Chapter 11
Qualitative Response Regression Modeling

Aliyu Olayemi Abdullateef
University Utara, Malaysia

ABSTRACT

In most regression models, readers have implicitly assumed that the dependent variable (regressand) Y is quantitative. On the contrary, explanatory variables could take the form of qualitative (or dummy), quantitative, or a triangulation thereof. This chapter discusses the observed fundamental differences between quantitative and qualitative models through a clear definition of their individual objectives. This chapter also considers many models in which the regressand is a qualitative variable, popularly called categorical variables, indicator variables, dummy variables, or qualitative variables. This chapter shows why it is not compulsory to restrict our dependent variable to dichotomous (yes/no) categories by establishing inherent benefits in estimating and interpreting trichotomous or polychotomous multiple category response variable. Relevant examples for developing, analyzing, and interpreting a probability model for a binary response variable using three known approaches (i.e. linear probability model, logit, and probit models) is also discussed.

LEARNING OBJECTIVES

Understand the differences between quantitative and qualitative regression models.

Understand how to develop, estimate and interpret qualitative response regression models.

- Comprehend the concept of linear probability models (LPM)? And what are its fundamental problems.
- Discuss alternatives to linear probability models.
- Does the conventionally computed R^2 have any value in qualitative response regression models?
- Beyond the dichotomous (yes/no) regressand variable, how can researchers estimate and interpret polychotomous multiple regression models?

DOI: 10.4018/978-1-4666-6371-8.ch011

QUALITATIVE AND QUANTITATIVE REGRESSION MODELS

As evident in many studies, researchers have used participatory tools in gathering information that are qualitative in nature. Some of the outcome in this type of research will help in addressing stated research questions, while others will provide general information in understanding a particular phenomenon. It is important to emphasize that quantitative method of modeling data analysis can be very important to researchers that is attempting to achieve reliable results that can be generalized from a qualitative data that is surveyed from large population.

Larger percentage of existing data analyses have been using endogenous and exogenous variables that are measured on a quantitative or continuous scales. In this chapter, our major focus is to discuss and analyze regression models that use constructs that are measured on qualitative or discrete scales. A major benefit of this approach is that it avails researchers the ability to separate a larger percentage of confounding factors that can potentially obscure findings in qualitative regression models. Suppose we want to conduct a study whose primary objective is to determine the impact of government loans on the performance of rural entrepreneur. An interactive interview with a focus group among the local entrepreneurs will provide valuable information for the research, but other complex factors such as available markets for the products, access to good road, profitability etc. will require certain degree of quantitative data analysis. Having required skills to separate quantitative components from a data will strongly assist researchers in identifying unique characteristics of qualitative data.

The structure of the data is another important factor to consider in analyzing and achieving reliable results in a qualitative data analyzed with quantitative methods. To do this, researchers should create special attention to data categorization that best fits different methods of collection. Thinking on how to structure our data will force researchers to concentrate on good methodologies to undertake in collecting qualitative information, thus helping in eliminating qualitative components that cannot be coded.

A major argument that we will like to establish in this chapter is that we don't have to restrict our selected response variables to dichotomous yes/no categories only. Suppose we want to embark on a study of Nigerian general election, where three parties have participated i.e. PDP, ACN and APGA. With a three parties scenario we will now have a response variable called trichotomous or multiple category polychotomous response variable.

QUALITATIVE AND QUANTITATIVE DUMMY MODELS

Dependent variables in regression analysis is not only influenced by ratio scale constructs (e.g. temperature, height, income etc.) but can also be influenced by variables that are essentially qualitative in nature i.e. nominal scale such as religion, party affiliation, race, sex etc. A good practical example is workers in developed countries are found to earn more than workers in developing countries or educated citizens are found to earn more than the uneducated citizens. Arguably, this type of phenomena may result in discrimination based on nationality or level of education, hence qualitative variables such as nationality and level of education that can influence the regressand should be included among likely regressors.

Another strong indication that supports this assertion is that they are nominal scales that have the ability to establish the presence or absence of an attribute or quality. In Gujarati (2003), he argued that a unique way through which researchers can quantify these attributes is by constructing artificial variables that takes values of 1 or 0, where 1 will indicate the presence of an attribute and 0 will indicate the absence of such attribute. Suppose will want to take 1 as the white Americans