroe, R. (2006). Combining lean and six sigmas for optimal result. In *Proc. the SME Summit and Annual Meeting*. California.

Dario, P., Deisell, M. D., & Howard, G. (2008). A lean six sigma case study: An application of the "5s" techniques. *Journal of Advances in Management Research*, 5 (1), 63.

Denzin, N. K., & Lincoln, Y. S. (2005). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (pp. 1–32). Sage Publications Ltd.

Frankfort-Nachmias, C., & Nachmias, D. (2008). Research methods in the social sciences (7th Edition). New York.

Furterer, S & Elshennawy, A. K. (2005). Implementation of TQM and lean six sigma tools in local government: A framework and a case stud. *Total Quality Management and Business Excellence*, 16 (10), 1179-1191.

Hsiao, C. (2005). Why panel data? *The Singapore Economic Review*, *50*(02), 143-154

Jeyaraman, K., & and Kee,T. L. (2010). A conceptual framework for critical success factors of lean Six Sigma. *International Journal of Lean Six Sigma*, 1 (3), 191

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert scale: Explored and explained. British

Journal of Applied Science & Technology 7(4), 396.

Kim J.H. (2019). Multicollinearity and misleading statistical results. Korean J Anesthesiol, 72(6):558-569. doi: 10.4097/kja.19087.

Kumar, M., Antony, J., Singh, R. K., Tiwari, M. K., & Perry, D. (2006). Implementing the lean sigma framework in an Indian SME: A case study. *Production Planning and Control*, 17 (4), 407-423.

Montgomery, D. C., & Woodall, W. H. (2008). An overview of six sigma. *International Statistical Review*, 76 (3), 329-346.

Munn, Z., Stern, C., Aromataris, E., Lockwood, C., & Jordan, Z. (2018). What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC medical research methodology*, *18*(1), 1-9.

Pepper, M. P. J., & T. A. Spedding, T. A. (2010). The evolution of lean six sigma. *International Journal of Quality and Reliability Management*, 27(2), 138-155.

Pusporini, P., Abhary, K., & Luong, L. (2013). Development of environmental performance model using design for six sigma (DFSS). *International Journal of Materials, Mechanics and Manufacturing*, 1 (1).

Sameer, K., & Michael, S. (2009), Using DMAIC six sigma to systematically improve shopfloor production quality and costs. *International Journal of Productivity and Performance Management*, 58 (3), 254.

Snee, R. D. (2010). Lean six sigma-getting better all the time. *International Journal of Lean Six Sigma*, 1 (1), 9-29

Souraj, S., Abdur, R., & Juan, A. C. (2010). The integration of six sigma and lean management. *International Journal of Lean Six Sigma*, 1 (3), 249

Su, C. T., Chiang, T. L., & Chang, C. M. (2006). Improving service quality by capitalising on

an integrated lean six sigma methodology. *International Journal Six Sigma and Competitive Advantage*, 2 (1), 1-22.

Vaske, J., Beaman J., & Sponarski, C. (2017). Rethinking internal consistency in Cronbach's Alpha. *Leisure Sciences*, 39(2), 163-173