Economic Analysis and Policy 55 (2017) 35–46

Contents lists available at ScienceDirect

Economic Analysis and Policy

journal homepage: www.elsevier.com/locate/eap

Full length article

Assessing consistency of consumer confidence data using latent class analysis with time factor

Sunil Kumar[a,*], Zakir Husain[b], Diganta Mukherjee[c]

[a] *Alliance University, Bangalore, India*

[b] *Indian Institute of Technology, Kharagpur, India*

[c] *Sampling and Official Statistics Unit, Indian Statistical Institute, Kolkata, India*

ARTICLE INFO

Article history:
Received 16 September 2016
Received in revised form 28 February 2017
Accepted 12 April 2017
Available online 19 April 2017

JEL classification:
C32
E31
E37

Keywords:
Latent class analysis
Reliability analysis
Consumer confidence survey
India

ABSTRACT

In many countries information on expectations collected through consumer confidence surveys are used in macroeconomic policy formulation. Unfortunately, before doing so, the consistency of responses is often not taken into account, leading to biases creeping in and, in turn, affecting the consistency of the indices hence created. This paper describes how latent class analysis may be used to check the consistency of responses and ensure parsimony in the questionnaire. In particular, we examine how temporal changes may be incorporated into the model. Our methodology is illustrated using three rounds of Consumer Confidence Survey (CCS) conducted by Reserve Bank of India (RBI).

1. Introduction

The effect of attitudes of non-financial organizations and consumers on economic activity is a subject of great interest to both policymakers and economic forecasters (see Galstyan and Movsisyan, 2009). In particular, expectations about macroeconomic variables may influence the effectiveness of monetary and fiscal policy and also the direction of the economy (see Phelps, 1968). This underlines the importance of incorporating such expectations in the process of formulation of macroeconomic policy. Information about such expectations is collected by policy makers in many countries through consumer confidence surveys. Examples of such surveys are the European Commission programme in the EU, University of Michigan survey of Consumers, Business Tendency and Consumer Surveys of the Organization for Economic Co-operation and Development in Armenia, The Nielsen Global Survey of Consumer Confidence and Spending Intentions in Hong Kong, Westpac-Melbourne Institute Survey of Consumer Sentiment, etc. Based on such surveys, consumer confidence indices are constructed and often used in policy making as a proxy for consumers' expectations about the future trend of the economy.

The consumer and industrial confidence indicators have been used in many studies mostly as a forecasting tool (Bialowolski, 2014; Golinelli and Parigi, 2004, among others), there has been almost no debate on their meaning and consistency. Apart from forecasting, there are only minor exceptions in which the confidence indicators have been analyzed in connection with their components (see Jansen and Nahuis, 2003; Ramalho et al., 2011). The choice of questions for

* Corresponding author.
E-mail addresses: sunilbhougal06@gmail.com (S. Kumar), dzhusain@gmail.com (Z. Husain), digantam@hotmail.com (D. Mukherjee).

http://dx.doi.org/10.1016/j.eap.2017.04.004
0313-5926/

confidence indicators lack proper justification. Mueller (1963, p. 901) states only that the questions serving as items in the American consumer sentiment indicator were chosen because "they relate to different important aspects of consumer sentiment and because they have been asked at least since 1952". Nevertheless, some general guidelines concerning the indicator of confidence can be found in the literature. Golinelli and Parigi (2004) claim that, given the rational expectations hypothesis, the indicator has to have additional information if it is defined as an expected value of macroeconomic variables. They also mention that consumer sentiment (confidence) is a more general concept that cannot be summarized only on the basis of some macroeconomic variables. Vuchelen (2004) states that consumer sentiment is not well understood and there is no agreement on its information content. Additionally, he adds that confidence might be associated with unobserved variables.

In India, information about the households' perceptions about the current economic situation and their expectations about future economic changes is collected on a quarterly basis by Reserve Bank of India since June 2010 through the Consumer Confidence Surveys (CCSs). The responses to these surveys are analyzed to obtain pictures of households' opinions on the overall economic situation (current and future) and their material security (current and future). In addition, the RBI calculates indices of current and future conditions using the formula:

$$\text{Overall Index} = \text{Average} * (100 + \text{Net response of selected factors}),$$

where Net response $=$ Percentage of positive responses $-$ Percentage of Negative responses.

The average net responses on current perceptions on factors like economic conditions, household circumstances, income, spending and price level are used to calculate the Current Situation Index, while average net responses on future perceptions on economic conditions, income, spending, price level and employment are used for the calculation of the Future Expectations Index (RBI Bulletin, 2012). The survey data are a potentially useful tool to monitor temporal changes in households' expectations. Accordingly, the survey results are used by the RBI to formulate monetary policy and determine key monetary variables like Cash Reserve Ratio, interest rates, etc.

Unfortunately, while constructing such indices, the consistency of responses is often not taken into account (see Katona, 1947, p. 459). This leads to overlooking of biases in consumer responses, and may even affect the consistency of the indices. The RBI, too, uses the survey response to form monetary policy variables without examining the consistency of responses, resulting in either under (over)-estimation of the true level of confidence. Depending upon the magnitude of inconsistent responses, this can lead to a substantial level of mis-reporting of consumer confidence and, if followed by policy response, can lead to measures that under (over)-boost consumer confidence.

This paper describes how latent class analysis (LCA) – a latent variable model with discrete latent and indicator variables – may be used to check the consistency of responses and identify the variables that may be used to construct a relatively precise index of consumer sentiment about the economy. Our methodology is illustrated using the CCS conducted by RBI. The usual candidate models for such analysis uses the concepts of Item response theory. In the present exercise, the data structure is discrete. Hence, any Item response theoretic modeling will have to be based on discrete conditional probabilities. This is achieved using LCA.

The paper is structured as follows. Section 2 describes the framework of the LCA model. Sections 3 and 4 present the analysis of three rounds of CCS data using the software application poLCA.[1] Finally, in Sections 5 and 6, we summarize our findings and identify potential areas for further research.

2. LCA models and response biases

2.1. LCA models

LCA is a statistical method for clustering the related cases (identifying latent classes) from multivariate categorical data, pioneered by Lazarsfeld (1950) and Lazarsfeld and Henry (1968). LCA models do not rely on traditional modeling assumptions like normal distribution, linear relationship, homogeneity etc. The conventional classification methods like Cluster analysis and Factor analysis implicitly or explicitly requires the data to be continuous and sometimes normally distributed. This is not true in the present situations. Both the latent and indicator variables of this model are discrete. So discrete data techniques are necessary. LCA is a subset of structural equation modeling, used to classify unobservable sub-groups or subtypes of cases in multivariate categorical data. As in factor analysis, the LCA can also be used to group cases according to their maximum likelihood class membership. LCA may be applied to classify types of attitude structures from survey responses, consumer segments from demographic and preference variables, or categorize subpopulations based on their responses to test items. LCA has the advantage of making no assumptions about the distribution of the indicators other than that of local independence which says that the indicators share a common latent variable, but the errors in measurement are uncorrelated. Subsequent development allows this assumption to be relaxed. Even when we have dependence between indicators, it can be incorporated via interaction term among the indicator variables (e.g. Harper, 1972; Vacek, 1985; Hagenaars, 1988; Espeland and Handelman, 1989; Sinclair Michael and Gastwirth, 1996; Reboussin et al., 2008; Bertrand and Hafner, 2011).

[1] The analysis is based on the most complete and most user-friendly package for the estimation of latent class models and latent class regression models in R (see Linzer and Lewis, 2011; R Development Core Team, 2011).