

Title: **Smoothers and the Cost of Model Selection**

Author(s): **Bradley Efron**

Technical Report number (Dept. of Statistics, Stanford Univ.): **2000-23B/209**

Date: **April 2000**

Abstract:

This paper concerns a particularly simple form of model selection, adaptive scatterplot smoothing. Two main questions are considered: How should the selection be done, and how much does it cost, compared to knowing beforehand the ideal smoothing parameter? For spline-like smoothers it is possible to derive simple formulas that give reasonably accurate small-sample answers. The most popular selection method, known in different contexts as C_p , AIC or GCV, turns out to be surprisingly vulnerable to a type of estimation instability. A competing method, GML, is much more stable but can be badly biased.