**Chapter 5: Nonlinearity and Nonadditivity**

Preview Questions

1. What is nonlinearity and nonadditivity?
2. How do you detect nonlinearity and nonadditivity?
3. How do you deal with nonlinearity?
4. How do you deal with nonadditivity?
5. What should we be cautious about regarding nonlinear and nonadditive specifications?

Reading Summary

Introduction

* Linearity and additivity are implicit assumptions in the regression model.
* Linearity is the assumption that for each independent variable Xi the amount of change in associated with a unit increase in Xi while holding all other independent variables constant is the same regardless of the level of Xi.
* Additivity is the assumption that for each independent variable Xi the amount of change in associated with a unit increase in Xi while holding all other independent variables constant is the same regardless of the other independent variables in the equation.
  + Don’t have to specify at which constant values the other independent variables are held.
* With nonlinearity and nonadditivity the relationship between the dependent variable and independent variables varies according to context.
  + Nonlinearity the relationship varies with the value of the independent variable.
  + With nonadditivity the relationship varies with the value of other independent variables.
* Many nonlinear and nonadditive equations can be converted to linear and additive form by performing transformations on the variables.
  + Nonlinear and nonadditive in terms of the variables but linear and additive in terms of the parameters.

Detecting Nonlinearity and Nonadditivity

* General process:
  + Use the theory underlying the model to hypothesize about the form of the nonlinearity and nonadditivity relationship
  + Specify the model reflecting this form
  + Estimate the parameters of the model
  + Use the statistics for the regression to evaluate the hypothesis
* Questions for evaluating nonlinearity
  + Can the slope of the relationship be expected to have the same sign for all values of the independent variable?
  + At what level of the independent variable can the sign be expected to change?
  + Should we expect the magnitude of the slope to increase or decrease?
* Similar questions can be asked about nonadditivity
* There are tests to detect nonlinearity and nonadditivity even when the precise nature of the relationship can’t be ascertained beforehand.
  + Scatter plot of dependent variable against the independent variable for visual inspection.
  + Divide the observations into subsamples, perform regression on each subsample, and compare slopes of the variables.
    - For sample divided in ns samples
      * ns – 1 dichotomous variables must be created
      * 2(ns – 1) terms included as independent variables in the test regression along with X
      * Easy to reject null hypothesis of linearity if ns is large.
      * If ns is small relative to n it is extremely difficult to reject the null hypothesis of linearity
  + Conduct a single regression using dummy variables.

Dealing with Nonlinearity

* There are a number of nonlinear specifications that are linear in terms of parameters where Ordinary Least Squares (OLS) regression can be used once a transformation is applied.
* Polynomial model
  + Appropriate where the slope of the relationship between an independent variable and dependent variable is believed to change sign as the value of the independent variable increases.
  + Y is viewed as a function of Xi and one or more powers of Xi
  + See Eq. 5.8 on p. 57
  + Number of bends almost always equals m-1
    - Point at which the slope equals zero and changes sign
  + Define the powers of Xi as distinct variables (e.g., X2j = , X3j = , etc.)
  + Each independent variable defined as functions of a single conceptual variable.
    - No perfect collinearity because variables are not linearly related.
    - Could still be highly correlated.
    - Check for multicollinearity.
  + Can include a second conceptual variable that is assumed to be linearly related to the dependent variable.
  + Typical interpretation of slopes not applicable for polynomial model.
    - Interpret regression coefficients in terms of the slope of the relationship and how it changes over key ranges in the value of the conceptual independent variable.
* Exponential model
  + Appropriate when the sloped of the relationship never changes sign but increase or decreases in magnitude as the value of X changes.
  + See Eq. 5.14 on p. 60
    - β > 1 🡪 slope increases as X increases
    - β < 1 🡪 slope decreases as X increases
    - Every one percent increase in X is associated with a β percent change in the expected value of Y.
  + Transform by taking the logarithm of both sides the perform regression
  + Note that error term is multiplicative, not additive.
    - Log of error has a mean zero rather than the value of the error have a mean zero.
* Hyperbolic model
  + Also called the reciprocal model
  + See Eq. 5.18 on p. 63
  + Y approaches α as X gets infinitely large
  + β > 0 🡪 Y approaches α from below
  + β < 0 🡪 Y approaches α from above
  + Y is always greater than α
  + Slope β eventually approaches zero
* e(α+βX+ε) exponential model
  + Y-intercept of eα
  + β < 0 🡪 curve has negative slope throughout
    - Slope β decreases in magnitude as X increases
    - Curve approaches X axis as Y gets infinitely large
    - Ratio of Y’ and Y’’ equals a constant value for any two values of X’ and X’’
    - If X’-X’’=1, then the ratio of Y’ and Y’’ equals eβ
  + Transform by taking the natural logarithm of both sides.

Dealing with Nonadditivity

* Independent variables interact in influencing the dependent variable
* Dummy variable interactive model
  + One of the independent variables is dichotomous
  + Estimate coefficients through contextual regression analysis
    - Develop two regression equations
* Multiplicative model
  + Two independent variables are both measured at the interval level
  + See Eq. 5.34 on p. 67
  + Y-intercept represents the expected value of Y when both independent variables equal zero
  + β3 equals the amount of change in the slope of the relationship between X1 and Y associated with a unit increase in X2 (or the other way around with the independent variables)
  + Create a new variable to equal to the multiplicative and perform regression
    - See Eq. 5.40 on p. 68
    - Perfect collinearity is impossible
    - High degree of multicollinearity is possible; test for multicollinearity
  + Interpretation of coefficients
    - β1 and β2 are estimates of conditional effects
      * Estimates of the change in Y associated with a unit increase in one independent variable under the condition that the other independent variable is equal to zero.
      * Conditional effects can be obtained for any fixed value of either independent variable.
    - Standard errors are also conditional.
      * Variability of β1 when X2 is fixed at zero
      * Variability of β2 when X1 is fixed at zero
      * Variability estimates will be different when the independent variables are fixed at nonzero numbers.
      * Significance tests must be conditional also.
      * Determine whether an independent variable has a significant effect on the dependent variable at a particular value of the other independent variable.
* Nonlinear interactive model
  + Both nonadditive and nonlinear
  + Modify exponential model to include a second independent variable.
  + Slopes of both relationships depend on the level of both independent variables.
  + Coefficients determined by taking the logarithm of both sides of the equation.
* All nonadditive models can be expanded to include more than two conceptual models.

Some Warnings about Nonlinear and Nonadditive Specifications

* Models may contain different types of error terms
  + Additive error term
  + Multiplicative disturbance term
  + Consider whether assumptions about the error term are consistent with the situation
* Ensure that the transformation is defined for all possible values of the variable
  + Measurement scale may need to be converted
  + Not appropriate if variable has a ratio scale
* Do NOT reject a nonlinear or interactive model in favor of a linear additive model simply based on a comparison of goodness-of-fit
  + Unit of comparison across models must be the same