Empirical Model Discovery

Sir David F. HENDRY

Nuffield College & Economics Department, Oxford, UK

Abstract

The talk summarizes a great deal of recent research, and explains how it facilitates the discovery of empirical models, greatly reducing the risks from model mis-specification and data contamination. Model evaluation concerns discovering what is wrong; robust statistics as discovering which sub-sample is reliable; non-parametric methods as discovering the functional form; and model selection as discovering which model best matches the given criteria. However, the high dimensionality, non-linearity, inertia, endogeneity, evolution, and abrupt change characteristic of economic data, which interact to make empirical modelling difficult and pose substantive risks of ending with an incorrect representation, make it essential to ***tackle all of these jointly***. Automatic methods enable formulation, selection, estimation, and evaluation on a scale well beyond the powers of humans alone, including when there are more candidate variables than observations, while allowing theory models to be embedded in the discovery process. Live computer illustrations using *Autometrics* show the remarkable power and feasibility of this exciting approach.