Competing Models[*]

José Luis Montiel Olea[†], Pietro Ortoleva[‡], Mallesh Pai[§], Andrea Prat[¶]

First version: October 2017
This version: March 29, 2019

Abstract

Different agents compete to predict a variable of interest related to a set of observables via an unknown data generating process. All agents are Bayesian, but have different priors over this data generating process: they may consider different subsets of observables to make their prediction. After observing a common dataset, who has the highest confidence in her predictive ability? We characterize it and show that it crucially depends on the size of the dataset. With *small data*, typically it is an agent using a model that is 'small-dimensional,' in the sense of considering few observables, even when the true data generating process involves many. With *big data*, it is instead typically 'large-dimensional,' possibly using more variables than the true data generating process. These features are reminiscent of model selection techniques used in statistics and machine learning. Crucially, here they emerge not normatively, but positively as the outcome of competition between Bayesian agents. The theory is applied to auctions of assets where bidders observe the same information but hold different priors.

Keywords: Bayesian, competition, model selection, Akaike Information Criterion.

JEL: D81, C11, C52.