Robust Model Selection: A Review

Jennifer L. Castle and David F. Hendry*
Economics Department and Institute for New Economic Thinking
at the Oxford Martin School, University of Oxford, UK

May 2016

Abstract

Complete and correct *a priori* specifications of models for observational data never exist, so model selection is unavoidable. The target of selection needs to be the process generating the data for the variables under analysis, while embedding the objective of the study, often a theory-based formulation. This requires starting from a sufficiently general initial specification that comprises all candidate variables, their lags in time-series data, and functional forms, allowing for possible outliers and shifts, seeking parsimonious final representations that retain the relevant information, are well specified, encompass alternative models, and evaluate the validity of the objective. Intrinsically, we seek robustness against many potential problems jointly: outliers, shifts, omitted variables, incorrect distributional shape, non-stationarity, mis-specified dynamics, and non-linearity, as well as inappropriate exogeneity assumptions. Our approach inevitably leads to more variables than observations, tackled by iteratively switching between contracting and expanding multi-path searches programmed in *Autometrics*. The paper explains the steps involved, based on *Empirical Model Discovery and Theory Evaluation* (Hendry and Doornik, 2014), specifically addressing indicator saturation to discriminate between outliers and large observations arising from non-linear responses. The analysis is illustrated using artifical data (outliers versus non-linearity), by empirical models of food demand in the USA (tackling outliers) and real wages in the UK (tackling non-linearities and location shifts), as well as by volcanic impacts on Northern-hemisphere temperature reconstructions (designed shift functions).

JEL classifications: C51, C22.
KEYWORDS: Model Selection; Robustness; Outliers; Location Shifts; Indicator Saturation; *Autometrics*.